This is to certify that the
dissertation entitled

IDENTIFICATION OF NOVEL GENES INVOLVED IN ZEBRA MUSSEL (DREISSENA POLYMORPHA) UNDERWATER ADHESION MECHANISM

presented by

WEI XU

has been accepted towards fulfillment
of the requirements for the

Doctoral degree in Pathology

Mohamed Faisal
Major Professor's Signature

12/16/2009
Date

IDENTIFICATION OF NOVEL GENES INVOLVED IN ZEBRA MUSSEL (*DREISSENA POLYMORPHA*) UNDERWATER ADHESION MECHANISM

By

Wei Xu

A DISSERTATION

Submitted to
Michigan State University
In partial fulfillment of the requirements
For the degree of

DOCTOR OF PHILOSOPHY

Pathology

2009

ABSTRACT

IDENTIFICATION OF NOVEL GENES INVOLVED IN ZEBRA MUSSEL (DREISSENA POLYMORPHA) UNDERWATER ADHESION MECHANISM

By

Wei Xu

Zebra mussels (*Dreissena polymorpha*) have invaded North America causing economic and ecologic devastation. The zebra mussel attaches firmly to underwater substrates through a complex system of exocrine glands, byssal threads, and adhesive plaques. In this study, tools were developed and experiments were designed in order to better understand the zebra mussel underwater adhesion mechanisms. A normalized cDNA library of the zebra mussel byssus was constructed with the subtractive suppression hybridization (SSH) technique. 750 non-redundant expressed sequence tags (ESTs) were obtained from the library. A cDNA microarray was developed with the PCR products of 716 ESTs selected from the cDNA library. The newly developed cDNA microarray was successfully used to compare between two groups of mussels with different byssogenenic activities. Between the two groups, 16 genes were differentially expressed with statistical significance. The results were validated by the quantitative PCR (qPCR). To follow up on genes that are either up or downregulated along the course of byssogenesis, a microarray time-course experiment was designed. Samples were collected at seven time intervals; 12 hours, 1 day, 2 days, 3 days, 4 days, 7 days, and 21 days after severing the byssal threads in the treatment group. The numbers of differentially expressed genes identified at these

time points were 13, 13, 20, 17, 16, 20, and 29 respectively. An additional experiment was designed to identify the differentially expressed genes in response to the changes of byssogenesis status and environmental factors, including temperature, current velocity, and dissolved oxygen levels, as well as the status of byssogenesis on the expression of foot-unique genes. The expression profiles of 18 genes were found to be altered by two experimental factors, while 117 genes had differentially expressed profiles in response to only one experimental factor. The numbers of the genes modulated by byssogenesis status, temperature, dissolved oxygen, and current velocity were 59, 27, 26, and 9, respectively. Seven genes identified by the time-course experiment and four genes identified by factorial analysis assay were validated by qPCR with the results consistent to microarray results. Detected by RNA fluorescent in situ hybridization (FISH), both genes BG20_A01 and BG97/192_B06 were found to be expressed within the thread-forming glandular cells. The genes BG15_F03 and BG16_H05 were expressed by stem-forming gland cells and plaque-forming gland cells, respectively. A full length gene in the microarray, homologous to insect defensin A, was cloned from the zebra mussel foot. The analysis of the *D. polymorpha* defensin (Dpd) suggested that the Dpd is homologous to the insect defensin A. The expression of Dpd in hemocytes was found to be inducible by stimulation with lipopolysaccharides (LPS), peptidoglycan (PGN), and zymosan (ZYM) using qPCR. The mature recombinant Dpd showed inhibition activity against four strains of Gram-negative bacteria and one Gram-positive bacterium. Dpd was present not only in the foot, but also in a number of other tissues. Interestingly, expression levels of Dpd during the early stage of byssogenesis were mostly higher than the non-byssogenesis status, which suggested that the activities of Dpd were associated with byssogenesis.

DEDICATION

This dissertation is dedicated to my wife, my parents, and my dearest friends.

ACKNOWLEDGEMENTS

It is a pleasure to thank those who made this thesis possible.

Foremost, I would like to express my sincere gratitude to my major advisor Prof. Mohamed Faisal for the motivation, enthusiasm and inspiration, which was always there when I needed it. This thesis would not have been possible without his sincere guidance and encouragement.

I would also like to thank the rest of my thesis committee: Prof. Paul Coussens, Prof. Matti Kiupel, Prof. Orlando Sarnelle, and Prof. Robert Bowker, for their greatest support during my Ph.D. study and invaluable advice.

Special appreciation goes to Prof. Robert Tempelman who gave great help in experimental design and microarray data analyses.

I am indebted to my many of my colleagues for a stimulating and fun environment in which to learn and grow, especially grateful to Thomas Loch, Michelle Gunn, Carolyn Schultz, Robert Kim, Elena Millard, and Drew Winters.

I must thank the current and former members in the Center for Animal Functional Genomics, Sue Sipkovsky, Christopher Colvin, Xiaoning Ren, and Robert Halgren for their effort in developing the microarray.

I would like to extend my appreciation to Michigan Sea Grant College Program, from National Sea Grant, NOAA, U.S. Department of Commerce, and Dissertation Completion Fellowship from MSU Graduate School.

Last, but not least, I wish to thank my wife, Qianqian, for her understanding and love during the past few years. Her support and encouragement was also the source of my

strength and motivation. My parents, Jianyun and Ruiyi, receive my deepest gratitude and love for their dedication and the many years of support through my candidature.

Preface

Native to Russian waters, the zebra mussel (*Dreissena polymorpha* Pallas 1771) was first introduced to the Great Lakes in the late 1980s. The non-native species has colonized an alarming number of North American inland waterways and lakes (Benson & Raikow 2009). Zebra mussels reproduce and spread rapidly, leaving behind a trail of habitat destruction, harmful algal blooms, loss of native species, and economic devastation. Developing an effective strategy to eradicate zebra mussels, or to at least control their spread, remains a remote possibility despite concerted efforts by scientists and managers.

Why are zebra mussels so abundant when most bivalve mollusks populations worldwide are either declining or at the brink of extinction? One likely reason is the zebra mussel's extraordinary ability to attach tightly and expeditiously to any underwater hard surface including adults of their own species (Johnson & Padilla 1996). Additionally, optimal environmental conditions, lack of competition, and scarcity of natural predators or pathogens may have also contributed to the explosive increase in zebra mussel populations in the Great Lakes.

Zebra mussel invasions have severe impacts on the health of the ecosystem and the local economy. Industrial facilities devote significant efforts to the cleaning of water intake pipes and heat exchangers clogged with zebra mussels. Water circuits in the facilities must be cleaned manually, where possible, or with various chemicals in processes that are both expensive and environmentally harmful (Cope *et al.* 1997). The profound economic impact of zebra mussel expansion in the Great Lakes basin alone has mounted billions of dollars in increased operating costs of affected industries and

municipalities over decades (Benson & Raikow 2009). This situation has overwhelmed North American scientists and water resource managers who unexpectedly became confronted with an ecologic catastrophe of an unprecedented magnitude. The major obstacle that impedes immediate intervention or development of an effective management strategy is the dearth of knowledge on the basic biology of zebra mussels and the survival strategies used by this successful intruder.

The zebra mussel is unique in many biological properties. For example, zebra mussels and other dreissenids are the only freshwater bivalves to produce highly dispersible planktonic veligers and to retain their attachment organ into adulthood (Morton 1993). These two properties give zebra mussels the capability to spread quickly and adhere tenaciously in the Great Lakes. The attachment apparatus of the zebra mussel is an extraorganismic, shock absorbing structure, called the byssus. The molecular mechanism of the zebra mussel attachment performed by the byssus remains largely unexplored. To this end, this study attempts to unravel the nature and regulation of zebra mussel byssus activities at the molecular level. The overarching hypothesis is that byssogenessis in *D. polymorpha* is mediated through expression of multiple genes; each should play a role in a cascade of events. Proteins of some of the expressed genes should possess a regulatory function during a certain stage of zebra mussel byssal thread production. The expression of genes involved in byssogenesis of dreissenids can be altered by a number of environmental factors such as temperature, current rate, and dissolved oxygen levels.

To test these hypotheses, a number of objectives have to be identified: 1) To identify the genes unique to zebra mussel foot and determine if they can be potentially

involved in the activities of zebra mussel byssus; 2) To develop a cDNA microarray containing all the genes unique to zebra mussel byssus; 3) To find out the genes up- or downregulated in response to the changes of environmental factors, as well as the change of byssogenesis status.; 4) To identify the zebra mussel byssus unique genes with differential expression profiles at the different stages of byssogenesis; and 5) To study the molecular characteristics, protein activities and potential functions of the genes identified by the microarray analyses.

Beginning with the construction of a normalized cDNA library containing the genes unique to zebra mussel byssus, this study applied a high-throughput technique, cDNA microarray to identify the gene expression profiles during different byssus activities in large scale. The cDNA library described in chapter II provided more than 700 expressed sequence tags (ESTs) that are unique to zebra mussel byssus for the development of the zebra mussel cDNA library. The putative functions of the proteins encoded by these ESTs have a broad spectrum such as, foot structural proteins (Dpfp), exocrine gland peptides (EGP), host defense related, other normal functional proteins, and the proteins without known homologues in database.

The purpose of the development of cDNA microarray in chapter III is to study the expression patterns of the zebra mussel byssus unique genes during the activities of the byssus. For example, in this chapter, two different zebra mussel attachment statuses were created, attachment and non-attachment. By comparing the expression profiles of the genes between these two status, we found seventeen genes were either up or downregulated in the attachment. It is the first time to use microarray technique to study

the zebra mussel attachment in molecular level. Also, the results of this study suggest that the newly developed zebra mussel microarray is efficient and reliable.

The zebra mussel microarray was also applied in chapter IV and V in order to identify the genes differentially displayed along the course of byssogenesis (CHAPTER IV) and those modulated in response to changes in environmental factors (CHAPTER V). From both of the assays, a number of genes encoding the proteins homolgous to *D. polymorpha* foot protein (Dpfp), EGP, and neuopeptide-like proteins (nlp) were identified. This suggested that the three families of proteins are very likely to play important roles in the different activities of zebra mussel byssus and possibly as new foot proteins that have never been reported before.

Interestingly, a host defense related molecule, *D. polymorpha* defensin (Dpd), was found significantly expressed by zebra mussel byssus gland cells and also had the expression profiles associated with the production of zebra mussel byssal threads. It is suggested that beside the basic antimicrobial activities reported in CHAPTER VI, Dpd may be necessary in the process of attachment by protecting byssal threads against microbial degradation.

The successful application of cDNA microarray to this study provided a great tool to understand byssus activities of the zebra mussel. The differentially expressed genes in response to the change of byssogenesis status, environmental factors and stages of byssogenesis identified in this study are of interest in understanding the mechanism of zebra mussel underwater adhesion.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Native to Russian waters, the zebra mussel (*Dreissena polymorpha* Pallas 1771) was first introduced to the Great Lakes in the late 1980s. The non-native species has colonized an alarming number of North American inland waterways and lakes (Benson & Raikow 2009). Zebra mussels reproduce and spread rapidly, leaving behind a trail of habitat destruction, harmful algal blooms, loss of native species, and economic devastation. Developing an effective strategy to eradicate zebra mussels, or to at least control their spread, remains a remote possibility despite concerted efforts by scientists and managers.

Why are zebra mussels so abundant when most bivalve mollusks populations worldwide are either declining or at the brink of extinction? One likely reason is the zebra mussel's extraordinary ability to attach tightly and expeditiously to any underwater hard surface including adults of their own species (Johnson & Padilla 1996). Additionally, optimal environmental conditions, lack of competition, and scarcity of natural predators or pathogens may have also contributed to the explosive increase in zebra mussel populations in the Great Lakes.

Zebra mussel invasions have severe impacts on the health of the ecosystem and the local economy. Industrial facilities devote significant efforts to the cleaning of water intake pipes and heat exchangers clogged with zebra mussels. Water circuits in the facilities must be cleaned manually, where possible, or with various chemicals in processes that are both expensive and environmentally harmful (Cope *et al.* 1997). The profound economic impact of zebra mussel expansion in the Great Lakes basin alone has

mounted billions of dollars in increased operating costs of affected industries and municipalities over decades (Benson & Raikow 2009). This situation has overwhelmed North American scientists and water resource managers who unexpectedly became confronted with an ecologic catastrophe of an unprecedented magnitude. The major obstacle that impedes immediate intervention or development of an effective management strategy is the dearth of knowledge on the basic biology of zebra mussels and the survival strategies used by this successful intruder.

The zebra mussel is unique in many biological properties. For example, zebra mussels and other dreissenids are the only freshwater bivalves to produce highly dispersible planktonic veligers and to retain their attachment organ into adulthood (Morton 1993). These two properties give zebra mussels the capability to spread quickly and adhere tenaciously in the Great Lakes. The attachment apparatus of the zebra mussel is an extraorganismic, shock absorbing structure, called the byssus. The molecular mechanism of the zebra mussel attachment performed by the byssus remains largely unexplored. To this end, this study attempts to unravel the nature and regulation of zebra mussel byssus activities at the molecular level. The overarching hypothesis is that byssogenessis in *D. polymorpha* is mediated through expression of multiple genes; each should play a role in a cascade of events. Proteins of some of the expressed genes should possess a regulatory function during a certain stage of zebra mussel byssal thread production. The expression of genes involved in byssogenesis of dreissenids can be altered by a number of environmental factors such as temperature, current rate, and dissolved oxygen levels.

To test these hypotheses, a number of objectives have to be identified: 1) To identify the genes unique to zebra mussel foot and determine if they can be potentially involved in the activities of zebra mussel byssus; 2) To develop a cDNA microarray containing all the genes unique to zebra mussel byssus; 3) To find out the genes up- or downregulated in response to the changes of environmental factors, as well as the change of byssogenesis status.; 4) To identify the zebra mussel byssus unique genes with differential expression profiles at the different stages of byssogenesis; and 5) To study the molecular characteristics, protein activities and potential functions of the genes identified by the microarray analyses.

Beginning with the construction of a normalized cDNA library containing the genes unique to zebra mussel byssus, this study applied a high-throughput technique, cDNA microarray to identify the gene expression profiles during different byssus activities in large scale. The cDNA library described in chapter II provided more than 700 expressed sequence tags (ESTs) that are unique to zebra mussel byssus for the development of the zebra mussel cDNA library. The putative functions of the proteins encoded by these ESTs have a broad spectrum such as, foot structural proteins (Dpfp), exocrine gland peptides (EGP), host defense related, other normal functional proteins, and the proteins without known homologues in database.

The purpose of the development of cDNA microarray in chapter III is to study the expression patterns of the zebra mussel byssus unique genes during the activities of the byssus. For example, in this chapter, two different zebra mussel attachment statuses were created, attachment and non-attachment. By comparing the expression profiles of the genes between these two status, we found seventeen genes were either up or

downregulated in the attachment. It is the first time to use microarray technique to study the zebra mussel attachment in molecular level. Also, the results of this study suggest that the newly developed zebra mussel microarray is efficient and reliable.

The zebra mussel microarray was also applied in chapter IV and V in order to identify the genes differentially displayed along the course of byssogenesis (CHAPTER IV) and those modulated in response to changes in environmental factors (CHAPTER V). From both of the assays, a number of genes encoding the proteins homolgous to *D. polymorpha* foot protein (Dpfp), EGP, and neuopeptide-like proteins (nlp) were identified. This suggested that the three families of proteins are very likely to play important roles in the different activities of zebra mussel byssus and possibly as new foot proteins that have never been reported before.

Interestingly, a host defense related molecule, *D. polymorpha* defensin (Dpd), was found significantly expressed by zebra mussel byssus gland cells and also had the expression profiles associated with the production of zebra mussel byssal threads. It is suggested that beside the basic antimicrobial activities reported in CHAPTER VI, Dpd may be necessary in the process of attachment by protecting byssal threads against microbial degradation.

The successful application of cDNA microarray to this study provided a great tool to understand byssus activities of the zebra mussel. The differentially expressed genes in response to the change of byssogenesis status, environmental factors and stages of byssogenesis identified in this study are of interest in understanding the mechanism of zebra mussel underwater adhesion.

CHAPTER I

Literature Review

Invasion of Zebra Mussels

In 1988, disturbing reports announced that zebra mussels (*Dreissena polymorpha*) were found in Lake St. Clair; a water body connecting Lake Huron and Lake Erie. This non-native species found a refuge in the Laurentian Great Lakes watershed and has successfully invaded the majority of the water shed (Hebert *et al.* 1989). By 1990, zebra mussels had been found in all the Great Lakes. Zebra mussels escaped the Great Lakes basin and found their way into the Illinois and Hudson rivers in 1991. Invasion into the Illinois River led to their introduction into the Mississippi River drainage system which covers over 1.2 million square miles. By 1992, populations of zebra mussels were established in the following rivers: Arkansas, Cumberland, Hudson, Illinois, Mississippi, Ohio, and Tennessee. Zebra mussels continued their spread across the continental U. S. and by 1994 were reported in water bodies of the following states: Alabama, Arkansas, Illinois, Indiana, Iowa, Kentucky, Louisiana, Michigan, Minnesota, Mississippi, Missouri, New York, Ohio, Oklahoma, Pennsylvania, Tennessee, Utah, Vermont, West Virginia, and Wisconsin (Benson & Raikow 2009). More recently, Connecticut (2002), Nebraska (2003), and South Dakota (2003) have been listed as additional states where zebra mussels were found (Benson & Raikow 2009). The zebra mussel has continued its spread west, south, and east; in January 2008, zebra mussels reached the southwestern United States, in the San Justo Reservoir in central California (Benson & Raikow 2009).

Currently in june 2009, zebra mussels were found in Lake Texoma on the border of Texas and Oklahoma (Benson & Raikow 2009). It is believed that the zebra mussel was first introduced to North America through the ballast exchange of a commercial cargo ship traveling from the north shore of the Black Sea to the Great Lakes of North America (McMahon 1996). This rapid range expansion by the zebra mussel was accomplished not only by the passive drifting of the larva, but also by its ability to attach to boats navigating lakes and rivers (Benson & Raikow 2009). An adult zebra mussel under cool and humid conditions stays alive for several days when taken out of water. That provides the mussel enough time to be transmitted to inland lakes and rivers through trailers and vehicles (Benson & Raikow 2009).

Impacts to Economy and Ecology

The underwater attachment of the zebra mussel has made this invasive species notorious for its biofouling capabilities. Zebra mussels colonize in underwater constructions, such as water supply pipes of hydroelectric and nuclear power plants, public water supply plants, and industrial facilities. Moreover, zebra mussels also cause deterioration and corrosion of underwater structures such as navigational buoys, fishing gears, boats, and docks (Benson & Raikow 2009). The impacts of the zebra mussel to the ecological system are even more profound. The huge consumption of phytoplankton, as well as other suspended materials including bacteria, protozoa, microzooplankton, and silt, has seriously reduced the biomass of local aquatic systems. The increased water clarity caused by the decrease of biomass allows light to penetrate further to the deep level of the water body which potentially promotes macrophyte populations (Skubinna *et al.* 1995). As most of the phytoplankton are consumed by the zebra mussel, the

concentration of dissolved organic carbon (DOC) can drop to an extremely low level; which has been reported in some inland lakes colonized by zebra mussels (Benson & Raikow 2009). The negative effects of zebra mussels on the ecosystem have induced major disturbances in the foodweb. The reductions of phytoplankton and zooplankton biomass cause huge impacts to fish populations, such as the increase of competition, decrease of survivals and the extinction of planktivorous fish. Since the microzooplanktons are more heavily affected by zebra mussels, fish larvae, which feed on the microzooplanktons may be more greatly affected than other later life stages (Raikow 2004). Moreover, native unionids were also affected by the zebra mussel invasion (Schloesser *et al.* 1996; Baker & Hornbach 1997). By simply attaching to the shells of unionids, the zebra mussels restrict valve operation, cause shell deformity, smother siphons, compete for food, impair movement, and deposit metabolic waste onto unionid clams. According to a previous study, the survival rates of native unionid mussels in the Mississippi River, Minnesota significantly declined with the zebra mussel colonization (Hart *et al.* 2001). The unionids in Lake St. Clair and western Lake Erie, to date, have been extirpated (Benson & Raikow 2009).

Control Methods against Zebra Mussel

The profound economic and ecological impact of zebra mussel expansion in the Great Lakes basin has mounted to over billions of dollars in increased operating costs of affected industries and municipalities (Roberts 1990). A variety of control methods have been developed against the attachment and distribution of zebra mussels. Physically, zebra mussels can be detached by exoteric force. Many types of equipment have been invented to remove the attached zebra mussels from underwater substrates, such as high

pressure water hoses and acoustical vibration. Alternatively, a number of chemical molluscicides including chlorine, chlorine dioxide, and other oxidizing and non-oxidizing chemicals have been applied in some areas to eliminate the zebra mussels. Some other physical changes of the environmental conditions may also lead to the mortalities or detachment of zebra mussels. Such physical elimination methods include dewatering/desiccation, thermal controlling, electrical current, CO_2 injection, and ultraviolet light. Instead of detaching the zebra mussels from the substrates, people also apply special coating materials on the surface of the structures, such as copper, zinc, and silicone-based paint (Benson & Raikow 2009). However, all the strategies mentioned above are either not sufficient or not environment-friendly. Even the most recently introduced biological control, such as the introduction of natural mussel predators (Bially & MacIsaac 2000), selective toxic microbes and parasites (Molloy 1998), and disrupting reproductive process (Snyder *et al.* 1997), have been proven to be ineffective. The ultimate goal of these control methods is to disrupt or prevent attachment of zebra mussels. However, the major obstacle that impedes immediate intervention or development of an effective management strategy is the dearth of knowledge on attachment mechanism employed by this successful intruder.

Zebra Mussel Byssus

The attachment organ of the zebra mussel is an extraorganismic, shock absorbing structure, called the byssus, which was first described by van der Feen in 1949 (Waite 2002). Except for the morphological characteristics, the biochemical nature of the byssus and processes involved in zebra mussel's adhesion to substrates remain largely unexplored. The byssus apparatus is considered to be a masterpiece of underwater

bioadhesion because of the rapid and robust production, as well as the relatively simple components of adhesive proteins (Waite 2002). The silver hair-like structures that can be seen out of shells in most natural conditions are called byssal threads. All byssal threads are generated from the stem located at the root of the foot. The enlarged plaque end of the byssal thread is the structure holding all the adhesive proteins at the interface of the byssal thread and substrata surface. In the zebra mussel, the byssus also includes three main glands located in the foot, namely, the stem-forming gland, thread-forming land, and plaque-forming gland. The proteins needed for byssal thread formation and adhesion are secreted by the glands and expelled to the groove of the ventral side in the foot and transported out of the organ (Figure 1-1). The anatomic structure of the zebra mussel byssus and the production of the adhesive proteins have been well described by Rzepecki and Waite (Rzepecki & Waite 1993a, b).

Two distinct types of byssal threads are produced by zebra mussels. One of them is the temporary byssal thread which possesses extraordinary plasticity. It is often seen in the juvenile stage and early stage of attachment of adult zebra mussels. This type of byssal thread helps to explore the precipitous surfaces of underwater substrates. Morphologically, the temporary threads are anchored by an elastic, mucous filament which is compositionally distinct from the threads and plaques (Rzepecki & Waite 1993a). The other type of byssal threads are permanent byssal threads, which appears mostly at the adult stage of the zebra mussel and help mussels for longer term settlement. There are two distinct sections observed in the permanent byssal threads, the proximal portion that is close to the stem of byssal threads, and the distal portion which is close to the attachment plaque of the thread. Observations of the exterior surface characteristics of

the byssal threads show that the surface of proximal region is relatively smooth, whereas the surface is rough at the distal portion (Eckroat *et al.* 1993).

Once the zebra mussels are relocated, the majority of byssal threads are formed within one week to ensure a rapid, strong attachment. Thereafter, the rate of byssal thread formation remains constant (Eckroat *et al.* 1993). As a zebra mussel prepares to produce byssal threads, the foot is moved along the surface of the substrate searching for suitable attachment site. At the same time, the space between the zebra mussel foot and the substrate is filled with mucus provided by the ventral groove. The mucus does not resemble the byssal threads, but is there to create a hydrophobic and clean environment for the formation of the byssal threads. The actual protein components are produced by the three major byssus glands which have been described above. The first formed section of byssal threads is the plaque; thereafter, the foot slowly retracts back into the shell while the clear zebra mussel byssal threads start becoming visible (Eckroat *et al.* 1993). Three main protein components have been reported to be involved in the formation of zebra mussel byssal threads, namely, *D. polymorpha* foot protein (Dpfp) -1, -2, and -3 (Rzepecki & Waite 1993b, a). However, only two of them, Dpfp-1 and -2 have been sufficiently purified for characterization. Preliminary studies on Dpfp-1 and -2 suggested that both of these two foot proteins may serve as the component maintaining byssal structures, rather than adhesive proteins (Anderson & Waite 2000). Dpfp-1 and Dpfp-2 contain a number of 3,4-dihydroxyphenylalanines (DOPAs) in their primary sequence. The DOPA proteins are tandemly repetitive with unique oligopeptide motif sequences (Rzepecki & Waite 1993b). Dpfp-1 was found to have at least 10-15 variants which constituted a polymorphic protein family specific to the zebra mussel (Rzepecki & Waite

1993b). Crosslinks formed by the Dpfp-1 variants along with DOPA create a dovetail frame, which is the primary component to maintain the byssal thread structures. Meanwhile, DOPA also participates in protein crosslinking during maturation of byssal threads, following oxidation to DOPA quinone. These crosslinks will structurally reinforce the weakest components of the byssus.

Similar structures have been observed in many marine byssus-carrying species, such as *Mytilus* sp. (Brown 1952; Price 1983; Waite 1983a, 1992). At least 12 proteins have been identified from the byssus of *Mytilus* species and eight of them are characterized as adhesive proteins located in the plaque. *M. edulis* foot protein-1 and -2 (Mfp-1 and Mfp-2) are the two most well studied foot proteins that can be easily purified from the *M. edulis* foot. The Mfp-1, which constitutes about 5% of the plaque matrix proteins, is a molecule to mediate bonding to the intended substratum during attachment (Benedict & Waite 1986; Waite *et al.* 1998; Sun & Waite 2005; Waite *et al.* 2005), while the Mfp-2 performs as a structural component comprising 25% of the plaque content (Papov *et al.* 1995). Also Mfp-1 covers all the exposed portion of the byssus to provide protection from the environment (Lin *et al.* 2007). Mfp-3 is the real adhesion proteins acting as glue in the plaque proteins (Lin *et al.* 2007). Mfp-4, with high levels of histidine, lysine and arginine in the primary structure, is proposed to be the protein that serves as a coupling agent in the thread plaque junction (Cha *et al.* 2008). Mfp-5 is a plaque specific protein with the highest 3,4-dihydroxyphenyl-L-alanine (DOPA) level (~30 mol.%) among all the foot proteins (Waite & Qin 2001).

Previous studies demonstrate that mussel adhesion largely depends on DOPA (Deming 1999; Waite 2002), the higher the DOPA content, the stronger the adhesion (Yu

& Deming 1998; Lee *et al.* 2006). Mfp-3 and Mfp-5 are two adhesive molecules that exist in the contact area between plaques and solid surfaces in the water (Waite 2002; Zhao *et al.* 2006; Zhao & Waite 2006). The main component of the byssal threads are two collagens, preCol-D and –NG. They dominate from the fibrous core of byssal threads (Waite *et al.* 1998) to the plaque matrix with their frayed ends (Waite *et al.* 2005). Recently, another plaque protein, Mfp-6, was identified from *Mytilus californianus*, which is a close relative to *M. edulis*. Mfp-6 is a thiol-rich protein that has potential function of mediating the cysteinyldopa cross-link between the adhesive proteins and the solid surface (Zhao & Waite 2006). The distributions of these foot proteins of Mytulids are shown in Figure 1-2.

Details about byssal thread formation also mostly comes from *M. eulis*, a marine byssate species. The medium for habitat of water mussels, perhaps poses some challenges to adhesion (Cayless 1991). Interactions between the adhesive proteins in the plaque and solid surfaces in the water are mostly noncovalent including charge-charge, hydrogen bond, and van der Waals forces which contain dipole-dipole, induced dipole-dipole, and nonpolar coupling interactions. However, the resistance of adhesion formation with the noncovalent bonds in water is about 80 times stronger than that in the air. If the mussel attaches to a hard surface in water, the individual should be able to remove weak boundary layers including water, dirt, microbes, etc, from the surface. Moreover, the mussel has to keep noncovalent bonds of adhesion from being subverted by water all around the plaques. In fact, adhesion of the mussels relies on different interfacial interactions, and can happen on a wide range of underwater materials. That suggests that the mussel is able to recognize and respond to different kinds of surfaces (Waite 2002).

Information about the mechanism of byssal thread formation is very limited according to the previous studies. In *M edulis*, the byssal proteins are synthesized by byssus glands and stocked in the foot. During thread formation, granular thread precursors are secreted by the glands and collected in the ventral groove in the foot. With the contraction of the foot, the secreted precursors are mixed and shaped, thereafter loaded onto the surface of the substratum. The maturation of byssal threads is very similar to formation of intermolecular cross-links derived from oxidized DOPA (McDowell *et al.* 1999).

Although the Dfp and Mfp have some similarities in their characteristics, the classification of dreissenids indicates that adhesion mechanisms employed by zebra and quagga mussels will be different than that of blue mussels. The *Dreissena* is classified in the superorder Veneroida with a non-byssate North American invader, *Corbicula fluminea*; while the mytilids belong to the superorder Pterioida (Allen 1985). This suggests that the dreissenid byssus evolved independently of that in mytilids (Morton 1993). Therefore, there may be other adhesive proteins that haven't been isolated or regulation mechanisms of the two byssate mussels are completely different.

Zebra Mussel Byssogenesis and Environmental Factors

In adult stages of the zebra mussel, a number of environmental factors have been reported to influence byssogenesis, or the rate of production of new byssal threads. Clarke and McMahon designed a series of experiments to test the effects of physical factors on byssogenesis (Clarke & McMahon 1996b, c; Clarke & McMahon 1996a; Clarke 1999). The byssogenesis rate of zebra mussels at the first week was boosted with an increase of temperature, within the range of 5 – 30 °C (Clarke & McMahon 1996c). Maximum byssal thread generation rate appeared when the current velocity was 0.2 m/s.

Byssogenesis rate was decreased with current speeds higher or lower than 0.2 m/s (Clarke & McMahon 1996a). With dissolved oxygen level at 30.9 torr, the zebra mussels had maximum byssogenesis rate while low dissolved oxygen level led to hypoxia, which caused a significant decrease of byssogenesis rate, as well as mussel mortalities. The byssogenesis rate decreased when dissolved oxygen levels increased to a very high level at 154.3 torr (Clarke & McMahon 1996b). Interestingly, the byssogenesis rate is also higher when more food is available (Clarke 1999). These results indicate that, at the molecular level, expression levels of genes involved in byssogenesis regulation can be affected by changes in environmental factors. The response to the changes in environment may be controlled by different byssogenesis genes. What are those genes involved in the byssogenesis and how do these genes regulate the byssogenesis to accommodate the environment are two very interesting questions about the byssogenesis mechanism.

In fact, the zebra mussel underwater adhesion mechanism does not only include the byssal thread structural proteins and plaque adhesive proteins, but also combines many other regulation processes including protein release and modification, byssus protection, and even signal transduction. The study of individual molecules has helped us learn more about the characteristics of the byssus components; however, it did not give the answer to the questions about the molecular regulation mechanism of zebra mussel attachment. At this point, a new strategy of the zebra mussel adhesion mechanism study is necessary to solve this complex puzzle.

cDNA Library Construction and Suppression Subtractive Hybridization cDNA

Given that a number of proteins participate in the regulation of zebra mussel attachment, conventional methods focusing on a few zebra mussel proteins may not be efficient enough to understand the complicated mechanism employed by zebra mussels during attachment. Therefore, we opted to employ assays from which we can obtain accurate information on a relatively large number of molecules of relevance to the attachment process. Most high throughput approaches in the field of functional genomics start with the construction of a complimentary DNA (cDNA) library to messenger RNA (mRNA) of expressed genes at the time of collection. The mRNA collected from target cells is reverse-transcribed to stable complementary DNA. One potential benefit to this approach is that cDNA libraries lack information about enhancers, introns, and other regulatory elements found in a genomic DNA library.

Although the cDNA library contains much less redundant sequence information compared to a genomic DNA library, the abundance of each gene in the library is tremendously different than the others. A typical eukaryote somatic cell contains three different classes of mRNAs based on their frequencies (Bishop *et al.* 1974; Davidson & Britten 1979). The frequencies of highest and lowest are 40-45% and 10% respectively. On the average, in a single cell there are about 10 most prevalent mRNA species and each of them is represented by 5,000 copies. On the other hand, the class of high complexity category includes about 15,000 mRNA species, with each of them represented by only 1-15 copies (Soares *et al.* 1994). Hence, the identification of the rare mRNA species in regular cDNA libraries is extremely difficult compared to mRNAs with

high frequencies in cells. It is therefore highly recommended to normalize the cDNA library so that each expressed gene is represented at similar frequency in the cDNA library.

Suppression subtractive hybridization (SSH) is a simple and efficient method for generating cDNAs highly enriched for differentially expressed genes of both high and low abundance (Diatchenko *et al.* 1996; Diatchenko *et al.* 1999; Wang & Feuerstein 2000; Ji *et al.* 2002). A high level of enrichment of rare transcripts has been reported from a cDNA library normalized by SSH; a 1,000- to 1,500-fold enrichment for rare cDNAs (Diatchenko *et al.* 1996). SSH has been successfully applied in studies on tissues (Li *et al.* 2000; Villalva *et al.* 2001) and cell lines (Eleveld-Trancikova *et al.* 2002; Langley *et al.* 2003).

In SSH, cDNAs synthesized from mRNAs of target tissues or cells are used as tester cDNA while the cDNAs from the mRNA isolated from control tissues or cells are drivers (Figure 1-3). The objective of SSH is to reduce cDNAs that are not unique to tester tissues or cells by subtracting cDNAs abundant in both testers and drivers from the tester cDNAs. Briefly, the tester cDNAs are divided into two groups, tester 1 and tester 2, and connected with two different DNA adapters, namely, adapter 1 and adapter 2 respectively. An excessive amount of driver cDNAs without any adapters is used to hybridize with tester 1 and tester 2, respectively. The hybridized tester 1 and tester 2 are then hybridized to each other. Thereafter, the cDNA pool contains a variety of hybridization products and is used for PCR with a pair of primers designed based on the sequence of adapter 1 and adapter 2. The only type of cDNA strand that can be amplified

by PCR is the one with different adapters connected at the 5' and 3' ends; this cDNA strand is also unique to tester tissues or cells (Figure 1-3).

Anatomic studies of the zebra mussel foot indicated that other than muscular tissue and epithelial covering, there are a number of byssus glands responsible for the development and structure of byssal threads and end plaques. If the cDNAs synthesized from the foot mRNA are used as a tester, the cDNAs abundant in the muscle cells can be greatly subtracted with the cDNAs from adductor or retractor mussels of the zebra mussel (Figure 1-4). The enriched cDNAs in the tester are likely to reflect the composition of mRNA in zebra mussel byssus glands.

cDNA Microarray

The normalized cDNA library will provide a number of genes unique to the zebra mussel byssus; however, identification of expression profiles and the potential functions involved in zebra byssus activities requires another technique to analyze gene expression on a large scale. Evolved from Northen blotting techniques, the DNA microarray technique was first applied and described as a high throughput method to identify genes whose expression is modulated by interferon (Kulesh *et al.* 1987). With further development and improvement during the past two decades, the DNA microarray technique has been applied to many fields, such as discoveries of the exons and genes for annotation of human genome draft sequence (Shoemaker *et al.* 2001), the analysis of genomic DNA for detection of amplifications and deletions in tumors (Hodgson *et al.* 2001; Fritz *et al.* 2002), and the differential gene expression analysis of the networks of genes within common pathways of regulation (Zhu *et al.* 2000; Miki *et al.* 2001).

Two main types of DNA microarray have been developed, spotted DNA microarrays and in situ synthesized oligonucleotide arrays (Affimetrix) (Dufva 2009). Although the Affymetrix microarray is more advanced in many aspects, including the lower background of hybridization, higher specificity, and more accessibility for hybridization, the spotted DNA microarray is currently widely used because it can be produced in academic laboratories with affordable prices (Bowtell & Sambrook 2003). In the cDNA microarrays, singled strand DNAs obtained from PCR are spotted and fixed in a glass slide. Separately labeled with different fluorescent dyes (usually green and red), two cDNA samples from different cells populations or tissues are mixed and then hybridized simultaneously with one microarray. These two cDNA probes competitively anneal to their complementary nucleotide sequences on the microarray slide and the rest of the probes that cannot be anchored by the DNA strands on the slides will be washed away. The expression levels of each probe from each sample can be detected by capturing the color and strength of the signals in each spot on the microarray. The intensity ratio at a spot is thus a measure of the relative abundance of the gene in the two samples (Nguyen *et al.* 2002; Stears *et al.* 2003).

In the present study, the use of the cDNA microarray has enabled us to understand the genetic regulation of the byssus unique genes involved in the process of byssogenesis, from the beginning of byssal threads being produced, until they mature within 21 days. As discussed above, byssogenesis can be divided into a number of stages such as the secretion of byssal proteins, the generation of temporary byssal threads, the formation of permanent threads, and structural modifications of the permanent threads. These steps involve various proteins, and therefore, a time-course microarray, as applied in this study,

proved helpful in the identification of the genes differentially regulated over the course of byssogenesis. The time-series analyses of microarray have been applied since the invention of the cDNA microarray technology. For example, DeRisi *et al* in 1997 successfully applied a time-course analysis on the study of the gene expression patterns in yeast during metabolic shift from fermentation to respiration (DeRisi *et al.* 1997). A similar study was performed by Chu *et al* in 1998 also with time-course microarray design (Chu *et al.* 1998).

The effects of the surrounding environmental factors on byssogenesis are complex to study, whether single or combined. Therefore, the non-reference factorial design of microarray analysis has been used in this study, based on the recommendations of a number of statisticians (Draghici *et al.* 2001; Churchill 2002; Kerr 2003). A particularly popular non-reference factorial design, the loop design (Kerr & Churchill 2001), involves at least two different arrangements of biological replicates (Dobbin *et al.* 2003). For an experiment dealing with the effects of more than two factors, the interwoven loop design (Figure 1-5) with more than one loop, is statistically more efficient than a common reference design, as it leads to less disparity in precision and power comparisons between any two treatments (Tempelman 2005). Indeed this approach allows the researcher to obtain more information from the microarray experimental design including multiple environmental factors with least number of hybridizations.

Zebra Mussel Host Defense Mechanism and Byssus Activities

Along with the generation of byssal threads, other proteins are probably being secreted along with the byssal threads. The highly proteinaceous components of fresh byssal threads are likely to be degraded by microbial communities, if there were no efficient mechanisms employed by the zebra mussels to protect their threads. On the other hand, the biofilms existing on the surface of underwater substances form a barrier between the substrates and adhesive proteins at the ends of the byssal threads. Some biofilms have negative effects on adhesion of both larval (Kavouras & Maki 2003a) and adult stages (Kavouras & Maki 2003b, 2004; Angarano *et al.* 2009) of zebra mussel byssus.

Like other bivalve mollusks, the host defense system employed by the zebra mussel consists of both cellular and humoral components. The cellular defense is primarily performed by hemocytes, which are found in a semiclosed circulation system, including the heart, vessels, and sinus with different sizes located in major organs (Auffret 2005). Phagocytosis by hemocytes represents the main process of cellular defense and goes through a number of phases that encompass recognition, adhesion, ingestion, destruction and elimination of foreign cells (Pipe 1990, 1992; Tiscar & Mosca 2004). There are different methods for hemocytes to destroy the pathogens, such as using lysosomal enzymes and respiratory burst. Similar to mammalian phagocytes, the respiratory burst of bivalve hemocytes can be induced by certain stimuli. A number of reactive oxygen species (ROS) will be produced during the respiratory burst activity, such as superoxide, the hydroxyl radical, singlet oxygen, hydrogen peroxide, hypohalides,

halidamines, nitric oxide and the peroxynitrite. The reactions of these oxidants involved in the respiratory burst has been identified from different mollusc hemocytes such as *Patinopecten yessoensis*, *Pecten maximus*, *Crassostrea virginica*, *C. gigas*, *Ostrea edulis*, *Mya arenaria*, *Mercenaria mercenaria*, *Mytilus edulis* (Roch 1999), *M. galloprovincialis* (Arumugam *et al.* 2000), and *Ruditapes decussatus* (Tafalla *et al.* 2003). However, the mechanism of ROS production and regulation against infections of pathogens remains largely unknown (Roch 1999).

The humoral defense system also includes lysosomal enzymes (Tiscar & Mosca 2004), opsonins (Renwrantz & Stahmer 1983) and antimicrobial peptides (AMPs) (Hubert *et al.* 1996; Mitta *et al.* 2000a). Produced in hemoytes, the lysosomal enzymes (β-glucuronidase, acid and alkaline phosphatase, lipase, aminopeptidase and lysozyme) are released into serum during phagocytosis (Pipe 1990). The lysosomal enzymes can also be found in digestive glands due to the fact that filtered bacteria represent nourishment for marine bivalves and the enzymes produced by digestive glands have to function in both digestion and protection simultaneously (Tiscar & Mosca 2004). Two major types of opsonin, agglutinins and lectins have been found in many bivalve tissues. The main targets of the agglutinins are erythrocytes, bacteria, protozoa and algae (Chu 1988), while the lectins have a specific opsonizing function in hemocyte aggregation and foreign cell agglutination. The lectins also manage to promote the recognition of foreign cells because they are specific for hemocyte and bacteria glycoconjugates (Tiscar & Mosca 2004).

As a main component of the bivalve humoral defense system, the AMPs have been identified not only from bivalve mollusks, but also from a variety of organisms in

other phyla (Boman 2003; Bulet *et al.* 2004). Over 1,000 identified AMPs have been classified in three groups based on their secondary structures (Bulet *et al.* 2004), including linear peptides with α-helices, highly disulphide-bonded (cysteine-rich) β-sheets, and those with proline- or glycine-rich character (Douglas *et al.* 2003; Gueguen *et al.* 2006). By inserting themselves into membranes, the AMPs cause the malfunctioning of pathogen membranes, thereby leading to lysis of pathogens. Moreover, depending on the tissue distribution, the AMPs are also involved in a number of other reactions, such as mediating inflammation (Hancock & Lehrer 1998). To date, AMPs have been identified from several bivalve species and defined as several types, such as defensins, mytilins, myticins, and mytimycin. Molluskan AMPs have been found to inhibit the growth of a broad spectrum of pathogens, primarily bacteria and fungi (Mitta *et al.* 2000b).

Studies on zebra mussels, in past decades, have provided us with basic knowledge about zebra mussel underwater adhesion, such as the biochemical characteristics of the foot proteins, the structures of the byssus apparatus, the stage of byssogenesis, and the activities of byssus under the impact of different environmental factors. This suggested that the zebra mussel employs a sophisticated system to produce, maintain, and modify the byssal threads. This process may involve several regulation systems that cannot be simply reflected by the morphological change of byssal threads. To this end, this study was undertaken with the following objectives: 1) to obtain the genes unique to zebra mussel byssus and participate in zebra mussel byssus activities; 2) to identify the genes differentially expressed during the formation of byssal threads; 3) to determine how the environmental factors affect the zebra mussel byssogenesis on gene level; 4) to find out the molecules with potential functions related to byssal threads protection.

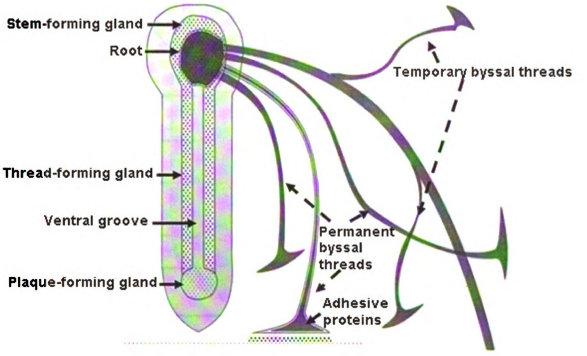


Figure 1-1 The anatomic structure of zebra mussel foot and byssus.
The stem-forming gland surrounds the root of byssal threads. The cells of thread-forming gland are distributed along the ventral groove in the middle of foot. The plaque-forming gland is located at the tip of zebra mussel foot. The foot proteins produced by thread-forming gland and the thread-forming gland releases the protein into the ventral groove transported to the root of the foot. Two types of byssal threas, temporary threads and permanent threads, are produced with the protein materials secreted by the three byssus glands.

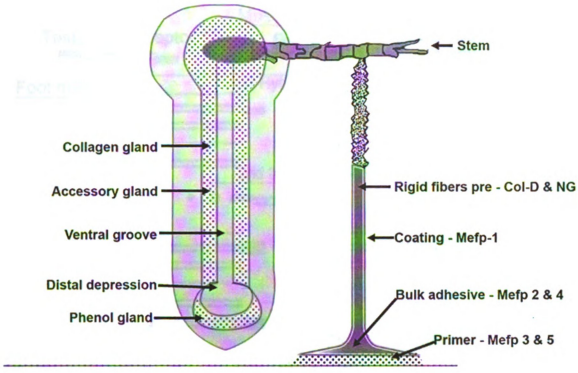


Figure 1-2 The structure of Mytilid byssus and distributions of foot proteins.
Three byssus glands were located in the foot of Mytilid. Seven types of foot proteins, by far, have been identified from a variety of Mytilids. The structural proteins of Mytilids byssal threads are pre-Col-D and NG coated with Mefp-1. The proteins filled in the plaques are bulk adhesive proteins Mefp 2 and 4, and the primer proteins, Mefp 3 and 5.

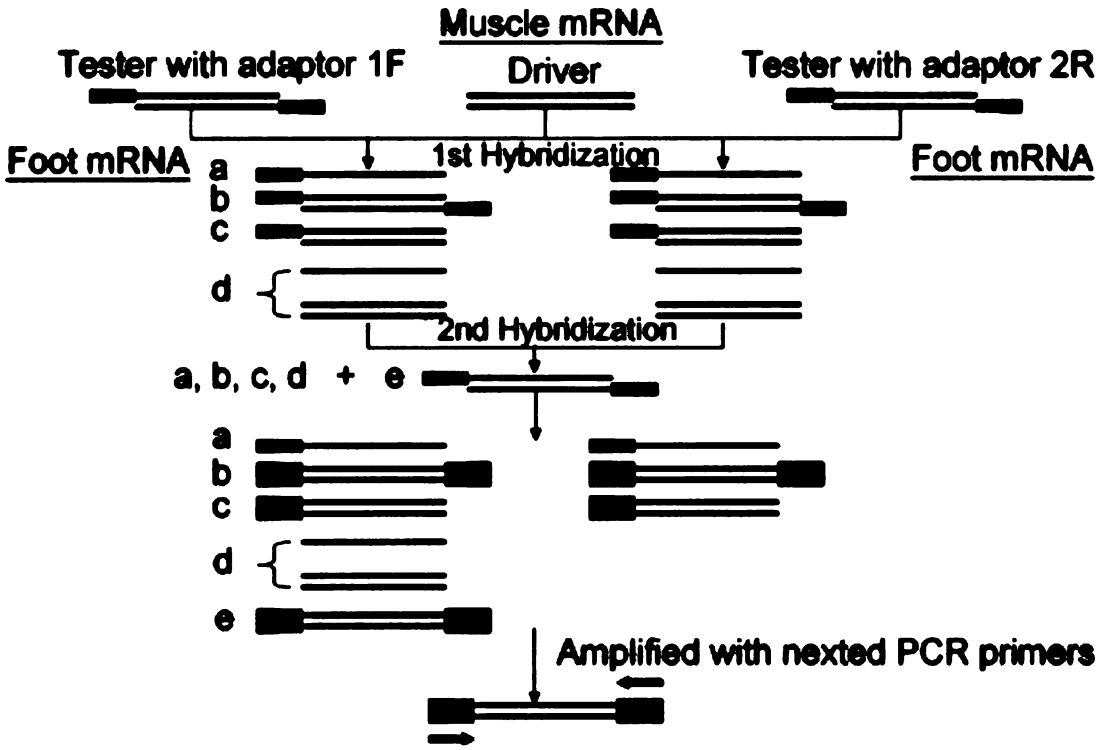


Figure 1-3 The construction of suppression subtractive hybridization cDNA library. Tester cDNAs are connected to different adaptors, 1F and 2R. After two rounds of hybridizations, there are many forms of cDNA templates that were generated; however, only the templates connected to different adaptors at 5' and 3' ends can be amplified by the adaptor specific primers. These amplified templates were originally unique in tester cDNAs.

This figure is in color

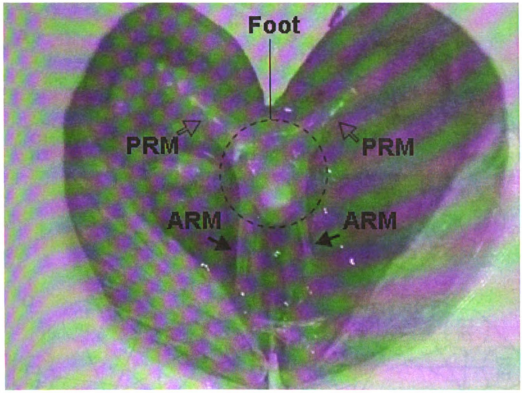


Figure 1-4 The anatomic structure of zebra mussel foot and retractor muscles.
The foot is circled by dashed line. The foot is connected with the anterior part of the shell through anterior retractor muscle (ARM) and anchored on the posterior side of the shell through posterior retractor muscle (PRM).

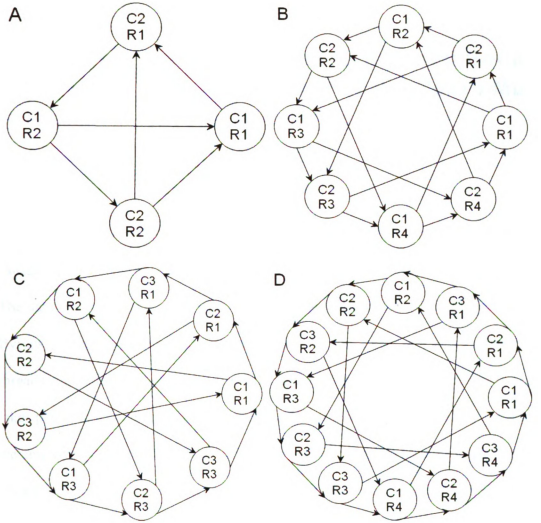


Figure 1-5 The non-reference interwoven loop design with different numbers of conditions and replicates.
A: the design for 2 conditions with 2 replicates in each condition; B: the design for 2 conditions with 4 replicates in each condition; C: the design for 3 conditions with 3 replicates in each condition; D: the design for 3 conditions with 4 replicates in each condition.

CHAPTER II

Putative Identification of Expressed Genes Associated with Attachment of the Zebra Mussel (*Dreissena polymorpha*)

Abstract

Through its firm attachment to substrata, the zebra mussel (*Dreissena polymorpha*) has caused severe economic and ecologic problems since its invasion into North America. The nature and details of attachment of this nuisant mollusk remain largely unexplored. Byssus, a special glandular apparatus located at the root of the mussel's foot produces threads and plates through which firm attachment of the mollusk to underwater objects takes place. In an attempt to better understand the zebra mussel's adhesion mechanism, we employed the suppression subtractive hybridization (SSH) assay to produce a cDNA library with genes unique to the mussel's foot. Analysis of the SSH cDNA library revealed the presence of 750 new expressed sequence tags (ESTs) including 304 contigs and 446 singlets. Using BLAST search, 365 zebra mussel ESTs showed homology to other gene sequences with putative functions. The putative functions of the homologues included proteins involved in zebra and blue mussels' byssal thread formation, exocrine gland secretion, host defense, and housekeeping. The generated data provide, for the first time, some useful insights into the foot structure of the zebra mussel and its underwater adhesion.

Introduction

Originally from Eurasia, the zebra mussel (*Dreissena polymorpha*) invaded the Great Lakes in North America in the early 1980s causing economic and ecologic devastation. Through its extraordinary ability to attach to hard surfaces underwater, the zebra mussel has blocked water intake pipes for industrial facilities, eroded underwater structures, and destroyed native mussel species (Johnson & Padilla 1996). The zebra mussel attaches to hard surfaces by the byssus, a structure embedded in the mussel's foot and consists of excretory glands and byssal threads that end with adhesive plaques (Morton 1993). There is a dearth of knowledge on the biochemical and physiological processes leading to byssus formation, attachment and detachment in the zebra mussel (Frisina & Eckroat 1992; Bonner & Rockhill 1994b). Moreover, there seems to be a large number of proteins involved in byssus formation and maintenance. The nature of proteins associated with the byssal apparatus structure and their functions have not been thoroughly studied due to the technical difficulties associated the isolation and purification of proteins using standard biochemical procedure. This study is the first step needed to identify the genes encoding byssal proteins. To achieve this goal, the suppression subtractive hybridization (SSH) assay was employed to enrich genes encoding proteins unique to the zebra mussel foot.

Materials and Methods

Zebra mussel tissue source, total RNA isolation and mRNA purification

Zebra mussel samples were collected from the Huron River in Ann Arbor, MI, USA (Latitude: 42.270N; Longitude -83.726W), transported alive in river water to the Aquatic Animal Health Laboratory at Michigan State University, East Lansing, Michigan. This sampling site was selected due to the absence of the quagga mussel (U.S. Geological Survey Investigation 2007). The mussels were then identified according to their morphological criteria as described elsewhere (Pathy & Mackie 1993). The mussels were then allowed to acclimate for eight weeks in glass aquaria kept at room temperature and supplied with well water. While the majority of the mussels were allowed to attach by placing them on the ventral side, the rest of mussels were placed on their dorsal side and stayed, therefore, unattached. Both attached and detached mussels were fed a pure culture of the algus *Ankistrodesmus falcatus* once a week. Five individual mussels from each attachment status were randomly selected to develop cDNA libraries that emcompass genes expressed in both attached and detached status. Shell length (measured from umbo to the opposite shell margin) of selected mussels ranged from 2 to 3 cm. Two types of tissues were excised from the zebra mussel: the entire foot, after removing byssal threads (ZMF), which consists of byssal glands and muscular tissues (Rzepecki & Waite 1993b, a); and the retractor muscles (anterior and posterior, ZMR, Figure 2-1). The Samples were immediately cryo-preserved in liquid nitrogen until RNA extraction. Total RNA was extracted with the RNeasy Protect Midi Kit (Qiagen, Valencia, CA) following the manufacturer's protocol. All the Poly A+ RNA (mRNA) templates were collected from the total RNA solution.

Construction of the suppression subtractive hybridization cDNA library

BD Clontech™ PCR-Select™ cDNA Subtraction Kit and PCR-Select™ cDNA Subtraction Kit (BD Biosciences, Palo Alto, CA) were applied in the construction of the cDNA library following the manufacturer's protocols. In this study, the ZMF samples were used as the testers that were connected to the specific adaptor offered by the kit before the first hybridization, and ZMR was treated as the driver, which was not connected to any adaptor. After two cycles of hybridization and nested PCR, only those expressed sequence tags (ESTs) that appeared in ZMF (tester) but not in the ZMR (driver) were maintained and enriched. The cDNA from the library was cloned using the QIAGEN PCR Cloning*Plus* Kit (Qiagen), which included the vector, T4 DNA ligase, and the competent cell strain for cDNA cloning. All the steps of the ESTs ligation, recombinant transformation and cells incubation were performed as described in the manufacturer's protocol.

ESTs sequencing

The *Eschrechia coli* cells were screened with isopropyl-β-D-thiogalactopyranoside (IPTG) and X-gal in Luria Broth (LB) agar plates (Invitrogen, Carlsbad, CA). The clones with white colors were picked for inoculation with LB under 200 rpm 37℃ overnight. The inoculated bacteria were used for sequencing. All the ESTs were sequenced using the Applied Biosystems 3730xl DNA Analyzer (Applied Biosystems, CA). The sequencing data are processed automatically and quality values are assigned using the Phred algorithm (Ewing & Green 1998; Ewing *et al.* 1998). The vector sequences were masked using the vectors database from GenBank.

Data Analysis

The assembling and clustering of ESTs were performed using CAP3 (Contig Assembly Program) (Huang & Madan 1999). Homologous gene searching was conducted using local BLAST (http://www.ncbi.nlm.nih.gov/BLAST) in conjunction with the GenBank protein database. All homologous sequences were confirmed using their respective E-value. All results with E-values less than 10^{-5} were considered highly homologous. The putative functions of the ESTs that had homologous genes within the GenBank database were predicted based upon the functions of their homologous genes.

The proteins encoded by identified ESTs were annotated to Gene Ontology (GO) terms with software package Blast2GO (Conesa *et al.* 2005). The GO terms for the annotated proteins were based on their molecular functions.

Results

ESTs sequencing and assembling

Sequencing and primary data analysis yielded 2336 ESTs (Table 2-1). The average read length for analyzed ESTs, using the Applied Biosystems 3730xl DNA Analyzer, was 673 bp. The 2336 ESTs were assembled and the results from CAP3 analysis are displayed in Table 2-1; 304 contigs were assembled with 1890 single ESTs, leaving 446 singlets that did not have overlaps with other ESTs.

Homologous gene searching and putative function prediction

Homologous genes for the 304 contigs and 447 singlets were searched using the BLASTx program. Among the results, 365 well-assembled ESTs hit homologous genes present within the GenBank database with putative functions. The remaining 385 ESTs either did not have homologous proteins or had no reported putative functions within GenBank. As displayed in Figure 2-2, 3.0% of the assembled ESTs had putative functions related to byssal proteins and byssogenesis. 37.0% of the ESTs were homologous with tick exocrine gland-secreted proteins. 1.0% of the ESTs were predicted to have some host defense functions involved in different pathways based upon the function of their homologues. The final 6.0% of the ESTs with putative functions were grouped into proteins that have basic cell functionality or with housekeeping gene characteristics. All the 725 assembled ESTs with length larger than 100 bp were submitted to GenBank and were assigned accession numbers from AM229723 toAM230448.

Of the sequences with putative byssal proteins and byssogenesis functions, 21 had homology with the zebra mussel byssal protein Dpfp1 (GenBank accession no.

AAF75279) (Table 2-2). The other ESTs in this group had homology with a number of proteins within GenBank that may be involved in zebra mussel attachment, such as the polyphenolic adhesive protein (S23760) from *Mytilus edulis* (Table 2-3). In the group of putative exocrine gland proteins, 110 ESTs (accession numbers in GenBank from AM229751 to AM229861 except for AM229807) had high homology with the western black-legged tick (*Ixodes pacificus*) transcriptome of the salivary glands NPL-2 (AAT92111) and the E-values of the homology are from 3×10^{-6} to 5×10^{-11}; 169 ESTs (The accession numbers of this group of ESTs are from AM229863 to AM230033 except for AM229950 and AM229951) were homologous to the black-legged tick, *Ixodes scapularis*, putative secreted salivary gland peptide (AAV80789) with the E-values from 7×10^{-6} to 6×10^{-12}; two of the ESTs in this group (The accession numbers of these two ESTs are AM230043 to AM230044) had homology with *I. scapularis* putative salivary protein that contains GYG repeats (AAY66521) and their E-values are 2×10^{-9} to 6×10^{-7} respectively.

Six ESTs showed sequence homology homologous to Rhinoceros beetle (*Oryctes rhinoceros*) defensin precursor (O96049), as well as some other host defense related proteins including different types of protease and protease inhibitors (Table 2-4). The other 122 ESTs with known putative functions were either housekeeping genes or involved in normal cell functions.

The annotation of the ESTs based on their molecular functions showed that eight ESTs encoded proteins that were involved in electron carrier activity. Another eight ESTs encoded binding proteins. Five proteins were predicted to have structural molecule

activity. Only one protein was classified into each of the following categories: receptor activity, receptor binding, transportor activity, molecular transducer activity, protein binding, and signal transducer activity. Another 21 proteins were found to have certain molecular functions (Figure 2-3). The graphic combination of the annotation displayed the relation between categories and the score of each node was shown in Figure 2-4.

Discussion

Since the SSH was applied, it is assumed that the generated cDNA library was dominated by the foot transcripts including those involved in the formation of byssal threads. Numerous ESTs in our cDNA library exhibited very high homology with the zebra mussel foot protein, Dpfp1, which is present in secretory granules within a gland that surrounds the ventral groove of the mussel's foot (Benedict & Waite 1986). The degree of homology with Dpfp1 varied among ESTs; while some showed extraordinarily high homology (e.g., BG12_F05 and BG22_E08 whose E-values were 3×10^{-69}) others exhibited barely significant homology (e.g., BG_CON_98 whose E-value was $2x10^{-8}$). Interestingly, the 24 sequences that exhibited high homology with the Dpfp 1 gene did not share the same sequence as was apparent from the cluster analysis. According to previous studies, the Dpfp1 protein is composed of tandemly repeated and segregated motifs (Anderson & Waite 1998). However, the absence of the overlaps among the 24 Dpfp1-like ESTs of our study demonstrated that these ESTs are probably different variants of byssal proteins rather than different tandem repeats of a single gene. This phenomenon can also be explained by *Mytilus edulis* foot protein-3 precursor (Mefp3) model, which encodes the major protein of the blue mussel adhesive plaque. Variants of Mefp 3 have been reported to exist due to the alternative splicing and RNA editing of the gene (Warner & Waite 1999). However, findings of this study cannot verify that if the multiple copies of Dpfp 1 in our library represent a number of variants or just the tandem repeats from a single gene. Also, from this study, it is apparent that zebra and the marine blue mussels (*Mytilus edulis*) share a number of similarities in their mechanisms of

underwater attachment. Two ESTs (BG13_D08 and BG02_A07) of the zebra mussel exhibited homology with the polyphenolic adhesive protein 1 of *M. edulis*, which is considered a major player in the blue mussel underwater attachment in the marine environment (Waite 1983b).

Another important finding of the current study is the relatively high number (37.0%) of zebra mussel ESTs that exhibited high homology to salivary gland peptides of two *Ixodes* spp. In the absence of knowledge about the functions of both families of proteins, it is impossible to make any inference about the functions of the zebra mussel molecules that resemble the tick salivary gland proteins. The zebra mussel foot harbors a number of exocrine glands, which are probably similar to salivary glands in terms of secretion and regulation. Whether these proteins form byssal threads or control their production remains to be elucidated.

It seems that the zebra mussel guards its foot and the byssus apparatus in the microbe-rich aquatic environments with a number of molecules whose homologues are involved in host defense mechanisms (Table 2-4). Homologues of antimicrobial peptides (e.g., defensin) may be vital in protecting byssal threads from bacterial degradation. Defensin, in particular, has been found in other mollusks such as the Mediterranean mussel (*Mytilus galloprovincialis*) (Yang *et al.* 2000), the American oyster (*Crassostra virginica*) (Seo *et al.* 2005) and the bay scallop (*Argopecten irradians*) (Zhao *et al.* 2007). These Molluscan defensins have high sequence homology with arthropod defensins (Seo *et al.* 2005). Additional defensive molecules, such as tumor suppressor factors and interleukin enhancer, may also be of importance to the zebra mussel's innate immunity, similar to those found in other bivalve species (Wiens *et al.* 1999).

Protease and protease inhibitors were identified in the zebra mussel SSH cDNA library (Table 2-4). Their exact role in the foot or in byssogenesis is currently unknown. The zebra mussel ESTs included also genes involved in signal transduction, cell division and development, metabolism, cell structure maintenance, and DNA/protein synthesis. Since the zebra mussel byssus is an extremely complex system, composed of a number of tissues including glands, it's very likely that the byssus employs multiple genes with various functions. Unlike the other molluscans, such as oysters (Ostreidae) and scallops (Pectinidae), there is very limited data available on this important organism.

In general, data obtained from this study constitute the first basis of knowledge gained and current studies are directed toward identifying their potential functions. Moreover, zebra mussel ESTs, with known and unknown potential functions, are currently being assembled into a microarray to further determine their roles in byssogenesis.

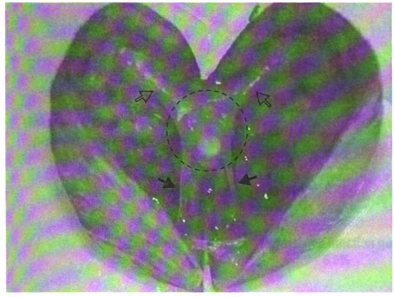


Figure 2-1 The structure of zebra mussel foot and retractor muscles.
The organ within the dotted circle is the foot (the source of ZMF-RNA). Black arrows point to zebra mussel anterior retractor muscles while the non-filled arrows point to the posterior retractor muscles. ZMR RNA was extracted fom both anterior and posterior muscles.

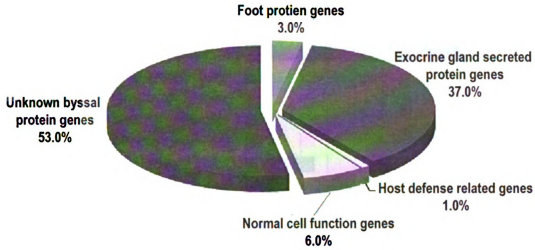


Figure 2-2 Percentage of ESTs with different possible functions.

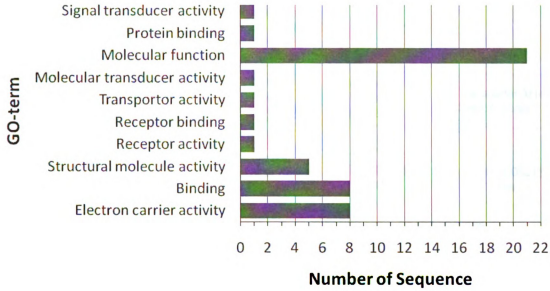


Figure 2-3 The annotation of proteins encoded by ESTs from cDNA library.
The proteins translated from the ESTs in the library were annotated and classified into 10 groups based on their molecular functions.

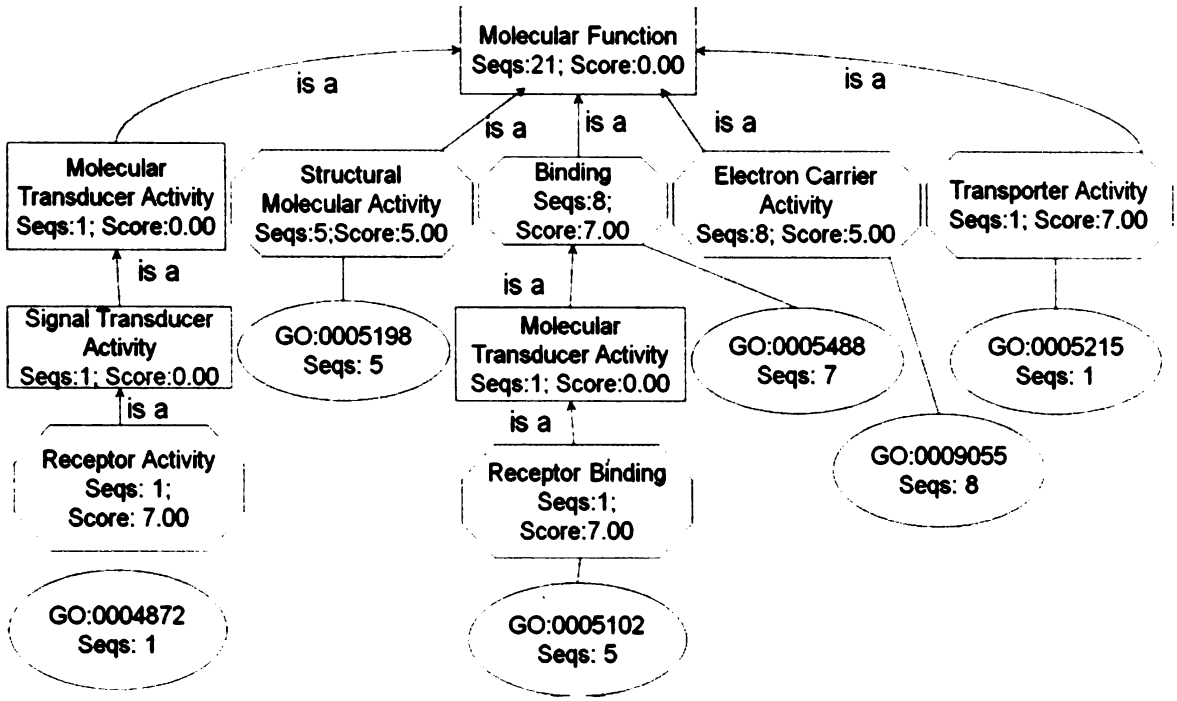


Figure 2- 4 The combined graphic of annotation of the zebra mussel byssus unique genes. The number of sequences and the annotation score are labled in each node.

Table 2-1 Summary of the ESTs sequenced from the SSH cDNA library.

ESTs	Overall
Number of ESTs	2336
Average length from the sequencer	673bp
Number of contigs[a]	304
The number of ESTs contained in the contigs	1890
Number of singlets[b]	446
Total ESTs after assembling	750

[a] The sequences assembled by many single ESTs with overlaps.

[b] The ESTs have no overlaps with any others in database.

Table 2-2 The ESTs with homology to zebra mussel (*D. polymorpha*) byssal protein Dpfp1.

Zebra mussels Clone IDs of this study (contigs)	Accession # given by EMBL	E-value	Zebra mussels Clone IDs of this study (singlets)	Accession # given by EMBL	E-value
BG_CON_212	AM229723	2.00E-63	BG12_F05	AM229736	3.00E-69
BG_CON_252	AM229724	3.00E-38	BG22_E08	AM229737	3.00E-69
BG_CON_279	AM229725	5.00E-26	BG09_G08	AM229738	3.00E-21
BG_CON_181	AM229726	8.00E-26	BG17_H08	AM229739	1.00E-12
BG_CON_70	AM229727	3.00E-20	BG13_D02	AM229740	4.00E-12
BG_CON_203	AM229728	1.00E-18	BG07_G07	AM229741	7.00E-12
BG_CON_149	AM229729	1.00E-12	BG08_F07	AM229742	2.00E-11
BG_CON_193	AM229730	5.00E-12	BG09_C04	AM229743	3.00E-11
BG_CON_135	AM229731	8.00E-10	BG15_D06	AM229744	5.00E-10
BG_CON_156	AM229732	1.00E-9	BG06_C02	AM229745	2.00E-9
BG_CON_98	AM229733	2.00E-8			

Table 2-3 The other foot proteins with putative potential attachment functions.

Zebra mussels Clone IDs of this study	Accession # given by EMBL	Functions of homologous proteins	IDs of homologous genes from GenBank	E-value
BG13_D08	AM229747	Polyphenolic adhesive protein 1	S23760	1.00E-9
BG02_A07	AM229748	Blue mussel (*Mytilus edulis*)		2.00E-7

Table 2-4 ESTs and their homologous proteins with putative host defense functions.

Zebra mussels Clone IDs of this study	Accession # given by EMBL	Functions of homologous proteins	IDs of homologous genes from GenBank	E-value
BG11_E08	AM230099	Defensin precursor Rhinoceros beetle (*Oryctes rhinoceros*)	O96049	4.00E-6
BG10_D04	AM230107	Hemicentin Human (*Homo sapiens*)	AAK68690	6.00E-18
BG23_A03	AM230108	Putative tumor suppressor Marine sponges (*Suberites domuncula*)	CAC80049	8.00E-48
BG_CON_268	AM230114	Kunitz-like protease inhibitor Dog hookworm (*Ancylostoma caninum*)	AAN10061	4.00E-8
BG_CON_124	AM230115	Tissue factor pathway inhibitor Rhesus monkey (*Macaca mulatta*)	AAB31955	2.00E-8
BG02_E12	AM230116	Cathepsin L precursor Flesh fly (*Sarcophaga peregrina*)	BAA03970	1.00E-9

CHAPTER III

Development of a cDNA Microarray of Zebra Mussel (*Dreissena polymorpha*) Foot and its Use in Understanding the Early Stage of Underwater Adhesion

Abstract

The underwater adhesion of the zebra mussel (*Dreissena polymorpha*) to substrates is a complex process that is controlled by a delicate apparatus, the byssus. As a critical activity of the byssus glands embedded in zebra mussel feet, byssogenesis is highly active to produce numerous byssal threads from the settled juvenile stage through the adult stage in its life cycle. This lifelong activity helps the zebra mussel to firmly attach to substrata underwater, thereby causing severe economic and ecologic impacts. In an attempt to better understand the zebra mussel's byssus activity, a cDNA microarray (ZMB) including 716 genes, generated from a Suppression Subtractive Hybridization (SSH) cDNA library, was printed and used for the comparison of gene expression during zebra mussel adhesion and non-adhesion. To better understand the byssogenesis mechanism, RNA samples from the zebra mussel feet with byssogenesis and without byssogenesis were used in a two-color hybridization to reveal the gene differential expression in the two states. Based on the *P* values ($P < 0.05$), Fifty-two ESTs were found as differentially expressed genes and were divided into two groups, upregulated and downregulated group according to there logFC values. With the false discovery rate (FDR) adjustment, seven were identified from the upregulated group and nine from the

downregulated group. Phylogenetic analysis indicated that the four excretory gland peptide-like proteins (EGP) encoding genes in the upregulated group are structurally different than the two in the downregulated list. The amino acid composition analysis on the proteins, which were encoded by the up- or downregulated ESTs without homologues (NH) suggested that seven of the NH proteins are biochemically similar to the novel foot proteins from other mussels. The quantitative reverse transcription PCR (QRT-PCR) proved the uniqueness of the templates in the array, and also confirmed the differentially expressed genes identified by microarray experiment. Our findings demonstrated that the zebra mussel byssus cDNA microarray is an efficient tool for the studies of differential gene expression in different byssogenesis states, thereby revealing important details of the underwater adhesion.

Introduction

The underwater attachment of the zebra mussel (*Dreissena polymorpha*) to underwater substrates causes catastrophic ramifications in the fragile ecosystem of the Laurentian Great Lakes, USA. Attachment of the zebra mussel is primarily mediated by a structure called the byssus which is comprised from exocrine byssal glands embedded in the mussel's foot, byssal threads located at the root of zebra mussel foot, and adhesive proteins forming a plaque at the distal end of the threads through which adhesion takes place (Waite 1992; Rzepecki & Waite 1993a, b). The unique characteristics of the byssus have helped the zebra mussel survive a number of physical, chemical and biological control methods. Indeed, in the absence of solid knowledge on the mechanism of zebra mussel attachment, an effective control strategy cannot be developed.

Current knowledge on the biochemical nature and functions of underwater adhesion by mussels is very scarce. Three families of *Dreissena polymorpha* foot proteins (Dpfp1-3) within the plaques of the byssal threads have been identified (Rzepecki & Waite 1993b, a). Dpfp-1 and -2 are believed to play a role in the maintenance of the byssal structures and have no adhesive properties (Anderson & Waite 2000). Studies performed on marine mussels of the genus *Mytilus* demonstrated the presence of 12 byssal proteins, eight of which are believed to play a major role in adhesion (Benedict & Waite 1986; Waite *et al.* 1998; Waite & Qin 2001; Sun & Waite 2005; Waite *et al.* 2005; Lin *et al.* 2007; Cha *et al.* 2008). How each of these proteins contributes to the underwater adhesion and potential differences in mussel attachment mechanisms between the freshwater and marine mussel species remains largely unexplored.

To better understand the molecular mechanism of zebra mussel attachment, a suppression subtractive hybridization (SSH) cDNA library was constructed using zebra mussel foot and muscle tissues (Xu & Faisal 2008) in which foot RNA was used to generate tester cDNA (ZMF) while the retractor muscle RNA was used as the templates for driver cDNA (ZMM). By subtracting ZMM from ZMF, all the cDNAs unique to the zebra mussel foot was assembled. This SSH cDNA foot-specific library entails 716 ESTs encoding proteins with various predicted functions such as adhesion, host defense and structure maintenance. Hence, the first aim of the study was to develop and validate a cDNA microarray that can be used to unravel mechanisms involved in byssogenesis of the zebra mussel.

Byssogenesis is influenced by a number of physical, chemical and biological factors, the most important of which is the presence of a suitable surface (substrate) to which the mussel can firmly attach (Clarke & McMahon 1996b, c; Clarke & McMahon 1996a; Clarke 1999). Indeed, it has been reported that the formation of zebra mussel byssal threads would not start unless the zebra mussel foot is tightly touching an underwater substrate (Eckroat *et al.* 1993). Whether the accessibility of the foot to attachment substrate triggers the synthesis of byssal thread proteins or modulates their release is yet to be elucidated. To this end, we report the use of the newly developed zebra mussel foot cDNA microarray to identify if access to substrates influences the differential expression of genes involved in byssogenesis.

Materials and Methods

Construction of zebra mussel foot cDNA microarray

The templates for microarray spotting were from the SSH cDNA library previously constructed in our laboratory(Xu & Faisal 2008). Accession numbers of the ESTs obtained from the library are from AM229723 to AM230448 in the GenBank. All ESTs were amplified using PCR before being spotted on the slides.

The preparation of DNA templates for microarray followed the protocols detailed in (Chandrasekharappa *et al.* 2003). Briefly, seven hundred and sixteen *Escherichia coli* clones were selected from thirty-four 96-well plates and rearranged into eight new 96-well plates. The overnight cultures from the eight 96-well plates were used to prepare the PCR templates. Ten microliters of each bacterial culture was taken to mix with 90 µl sterile double distilled water. The mixture was then incubated at 100 °C for 10 minutes followed by centrifugation for 15 min at 1200 *g*. The supernatant was used as PCR templates. M13 forward (5'- GTT TTC CCA GTC ACG ACG TTG -3') and reverse (5'- TGA GCG GAT AAC AAT TTC ACA CAG -3') primers were applied with GoTaq Green Mastermix (Promega Co., Madison, WI) to yield 100 µl PCR product for each EST. PCR products purification was done with the MinElute 96 UF PCR Purification Kit (Qiagen Inc., Valencia, CA). All purified PCR products were dissolved in 20 µl of 50% dimethyl sulfoxide (DMSO). One µl of each purified amplicon was loaded onto 1.2% agarose Tris-acetate-EDTA (TAE) gel to detect the presence of products. If a sample produced a faint band on gels, the concentration of this sample was determined using Qubit Fluorometer and Quant-iTTM dsDNA BR Assay Kit (Invitrogen, Carlsbad, CA).

The wells with template concentrations <100 ng/μl were not used and their contents replaced with freshly prepared templates whose concentration was ≥100 ng/μl. All purified amplicons were then transferred into two 384-well plates with the help of the Tecan Genosys 150 liquid handling robot (Tecan U.S., Research Triangle Park, NC) before microarray spotting.

The SuperAmine microarray substrate (ArrayIt®, TeleChem International Inc, Sunnyvale, CA) was used throughout this study. The cDNA microarray spotting was performed by the GeneTAC G3 arraying robot (Genomic Solutions, Ann Arbor, MI) equipped with a 48-pin head. The diameter of each pin is 200 μm. The zebra mussel foot cDNA microarray was designed to contain the 716 genes arranged in 48 sub-arrays (4 ×12). In each sub array, 64 spots were distributed as an 8 × 8 matrix. The 716 amplicons were spotted in triplicates, which are randomly located in each sub-array. Zebra mussel β-actin was used as the positive control spotted in every sub-array while the spotting solutions were used as negative controls in each sub-array. Meanwhile, 12 spots in each sub-array were left blank and considered as both background and negative controls (Figure 3-1).

cDNA microarray validation

This experiment was designed to ascertain that the ESTs used in the developed microarray are, indeed, more enriched in the foot tissue as compared to other major tissue types of the zebra mussel. Zebra mussels used for total RNA extraction were obtained from the Huron River in Ann Arbor, MI, USA (Latitude: 42°16'12''N; Longitude - 83°43'34"W). Hemolymph was extracted from the anterior adductor muscle sinus by using a 25G 5/8" needle inserted through the dorsal side of the mussel's shell. The

mussels were then dissected under a dissecting scope, and samples of foot, muscle, ctenidium, mantle, and gonad were collected separately. Two total RNA samples from each of the tissue sample were extracted with 5-PRIME PerfectPure RNA Tissue Kit (5 PRIME Inc, Gaithersburg, MD) and the concentrations were quantified by a Qubit Fluorometer (Invitrogen, Carlsbad, CA). Then samples were diluted to 10 ng/μl with RNase and DNase Free Water (Fisher Scientific, Pittsburgh, PA).

Eight genes were randomly selected for quantitative reverse transcription PCR (QRT-PCR). Table 3-1 displays the gene specific primers designed with software PimerExpress 2.0 (Applied Biosystem, Foster City, CA). As an internal reference, the 18S ribosomal RNA (rRNA) was used to eliminate potential errors brought by quantification. The QRT-PCR was applied with Brilliant SYBR Green QRT-PCR Master Mix (Stratagene, La Jolla, CA) in the real-time thermocycler, Mastercycler ep *realplex* S (Eppendorf, Westbury, NY). For each EST biological replicate, two technical replicates were applied. This gave four replicates for each reaction. Twenty nano gram total RNA from each sample was added in each PCR reaction. The program used for QPCR is: 30 minutes at 50 °C, 10 minutes at 95 °C, followed by 40 repeats of a 3-step temperature cycle (30 seconds at 95 °C, 1 minute at 55 °C and 30 seconds at 72 °C). The values generated by QRT-PCR were cycle threshold (Ct), which can be used as a parameter to indicate the relative abundance of the genes. The algorithm of gene differential expression was $2^{-\Delta\Delta Ct}$ as described by Livak and Schmittgen (Livak & Schmittgen 2001). The multiple comparisons were performed by Tukey's test after analysis of variance (ANOVA) test with SAS (SAS Institute Inc, Cary, NC).

False discovery rate (FDR) determination

Four RNA samples isolated from zebra mussel feet (designated T1, T2, T3, and T4) were used as biological replicates. T1 and T2 were labeled with Alexa555 and Alexa647 respectively and applied to the first array slide (MT1). T3 and T4, also labeled with Alexa555 and Alexa647 respectively, were applied to a second slide (MT2). MT1 and MT2 were used as replicates to determine FDR of the newly developed microarray.

Effects of the presence of attachment substrate on gene expression in the foot

Zebra mussels used in this experiment were also collected from the Huron River in Ann Arbor, MI, USA. The byssal threads of mussels were cut off with a sharp scalpel at the point of their emergence from the foot groove. Mussels were then randomly assigned to one of two groups: AD and NAD. Mussels in Group AD were allowed to attach to a matrix by placing them in glass petri dishes with their ventral sides facing down so that their feet have access to the glass. In the second group (NAD), mussels were placed in petri dishes on their dorsal sides; thereby giving their feet no access to a substrate to which they can attach. Petri dishes, containing both AD and NAD mussel groups were placed in an aquarium, supplied with 0.22 μm filtered Huron River water, and kept at room temperature. The mussels in both groups were fed daily with a pure culture of the alga *Ankistrodesmus falcatus*.

The generation and growth of byssal threads after their severing was observed and recorded. On days 1, 2, 4, 5, 6, 7, 8, 9, 12, 14, and 16, six mussels from each group were randomly selected and their byssal threads counted under a dissecting microscope.

In order to identify genes differentially displayed due to the ability of the mussel to attach, total RNA was extracted from mussels' feet 48 hours post-treatment as described above.

In each group, total RNA extraction involved four biological replicates, with four individual feet pooled in each replicate. The concentration and the integrity of the RNA samples were detected by 2100 Bioanalyzer (Quantum Analytics, Foster City, CA). Started with 10 μg total RNA, the cDNA synthesis and dye labeling were performed by SuperScript *Plus* Indirect cDNA Labeling Kit (Invitrogen) containing Alexa Fluor 555 & 647 dyes. The Oligo(dT) primer was used in cDNA synthesis. In each group, two cDNA replicates were labeled with Alexa 555 while the other two were labeled with Alexa 647. One cDNA sample from each group labeled with different dyes were mixed, therefore, two replicates for each combination were obtained: the combination of Alexa 555 labeled AD-cDNA and Alexa 647 labeled NAD-cDNA, as well as the combination of Alexa 647 labeled NAD-cDNA and Alexa 555 labeled AD-cDNA. The design for this study can be described as the following:

Slide 1: AD1 → NAD1; Slide 2: AD2 ← NAD2;

Slide 3: AD3 → NAD3; Slide 4: AD4 ← NAD4.

The start point of the arrow stands for being labeled with Alexa555 while the end point of the arrow means being labeled with Alexa647. Each mixed aliquot was transferred to a Microcon YM-30 Centrifugal Filter Unit (Millipore, Billerica, MA) to be concentrated. Before the samples were loaded on the GeneTAC HybStation (Genomic Solutions), 110 μl SlideHyb buffer 3 (Ambion, Austin, TX) was added to each concentrated mixture and subsequently incubated in 70 °C for 5 minutes. An 18-hour step-down protocol was applied to all hybridizations; briefly, 3 hours at 60 °C, 3 hours at 55 °C, 12 hours at 50 °C. The washing procedure followed the hybridizations, and was performed with three different washing buffers containing SSC buffer (Ambion) and sodium dodecyl sulfate

(SDS) under different conditions: 30 seconds in each cycle at 50 °C for three cycles with medium stringency wash buffer (2 × SSC and 0.1% SDS), 30 seconds in each cycle at 42 °C for three cycles with high stringency wash buffer (0.2 × SSC and 0.1% SDS), and 30 seconds in each cycle at 42 °C for three cycles with post wash buffer (0.2 × SSC). All slides were rinsed once in 0.2 × SSC and once in double distilled water after they were downloaded from the HybStation, and immediately dried by centrifugation at 1,000 *g* for two minutes.

Hybridized cDNA microarrays were scanned by GenePix 4000B two-laser Scanner (Molecular Devices, Downingtown, PA), and GenePix Pro 6.0 (Molecular Devices) software was then used for image processing and spot intensity file creation. The data obtained from the microarray and used in the following analysis have been deposited in Gene Expression Omnibus (GEO) at the National Center for Biotechnology Information (NCBI), with the series accession number GSE10851.

Microarray data analysis

The four spot density files output from GenePix Pro 6.0 were analyzed with Limma (Wettenhall & Smyth 2004) software package. The raw spot intensity data were normalized within and between the slides. Least square regression was also performed to determine the differentially expressed ESTs and their fold change. The *P* values were adjusted by Benjamini-Hochberg (BH) correction. The significance of the differential expression was also decided by *P* values and the FDR as described above.

Excretory gland peptide (EGP)-like zebra mussel protein amino acid analysis

This analysis was performed in order to determine the difference between the up- and downregulated EGP-like genes. EGP-like ESTs, from the up- and downregulated groups, were translated into amino acid sequences, a phylogenetic tree of the seven sequences constructed using Mega4 software (Tamura *et al.* 2007), and phylogenetic analysis performed with the Neighbor-Joining algorithm. Multiple alignments of the EGP-like peptides were done by ClustalW (Thompson *et al.* 1994).

Amino acid analysis of differentially expressed genes without putative function

Twenty-one differentially expressed ESTs without putative hits in both the up- and downregulated groups were translated with the aid of the proteomics and sequence analysis tools of the Expert Protein Analysis System (Gasteiger *et al.* 2003). Six translation frames were obtained from each EST, with the longest continuous amino acid sequence without a stop codon chosen for further analysis. Corrections were made manually and the percentage of each amino acid residue in each sequence calculated by Protein Stat in Sequence Manipulation Suite (Stothard 2000).

Validation of selected differentially expressed genes

QRT-PCR was utilized to verify the differentially expressed genes identified from the cDNA microarray. For this experiment, additional AD and NAD groups were created as described above. Total RNA samples were extracted from both groups two days after the mussels settled and were labeled as AD-2 and NAD-2. The QRT-PCR was also performed as described above. In this QRT-PCR assay, the NAD samples were treated as

control and the relative abundance of AD-2 was calculated also with the $2^{-\Delta\Delta Ct}$ algorithm as explained above. Four replicates for each sample were used in this experiment (two biological replicates with two technical replicates in each biological replicate). Statistical analysis was also done by ANOVA followed by Tukey's test (SAS).

Results

The tissue specificity of microarray templates

To test if the templates spotted on this zebra mussel byssus microarray was unique to zebra mussel foot. Eight ESTs that were randomly selected from the array were detected within different tissues of zebra mussel by QRT-PCR. As displayed in Figure 3-2, the QRT-PCR demonstrates that the eight selected differentially expressed genes were much more expressed in the zebra mussel foot than in other tissues tested. The differences in transcription levels of the selected genes among the tissues were highly significant ($P< 0.01$).

Byssogenesis in group AD and NAD

None of the mussels in the NAD group grew byssal threads during the 16-day observation period. Mussels in the AD group grew byssal threads that averaged 14.0 $\pm$ 1.8 (standard error) per individual within one day to 104.2 $\pm$ 3.9 threads/individual after 16 days of observation (Figure 3-3). Regression analysis indicated that the trend of the increase of byssal threads in AD group fits the linear model very well with the correlation coefficient $R^2 = 0.9705$. The ANOVA analysis with Tukey's test indicated that Day 2 is the earliest stage with the number of byssal threads significantly different than Day 0 ($P < 0.05$). Except for Day 1, the numbers of byssal threads at other time points are all significantly more than Day 0 ($P < 0.05$). The numbers of byssal threads increased significantly within every two days ($P < 0.05$), while it is mentionworthy that within the observation period the number of new byssal threads plateaued between the 7th and the

9th day (P>0.05), then continued to increase throughout the remainder of the observation period (Figure 3-3).

FDR determination

Results obtained from the two microarrays hybridized by differentially labeled AD-cDNA samples demonstrated that the FDR is much lower than expected at each P value level. If the P value was set at 0.01 as criteria, only one false positive gene was obtained (Figure 3-4), suggesting that our design of the ZMB microarray, combined with LOESS normalization and background subtraction, yields a very low rate of false positives.

Effects of adhesion status on gene expression

Data obtained from hybridized zebra mussel cDNA microarray using the SSH library and analyzed with the limma revealed 52 differentially expressed genes in the AD Group with $P < 0.05$. Twenty-four genes were upregulated (i.e., logFC > 0, Table 3-2), and 19 genes were downregulated (i.e., logFC < 0, Table 3-3). Using the cutoff of P<0.01, based on FDR determination experiment, seven upregulated and nine downregulated ESTs were found differentially expressed genes ($P < 0.01$). In tables 3-2 and 2b, the positive log values show the upregulated genes highly expressed in AD samples, while the negative ratios indicated the downregulated genes that are more abundant in NAD samples. Among the upregulated genes, three of them were homologous to excretory gland peptides (EGP) identified from western black-legged tick (*Ixodes pacificus*, GenBank accession no. AAT92111) or the black-legged tick (*I. scapularis*, GenBank accession no. AAV80789). One of the upregulated ESTs were homologous to shematrin

4 of the pearl oyster, *Pinctada fucata*. The rest of the upregulated genes had no homologues in GenBank database. The downregulated group had two sequences homologous to EGP of *I. scapularis* and one EST homologous to polypeptide release factor 3 from the yeast, *Yarrowia lipolytica*. The other downregulated ESTs had no putative functions based on BLASTx search.

With the cutoff of $P < 0.05$, twenty-four ESTs were found upregulated in AD group while 28 ESTs were classified as downregulated. Among the 24 upreguated genes, fifteen of them had no homologues in GenBank. Nine of the upregulated sequences have putative functions based on their homologues. AM229726 was found homologous to zebra mussel foot protein 1 (Dpfp-1) (Table 3-2). Among the downregulated genes, eighteen of were not homologous to any sequences in GenBank database (Table 3-3).

EGP-like sequences analysis

As displayed in Figure 3-5, the four upregulated EGP-like sequences clustered together, while the two downregulated EGP-like sequences clustered together. AM229866, AM229964, AM229813, and AM229911 showed the closest relatedness in the upregulated EGP-like ESTs clade. The AM229892 and AM229917 formed the other clade. Multiple alignments of the six EGP-like EST sequences demonstrated a similar dichotomy between the four upregulated EGPs and the two downregulated sequences. Within the sequences of the six EGP-like ESTs, The bases at both 3' and 5' ends exhibited more variance than those within the sequences (Figure 3-6).

Validation for microarray results

As displayed in Figure 3-7, QRT-PCR confirmed that AM229866, AM230104, and AM229726 were significantly upregulated (P <0.05) in the AD-2 group, while the AM230072, AM230157, and AM230114 were dramatically downregulated (P < 0.05). This was in accordance with the results from microarray analyses.

Discussion

The cDNA microarray developed along the course of this study is unique for a number of reasons. First, it is the first of the zebra mussel, a nuisance species that is causing severe economic and ecologic consequences. There are only a handful of DNA microarrays from other bivalve mollusks, such as a cDNA microarray of the American (*Crassostrea virginica)* and Pacific (*C. gigas*) oysters (Jenny *et al.* 2007), that consists of 27,496 ESTs obtained from sequences deposited in the GenBank. Second, it is the first microarray specifically designed to study underwater attachment mechanisms at the molecular level. Recently, a relatively small low-density oligonucleotide microarray has been constructed from 24 ESTs of *Mytilus* spp. from sequences deposited in the GenBank and used to determine gene expression levels in response to pollution stresses (Dondero *et al.* 2006). The ESTs in this *Mytilus* microarray are not related to the foot function. Last, the zebra mussel foot cDNA microarray was constructed based on an SSH cDNA library, which allowed the enrichment of foot-specific expressed genes, therefore, the likelihood of this microarray reveals novel attachment mechanisms is high. The validation experiment performed in this study (Figure 3-2) is evidance that the ESTs of the zebra mussel cDNA microarray are, indeed, highly expressed in the foot and not (or much less) in other mussel tissues.

False positive results have always been a problem in analyzing microarray data. The problems stem from a number of factors such as the proportion of truly differentially expressed genes, distribution of the true differences, measurement variability, and sample size (Pawitan *et al.* 2005). Among the tools used to control false positive rate, FDR correction is the most common statistical method. However, the FDR correction usually

gives very high adjusted *P* values when the sample of the experiment is small (Pawitan *et al.* 2005). In our study, when the BH correction was implemented, most adjusted *P* values increased substantially (table 3-2 and 3-3). It is likely caused by the small sample size of this study (n = 4). To avoid this problem, Nobis et al. (2003) suggested hybridization of a microarray slide with two identical samples, a step that helps determining the actual false discovery rate. When Nobis et al. (2003) protocol was applied in this study, 13 genes were found falsely identified as differentialy expressed at $P<0.05$; while at $P<0.01$, only one falsely positive gene was found. Therefore, subsequent analyses of the AD-NAD microarray experiment used $P < 0.01$ as the cutoff. This modification in analysis revealed that 16 genes (seven upregulated and nine downregulated) out of the 52 differentially expressed genes obtained using $P<0.05$ as the cutoff value (marked with * in Table 3-2 and 3-3). This method certainly increased the specificity of the microarray in finding the adhesion associated genes on one hand, however it decreased its sensitivity as shown by qRT-PCR assay of the EST AM229726 and AM230114 which were rejected by $P < 0.01$ but are definitely associated with adhesion (Figure 3-7).

Experiments performed in this study also demonstrated that the presence of a suitable substrate for attachment is vital for byssogenesis. In the AD mussel group, byssal threads grew in as early as two days and increased thereafter, while in the NAD mussel group, no byssal threads were formed over the entire observation period. Statistical analysis between the AD and NAD mussel groups indicated that the difference in the number of newly generated byssal threads becomes statistically significant as early as two days post-treatment, even though there have been no threads formed in the NAD mussel group. Based on this finding, samples were collected and microarray analysis was

performed to determine the differences in gene expression between the AD and NAD mussel groups at the early stage of byssal thread regeneration, an important stage of zebra mussel attachment.

Interestingly, the genes identified from this microarray represent a wide range of proteins with different putative functions. Some genes are known for their involvement in adhesion, such as EGP-like sequences (AM229866, AM229964, AM229911, AM229892, and AM229917 in Table 3-2 and 3-3) that are reported to function as the main adhesive host defense molecules in black tick's salivary gland (Francischetti *et al.* 2005; Narasimhan *et al.* 2007). The shematrin-like molecule is homologous to a shematrin isolated from the mantle of a pearl oyster *Pinctada fucata*, providing a framework for shell classification (Yano *et al.* 2006). The actual function of this shematrin-like protein of the zebra mussel remains unknown; however, the genes exclusively expressed in zebra mussel foot indicated that its function is more likely to be related to foot activity rather than to shell classification.

Surprisingly, none of the zebra mussel foot proteins, originally identified in the SSH cDNA library (Xu & Faisal 2008) were differentially expressed in this early phase of byssogenesis. This can be attributed to a number of reasons. First, it is possible that these genes become differentially expressed at a later stage of byssogenesis (i.e., later than 48 hours post-treatment). Indeed, some of these proteins function as links and dovetails between adjacent structural proteins, and therefore will only be needed at a later stage. For example, two proteins (preCol-P) have been identified from the blue mussel *M. edulis*, with one being distributed in the foot and the other joining the proximal threads to the byssal stem (Coyne & Waite 2000). Second, there is a possibility that these proteins

are constitutively expressed and their encoded proteins accumulated in the byssal glandular cells, yet their excretion to form new threads is triggered, directly or indirectly, by the presence of adhesion substrate. As previously reported (Eckroat *et al.* 1993; Anderson & Waite 2000), byssal proteins are produced prior to the thread formation, stocked in the foot's ventral groove, and then released upon attachment. Last, the accessibility to the substratum may not be the only triggering factor to stimulate the differential expression of foot proteins. Environmental factors, such as temperature, dissolved oxygen, current velocity and food availability are known to play an important role in byssal thread formation (Clarke & McMahon 1996b, c; Clarke & McMahon 1996a; Clarke 1999). On the other hand, some genes without homologues (NH) in our microarray exhibited the highest fold increase as demonstrated by microarray analysis and RT-PCR. This suggests that the transcripts of these genes are very likely involved in byssal thread regeneration 48 hours post-treatment. These NH protein fragments share some characteristics to foot proteins identified from marine or freshwater mussels in the amino acids composition. For example, in the novel Dpfp1, the proline, tyrosine, aspartic acid/asparagine, lysine, threonine, and glycine residues together account for more than 50% of the amino acid composition (Anderson & Waite 1998). It is also noted in other marine byssal precursors, such as *M. edilus* foot protein 1 (Filpula *et al.* 1990), *M. galloprovincialis* foot protein 1 (Inoue & Odo 1994), *M. coruscus* foot protein 1 (Inoue *et al.* 1996), and, to a lesser extent, *M. galloprovincialis* foot protein 2 (Inoue *et al.* 1995). Some other foot proteins, such as *M. californianus* foot protein 3 have amino acid compositions dominated by glycine, asparagine and tyrosine (Zhao *et al.* 2006). The high content of these amino acids in the sequences listed in Table 4 suggests that one of the

NH molecules upregulated in the early stage of byssogenesis may have similar biochemical characters with the novel foot proteins.

Homologous to a salivary gland peptide of *Ixodes pacificus* (Francischetti *et al.* 2005) and *I. scapularis* (Narasimhan *et al.* 2007), the EGP-like proteins were found in both up- and downregulated molecule groups. According to Fracischetti et al, the salivary gland peptides isolated from *I. pacificus* can be divided by 16 groups that are functionally related to anti-hemostatic, anti-coagulant, anti-microbial, oxidant metabolism, and housekeeping (Francischetti *et al.* 2005). The EGP-like proteins identified by microarray are all homologous to the salivary gland peptide with putative anti-microbial activity. However, there is some difference between the upregulated and downregulated EGPs. The phylogenetic analysis indicated that upregulated EGPs and downregulated EGPs belong to different clades (Figure 3-5). Structurally, the main difference is caused by the sequence areas close to 3' or 5' ends, as well as by a longer fragment consisting of six amino acids (Figure 3-6). It is not clear that if this divergence will cause difference in tertiary structures of the proteins; however, the structural difference has been reflected by differential expression in this study.

In general, experiments performed in this study underscore the importance of the newly developed cDNA microarray to better understand mechanisms of adhesion in the zebra mussel. The microarray analysis directed our attention to a number of proteins of importance to the early byssogenessis, such as the EPG-like peptides, shamatrin-like proteins, and some sequences with unknown functions. Future studies to determine their exact function, location within the byssus, and factors that regulate their expression will be necessary to unravel the mechanism of zebra mussel attachment.

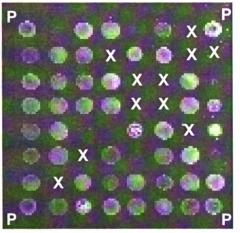


Figure 3-1 Zebra mussel byssus cDNA microarray.
This figure depicts one of the 48 sub-arrays after hybridization. The 64 printed spots were arranged as an 8 × 8 matrix. The spots (P) located at four corners are filled with β-actin DNA. The white crosses demarcate the blank spots that were used as negative controls.

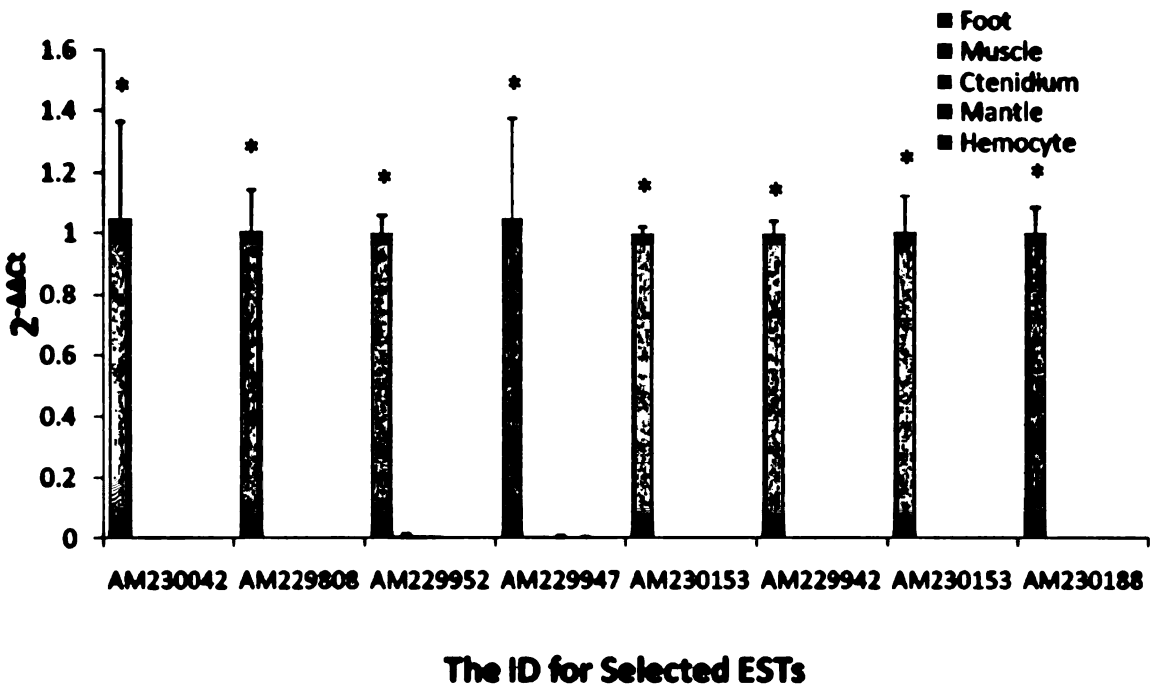


Figure 3-2 qRT-PCR with selected genes using different tissues.
The Ct values in foot tissue were used as control with the value 1. The relative expression levels of the genes in other tissues were calculated using $2^{-\Delta\Delta Ct}$ model. * Significant difference ($P<0.05$). Bars represent average of two biological replicates ±standard errors.

This figure is in color.

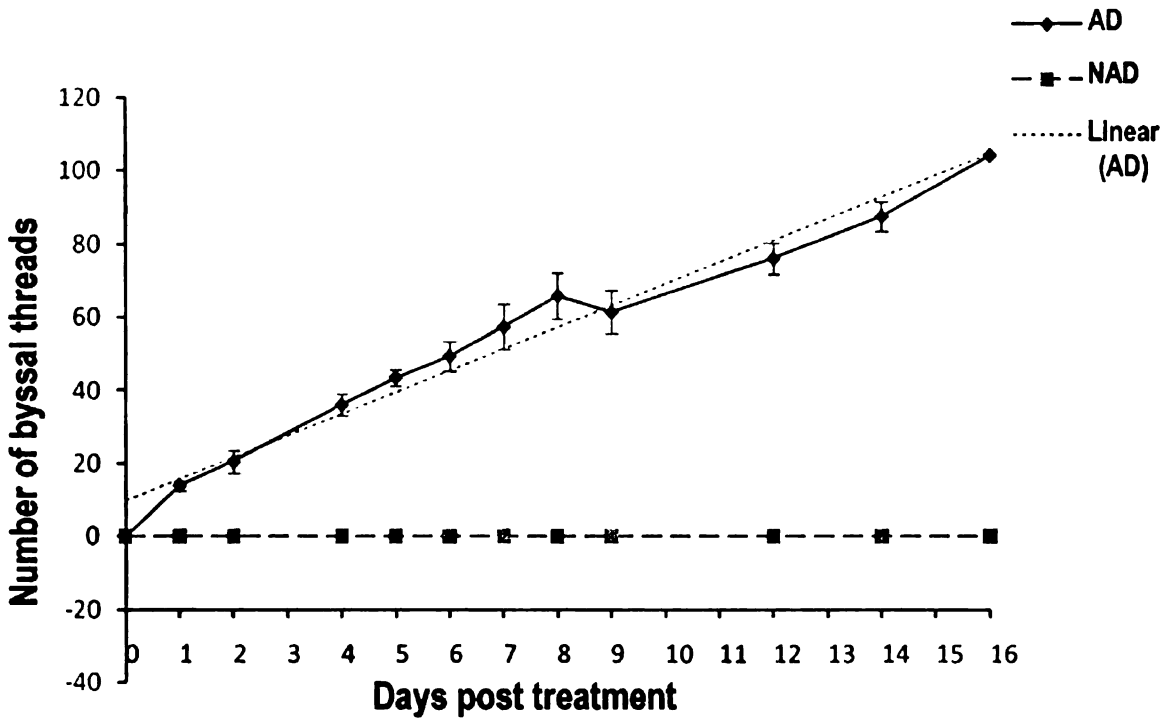


Figure 3-3 Number of regrown byssal threads over 16 days.
Each point represents an average and standard error of thread numbers of six randomly selected mussels. The linear regression analysis of the samples in AD group was performed by Excel with the equation $y = 5.9425x + 9.6942$. The correlation coefficient was calculated as $R^2=0.9705$.

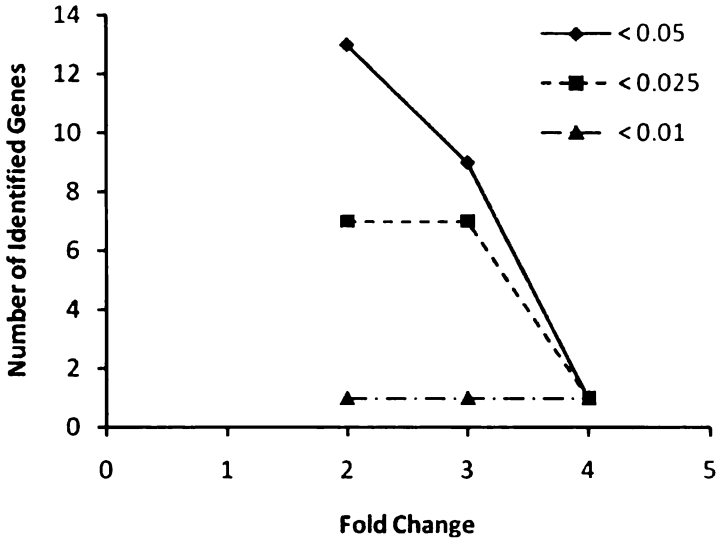


Figure 3-4 Histogram of distribution of the number of differentially expressed genes and fold-change.
The values of x-axis are fold changes. The y-axis stands for the number of the genes. Each curve is drawn under a certain *P* value cutoff: 0.05, 0.025 and 0.01. In each *P* value level, the numbers of differentially expressed genes (y-axis) with different fold change levels (x-axis) were plotted.

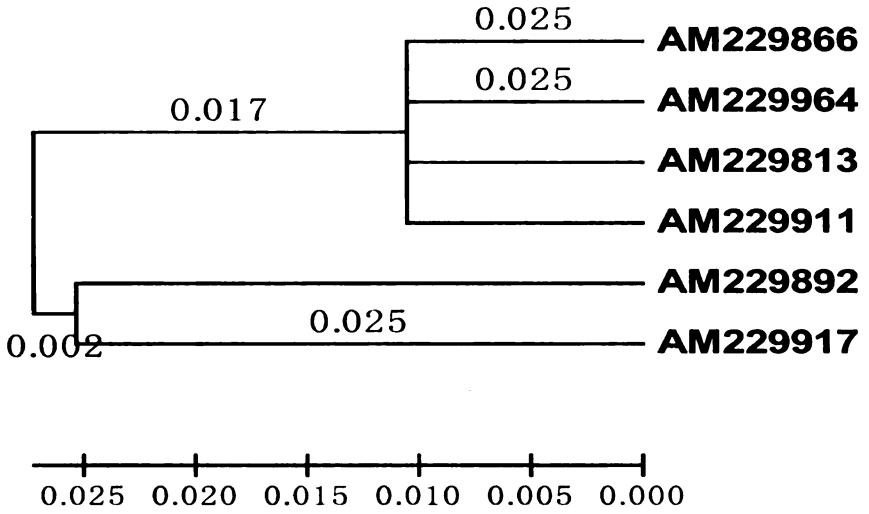


Figure 3-5 Neighbor-Joining phylogenetic analysis with excretory gland peptide-like molecules.
All the proteins selected in this phylogenetic tree are homologous to excretory gland peptide of the western black-legged tick (*Ixodes pacificus*, GenBank accession no. AAT92111) or the black-legged tick (*I. scapularis*, GenBank accession no. AAV80789). The upregulated AM229866, AM229964, AM229813, and AM229911 ESTs clustered together while the downregulated AM229892 and AM229917 ESTs formed another cluster. The number above each branch is the actual length of the branch.

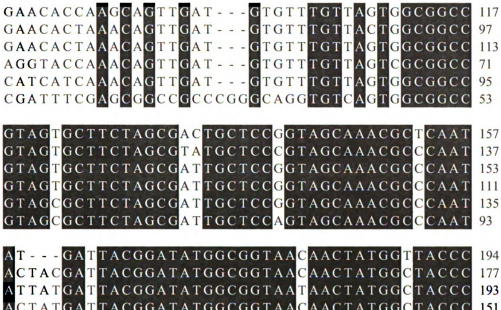


Figure 3-6 Multiple alignments with the seven EGP-like sequences.
EST AM229866, AM229964, AM229813, and AM229911 were all found upregulated in microarray experiment. Sequence AM229892 and AM229917 were proved to be downregulated. ClustalW was used to align all the EGP-like ESTs. The dark stripe indicates that more than 70% nucleotides in this site are identical. The grey shading highlights the site with similar nucleotides (similar biochemical characters). The hyphens show the gaps existing in the sequences.

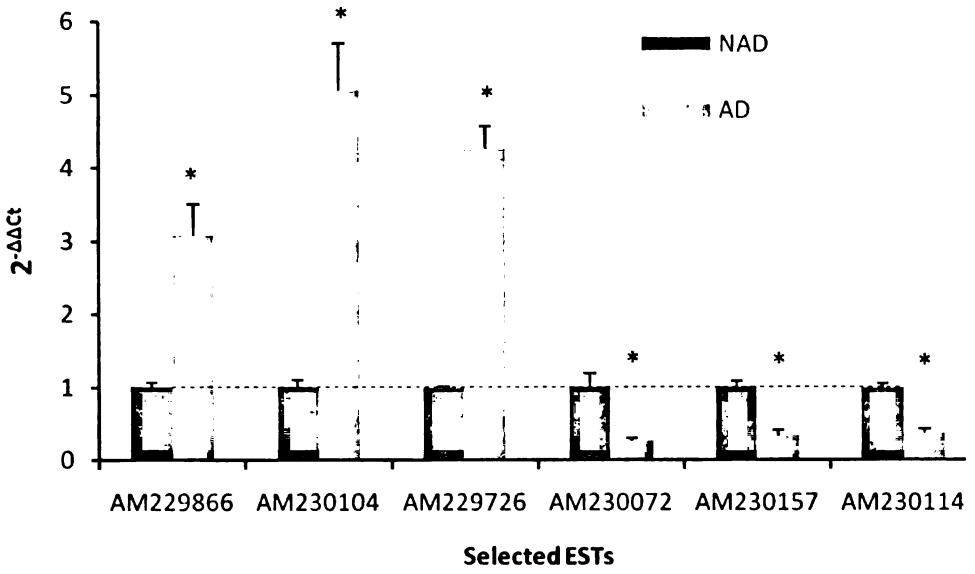


Figure 3-7 Validation of microarray results with qRT-PCR.
The Ct values of NAD were used as control with the value 1. A baseline was drawn to show the expression level of NAD group. The relative expression levels of the genes under attachment in each day were calculated using $2^{-\Delta\Delta Ct}$ model. * Significant difference (P <0.05).

Table 3-1 The primers used for QRT-PCR.

Primer names	Sequence
AM230188_Forward	5'- TGC GAG CCG ACG TTA TCC -3'
AM230188_Reverse	5'- GCT ATC GGC CGC AAC AAT AT -3'
AM229952_Forward	5'- GCG GCC GAG GTT TGT TAG T -3'
AM229952_Reverse	5' -ATC CGT AGT CAT AGT ATT GGG CG -3'
AM229808_Forward	5' -GAT GAA CAC CAA ACA GTT GAT GTG -3'
AM229808_Reverse	5' -TAT TGG GCG TTT GCT ACC G -3'
AM230072_Forward	5'- CTT CCA CGT TAT CCG CAT CA -3'
AM230072_Reverse	5'- TCC CCG TCT GGA GCC TAT C -3'
AM229947_Forward	5'- GGC AGG TAA AAC ACC TGA TGT G -3'
AM229947_Reverse	5'- TCA TAG TAT CGG GCG TTT GCT -3'
AM230157_Forward	5'- GAT CCT ATT GGC CTG GAC CAT -3'
AM230157_Reverse	5'- ATA CCC GCT TAC GGA ATT ACC TT -3'
AM229866_Forward	5'- CCT CCG AGC AGC GAC AAA -3'
AM229866_Reverse	5'- CTA CCG GAG CAG TCG CTA GAA -3'
AM230104_Forward	5'- TCC TTG GGT GTG GCA TCA G -3'
AM230104_Reverse	5'- GAC GAC TGC CCC AAG TTC TC -3'
AM229726_Forward	5'- CCG ATG GGC CAT ATG ATA AGA A -3'
AM229726_Reverse	5'- GAT GCC TTG CTT GTC TTG TTG A -3'
AM230114_Forward	5'- GAA CTG TGC GCT GGA TAC GA -3'
AM230114_Reverse	5'- GGA TCA CCC AGG CTC CAT T -3'
AM229942_Forward	5'- CCA GGA AAT TAT GGC GAC TAT GA -3'
AM229942_Reverse	5'- CCA CCC CCT CCG AGG AT -3'
AM230042_Forward	5'- GAT GAA CAC CAA CCA GCT GAT G -3'
AM230042_Reverse	5'- TAC TGG GCG TTT GCT ACC G -3'
AM230153_Forward	5'- TCC TAT TGG CCT GGA CCA TG -3'
AM230153_Reverse	5'- CCC GCT TAC GGA ATT ACC TTG -3'
*18SrRNA_Forward	5'- GAC ACG GCT ACC ACA TCC AA -3'
*18SrRNA_Reverse	5'- CTC GAA AGA GTC CCG CAT TG -3'

* The primers used as internal reference for QRT-PCR.

Table 3-2 The ESTs with upregulated expression level in AD group.

Accession #	Genes ID	LogFC	Fold-change	*P*	Putative Function
AM229866*	BG21_B08	2.402	5.285354	4.00E-05*	salivary gland peptide [*Ixodes scapularis*]
AM230108*	BG34_H05	2.679	6.404118	0.00018	N/A
AM230253*	BG28_F07	1.811	3.508854	0.00196	N/A
AM229964*	BG13_A07	1.535	2.897884	0.00244	salivary gland peptide [*Ixodes scapularis*]
AM229911*	BG33_C07	1.402	2.642677	0.00563	salivary gland peptide [*Ixodes pacificus*]
AM229799*	BG30_F03	1.567	2.96288	0.00737	Shematrin-4 [*Pinctada fucata*]
AM229776*	BG22_A01	1.348	2.54559	0.00778	N/A
AM229813	MF030105_E04	2.585	6.000156	0.01072	salivary gland peptide [*Ixodes pacificus*]
AM230109	BG03_A09	1.281	2.430074	0.01141	N/A
AM230094	BG33_H08	1.273	2.416636	0.01193	choriogenin Hminor [*Oryzias latipes*]
AM230117	BG33_H09	1.272	2.414961	0.01206	N/A
AM230104	BG10_D04	1.271	2.413288	0.01206	hemicentin 1 [*Homo sapiens*]
AM230205	BG20_C08	1.211	2.31498	0.01679	N/A
AM230183	BG14_F10	1.181	2.267339	0.01975	N/A
AM230042	BG25_H08	1.14	2.20381	0.02435	N/A
AM229726	BG20_A02	1.311	2.481135	0.02499	Byssal protein Dpfp1 [*Dreissena polymorpha*]
AM230399	BG18_F10	1.585	3.000078	0.02692	N/A
AM229876	BG14_C02	1.113	2.16295	0.02803	N/A
AM230435	BG28_B04	1.105	2.150989	0.02913	N/A
AM230179	BG08_B03	1.261	2.396618	0.03104	N/A
AM230100	BG34_A07	1.063	2.089272	0.03586	N/A
AM230249	BG28_H09	1.046	2.064797	0.03897	N/A
AM229789	BG06_C03	1.463	2.75681	0.04111	N/A
AM229935	BG97-192_E06	1.028	2.039195	0.04241	Shematrin-4 [*Pinctada fucata*]

* The differentially expressed ESTs with P <0.01.

Table 3-3 The ESTs with downregulated expression level in AD group.

Accession #	Genes ID	Coef	Fold Change	*P*	Putative Function
AM230401*	BG18_D08	-2.447	0.183392	1.35E-06	N/A
AM230072*	BG18_H07	-3.022	0.123108	2.00E-05	N/A
AM229892*	BG31_E06	-2.01	0.248273	0.00059	Salivary gland peptide [*Ixodes scapularis*]
AM229917*	BG33_F04	-1.98	0.25349	0.00071	Salivary gland peptide [*Ixodes scapularis*]
AM229853*	BG97-192_H11	-1.348	0.392836	0.00778	N/A
AM229749*	MF030105_C07	-1.543	0.343171	0.00833	N/A
AM230237*	BG34_G06	-1.326	0.398873	0.00884	N/A
AM230157*	BG34_A05	-1.318	0.401091	0.00929	Polypeptide release factor [*Yarrowia lipolytica*]
AM230394*	BG17_G10	-1.313	0.402483	0.00955	N/A
AM229918	BG18_D06	-1.292	0.408384	0.01072	N/A
AM230372	BG15_H07	-1.285	0.410371	0.01117	N/A
AM230377	BG16_D10	-1.792	0.288771	0.01234	N/A
AM229777	BG31_E04	-1.259	0.417833	0.01293	N/A
AM229781	BG29_F09	-1.245	0.421908	0.014	N/A
AM230109	BG23_D06	-1.235	0.424842	0.01479	N/A
AM230268	MF030105_B12	-1.422	0.373195	0.01502	N/A
AM229973	BG28_E10	-1.231	0.426022	0.01505	Salivary gland peptide [*Ixodes scapularis*]
AM230114	BG20_C01	-1.733	0.300826	0.01555	Kallikrein [*Sus scrofa*]
AM229947	BG26_E11	-1.209	0.432568	0.01694	Salivary gland peptide [*Ixodes pacificus*]
AM230178	MF030105_C01	-1.117	0.461052	0.02738	N/A
AM230258	BG16_G06	-1.287	0.409802	0.0278	N/A
AM229831	BG32_F09	-1.286	0.410086	0.0279	Salivary gland peptide [*Ixodes pacificus*]
AM230384	BG17_C09	-1.285	0.410371	0.02799	N/A
AM230406	BG18_G03	-1.249	0.42074	0.03277	N/A
AM229920	BG33_E08	-1.45	0.366021	0.04291	Salivary gland peptide [*Ixodes scapularis*]
AM229864	BG16_H05	-1.181	0.441046	0.04338	Salivary gland peptide [*Ixodes scapularis*]
AM229892	BG34_B08	-1.013	0.495515	0.0455	Salivary gland peptide [*Ixodes scapularis*]
AM230248	BG17_D01	-1.419	0.373971	0.04754	N/A

* The differentially expressed ESTs with P <0.01.

CHAPTER IV

Gene Expression Profiling During the Byssogenesis of Zebra Mussel (*Dreissena polymorpha*)

Abstract

Since its invasion to the North American waters 20 years ago, the zebra mussel (*Dreissena polymorpha*) has negatively impacted the ecosystems it has infested through its firm underwater adhesion. The molecular mechanisms governing the functions of the zebra mussel byssus, the main structure responsible for maintaining the underwater adhesion, have received little attention. Our previously developed zebra mussel foot byssus cDNA microarray was applied in this study to identify the genes involved in different stages of the byssal threads generation. Byssal threads of zebra mussels were manually severed under laboratory conditions and the formation of new byssal threads was followed over a three week course. By comparing the gene expression profiles in different stages of byssal threads generation (byssogenesis) to their baseline values, we found that the number of unique byssus genes differentially expressed at 12-hour, 1-day, 2 days, 3 days, 4 days, 7 days, and 21 days post-treatment was 13, 13, 20, 17, 16, 20, and 29, respectively. Comparisons were also made between two subsequent samples (e.g., 12h vs. 1d, 1d vs. 2d, 2d vs. 3d, and so on).

Seven differentially expressed genes were selected for validation by using quantitative reverse transcription PCR (qRT-PCR) and the results were consistent with those from the

microarray analysis. By using fluorescent in situ hybridization (FISH), we found that two microarray identified genes, BG15_F03-DPFP and BG16_H05-EGP, were expressed in two major byssus glands located in the zebra mussel foot: the stem-forming gland and plaque-forming gland, respectively. Moreover, the qRT-PCR of seven microarray identified genes with different zebra mussel samples suggested that they were also expressed in other mussel tissues beside the foot, albeit at much lower levels. This suggests that the microarray identified genes were produced primarily by the foot, and are likely associated with byssogenesis. The differentially expressed genes identified in this study suggested that multiple molecules are involved in byssogenesis, most likely performing multiple functions during the generation of byssal threads. These results obtained herein represent the first logical step toward understanding underwater attachment mechanisms employed by this invasive species.

Introduction

Native to eastern Eurasia, the zebra mussel (*Dreissena polymorpha*) has invaded North America since the late 1980s. Zebra mussel invasion led to economical and ecological devastation. The mussel attaches to substrates underwater with hair-like byssal threads produced by glands within the zebra mussel foot (Morton 1969; Claudi & Mackie 1993; Benson & Raikow 2009). The strength of zebra mussel attachment is mostly decided by the number of its byssal threads (Dormon *et al.* 1997), thus the generation of byssal threads, also known as byssogenesis, is critical for zebra mussel underwater attachment and survival. The mechanism of zebra mussel byssogenesis remains largely unknown, especially at the molecular level. Based on previous morphological and biochemical studies, it has been determined that byssogenesis in the zebra mussel foot goes through multiple stages, ending up with the formation of byssal threads with adhesive pads (Eckroat *et al.* 1993).

Three major byssus glands are involved in this process. The stem-forming gland, located at the root area of the zebra mussel foot, was the starting point of the zebra mussel byssal threads. The thread-forming gland cells evenly distributed along the ventral groove in the middle of the foot produce two main zebra mussel foot proteins, *D. polymorpha* foot protein -1 and -2 (Dpfp-1 and -2), which serve as the component maintaining the byssal structures (Rzepecki & Waite 1993b, a). Another major byssus gland, the plaque-forming gland, is located in the tip area of the foot and involved in the formation of byssal thread plaques (Rzepecki & Waite 1993b). The byssogenesis in the zebra mussel foot ceases when the mussels are fully attached to suitable substrates (Eckroat *et al.* 1993). On the contrary, the byssogenesis can be initiated by the removal of

existing byssal threads (Clarke & McMahon 1996c, b; Clarke & McMahon 1996a; Clarke 1999).

Morphological studies of the zebra mussel foot revealed the presence of three major exocrine glands that are involved with byssal thread formation: the stem-forming gland, thread-forming gland, and plaque-forming gland (Rzepecki & Waite 1993a). Located at the root of the zebra mussel foot, the stem-forming gland forms and secretes the proteins that build the trunk of the byssal threads. This process is considered the initiation event of byssogenesis. The thread-forming gland cells are distributed along the foot's ventral groove and produce a large amount of Dpfp-1 and -2, both important in holding the byssal thread structure. The plaque-forming gland, involved in the formation of the enlarged end of byssal threads, is located at the tip section of the zebra mussel foot. Along with the thread-forming gland, the plaque-forming gland releases numerous foot proteins into the ventral groove. The mixed proteins in the ventral groove are transported to the root of the foot, and thereafter, expelled into the hydrophobic environment created between the foot and underwater substrates to form the complete byssal thread (Eckroat *et al.* 1993). Clearly, the process of byssogenesis seems to be regulated by a number of intrinsic and extrinsic factors.

Most recently, we have developed a cDNA microarray (716 genes) based on normalized zebra mussel foot cDNA library (Xu & Faisal 2008; Xu & Faisal 2009a). In a previous study, we identified molecules with potential involvement in zebra mussel attachment which are differentially displayed when the mussels have an access to attachment substrates for 48 hours. Since the byssal threads take up to 21 days to be fully formed (Waite 2002; Farsad *et al.* 2009), this study was designed with a holistic approach

the would allow the identification of shifts in gene expression during the entire length of byssogenesis. In specific, we employed cDNA microarray to compare gene expression patterns between quiescent, fully attached mussels, and those in mussels at different time points during the formation of new byssal threads.

Materials and Methods

Zebra mussel sample collection and experimental design

Zebra mussels used in this study were collected from the Vineyard Lake in Brooklyn, MI, USA (Latitude: 42°4'59''N; Longitude: 84°12'34"W). In the laboratory, the mussels were thoroughly cleaned, placed in submersed glass PYREX® 100 × 15mm Petri-dish (Corning Inc, Corning, NY) in aerated, filter-sterilized (0.22 μm filter) Vineyard Lake water in a glass aquarium (30 gal). The mussels were fed weekly with a pure culture of the algus *Ankistrodesmus falcatus* and were allowed to be acclimated to the laboratory conditions for eight weeks. The adapted mussels were then divided into two groups. Mussels of the first group were detached, their byssal threads severed as close as possible to their roots by a sharp scalpel, and then allowed to reattach in the submersed Petri-dishes. In the second group, the mussels were allowed to maintain their attachment status and were considered the control group.

Total RNA was extracted from zebra mussel feet of both groups at 12 hours, 1 day, 2 days, 3 days, 4 days, 7 days, and 21 days post mussel reattachment. For Group 1, four biological replicates were obtained at each sampling time point and each replicate contained total RNA pooled from five mussel feet. RNA Samples were collected from mussels in Group 2 in the same manner at the same time intervals. Twenty-six microarray slides were used for hybridizations. Each slide was hybridized with two different samples labeled with Alexa 555 and Alexa 647, respectively. Thus, the gene expression profiles of the two samples on each slide could be compared. Fourteen slides were used to compare mussels in Group 1 and Group 2. The other 12 slides were applied to differentiate the

gene expression profiles between any two Group 1 samples at connected time points (i. e., 12 hours vs. 1 day, 1 day vs. 2 days, and so on). The detailed sample labeling and hybridization combination can be found in the description of the deposited microarray data in NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) with the series accession number GSE16407.

RNA extraction and cDNA synthesis

The total RNAs used for microarray hybridizations were extracted with 5-PRIME PerfectPure RNA Tissue Kit (5 PRIME Inc, Gaithersburg, MD) and the protocol followed the manufacturer's instructions for the kit. The reverse transcription of the total RNAs was performed with SuperScript *Plus* Indirect cDNA Labeling Kit (Invitrogen, Carlsbad, CA) containing Alexa Fluor 555 & 647 fluorescent dyes. Fifteen micrograms of total RNA from each sample was used for cDNA synthesis. The procedures of cDNA syntheses and labeling were all described by the instructions for the kits.

Microarray hybridization and data analyses

Details of the zebra mussel foot cDNA microarray used in this study can be found in our previous report (Xu & Faisal 2009a). Microarray hybridizations were performed on the GeneTAC HybStation (Genomic Solutions, Ann Arbor, MI), and the 18-hour step-down protocol for hybridization was used as described previously (Xu & Faisal 2009a). The hybridized dual-channel array slides were scanned by GenePix 4000B two-laser Scanner (Molecular Devices, Downingtown, PA), and the images from the scanner were processed by GenePix Pro 6.0 (Molecular Devices) software to generate the spot intensity file of each array image. The spot density files were analyzed with Limma software

package (Smyth 2005). The preliminary data processing includes background correction, within array normalization and between array normalization. The differentially expressed genes were determined by least square regression. The relative expression level of each gene at each time point was calculated compared to the common reference. Meanwhile, the differential expression of each gene between any of the two time points was also calculated and demonstrated by the logarithmic value of the ratio between the two time points. The hierarchical clustering was done by the software Genesis (Sturn *et al.* 2002) and the dataset of this microarray study was also deposited in NCBI GEO (series accession number GSE16407). The putative functions of the identified genes were predicted by the homologues searching method, Basic Local Alignment Search Tool (BLAST) (Altschul *et al.* 1990) with GenBank nucleotides and protein database. For the genes without homologues in the database, the sequences were scanned with InterProScan function (Zdobnov & Apweiler 2001) in software package Blast2GO (Conesa *et al.* 2005) to detect the signature sequences of unknown molecules.

Quantitative reverse transcription PCR (qRT-PCR) validation of microarray data

The qRT-PCR was applied to validate the expression profiles of the genes that were affected by attachment status identified by microarray. Additional zebra mussels were collected and maintained as described above. The treatment and RNA sampling of mussels were the same as for the samples used for microarray study. The one-step qRT-PCR was performed with Power SYBR® Green RNA-to-CT^{TM} 1-Step Kit (Applied Biosystems Inc, Foster City, CA) and Mastercycler ep realplex S thermocycler system (Eppendorf, Westbury, NY). The thermocycler program used for the qRT-PCR was 30 minutes at 48 °C, 10 minutes at 95 °C and 40 repeats of a 2-step temperature cycle (15

seconds at 95 °C and 1 minute at 60 °C). The relative expression levels of the gene expression were calculated with $2^{-\Delta\Delta Ct}$ algorithm using cycle threshold (Ct) of PCR described by Livak and Schmittgen (Livak & Schmittgen 2001). The gene specific primers and the internal reference control used in this assay were designed by Primer ExpressTM (Applied Biosystem Inc.), listed in Table 4-1.

Visualization of two differentially expressed genes in foot tissues using RNA fluorescence in situ hybridization (FISH)

Fluorescence *in situ* hybridization (FISH) was applied in order to locate the transcripts of two genes *in situ*, of potential importance to byssogenesis. The product of BG15_F03-DPFP is homologous to *D. polymorpha* foot protein -1 (Dpfp-1), which is believed to play a major role in the maintenance of the byssal thread structure, while the protein encoded by BG16_H05-EGP is homologous to one of the peptides secreted by *I. pacificus* salivary gland. The zebra mussel foot tissue was immediately fixed in 2% paraformaldehyde phosphate buffer saline after severing from the mussel. The fixation solution contained 0.02 M NaH_2PO_4, 0.0077 M Na_2HPO_4, 1.4 M NaCl, and 2% w/v paraformaldehyde (pH 8.0). Excised tissue was incubated in the fixation buffer for five hours at room temperature and then kept in 100% ethanol until use. The dehydrated tissue sample was then covered totally with paraffin and longitudinally sectioned (5 μm). Ten serial longitudinal sections from the middle of the foot were mounted on RNase- and DNase-free glass slides (Corning Inc.).

The probes used for the FISH were RNA probes which were complementary strands of the transcripts of the two selected genes. The PCR product of the gene

fragment was cloned into pGEM®-T Easy Vector Systems (Promega U.S., Madison, WI). Then the positive colony was picked for sequencing with M13 forward primer to clarify the direction of the insert of the recombinant plasmid. The PCR was applied on the selected colonies with both M13 forward and reverse primers. The PCR products were purified with Wizard SV Gel and PCR Clean-Up System (Promega) and used as the templates of RNA transcription. The T7 or Sp6 RNA polymerase was selected according to the insert direction of the recombinant plasmid to synthesize the antisense RNA. Besides the genes BG15_F03-DPFP and BG16_H05-EGP, RNA probe of zebra mussel cytoplasmic actin (AF082863) gene was also produced in the same way described above. The probe BG15_F03-DPFP (DPFP-G) was labeled with fluorescent dye Alexa 488 (Invitrogen) while the probe BG16_H05-EGP (EGP-R) was labeled with Alexa 594 (Invitrogen). The control probe cytoplasmic actin was divided into two groups and labeled with Alexa 594 (ACT-R) and Alexa 488 (ACT-G), respectively. The steps of the RNA synthesis and labeling can be found in the instructions of the two kits: FISH Tag™ RNA Green Kit (Invitrogen) and Fish Tag™ RNA Red Kit (Invitrogen). The probes DPFP-G and ACT-R were used to hybridize with same zebra mussel foot section slide while the EGP-R and ACT-G were applied on the other slide. The RNA-FISH was performed as recommended by manufacturer (Invitrogen). The hybridized slides were visualized by Olympus BX41 (Olympus America Inc., Center Valley, PA) microscope under the excitation of mercury lamp and the images were captured by Olympus DP25 digital camera (Olympus America Inc.). The image processing software DP2-BSW (Olympus America Inc.) was used to combine the different color channels into a single image.

Tissue distribution of representative expressed genes

In order to find whether the the differentially expressed genes (DEG) are also expressed in other tissues besides the zebra mussel foot, the one-step qRT-PCR was performed to detect the amount of transcripts of selected DEG in zebra mussel hemocyte, retractor mussel, ctenidium, mantle, as well as foot. The mussels were collected, maintained, and handled similarly to the mussels described above. Immediately following the acclimation period, RNA samples were extracted from hemocytes, feet, muscles, ctenidia, and mantles of zebra mussels. The protocol for hemocyte collection was described by our previous studies (Xu & Faisal 2007, 2009b) and the other samples were taken by sterilized scissors and scalpels. The primers used in this study were listed in Table 4-1.

Results

Differentially expressed genes during the byssogenesis identified by microarray

Compared to the non-byssogenesis group, several genes have been up- and downregulated at the different time points during byssogenesis; their numbers and nature varied, however, along the course of byssogenesis. Numberwise, 9, 8, 3, 2, 10, 12, and 15 genes were upregulated at 12 hours, 1 day, 2 days, 3 days, 4 days, 7 days, and 21 days post-treatment, respectively. On the contrary, 4, 5, 17, 15, 6, 8, and 14 genes were downregulated after 12 hours (Table 4-2), 1 day (Table 4-3), 2 days (Table 4-4), 3 days (Table 4-5), 4 days (Table 4-6), 7 days (Table 4-7), and 21 days (Table 4-8), respectively. Comparative gene expression studies were performed between any two different time points for each differentially expressed gene.

Based on the expression patterns of each differentially expressed gene along the time course of the byssogenesis, the genes were hierarchically arranged into 12 clusters (A-L, Figure 4-1). As demonstrated in Figure 4-1, cluster A included 16 genes which were mostly upregulated during the early stages of byssogenesis (12 hours to 1 day of byssogenesis) and the upregulation was diminished after that. The eight genes included in cluster B had opposite expression patterns than cluster A, which were not upregulated until 21 days of byssogenesis. The expression patterns of the eight genes in cluster C were similar to those in cluster B; however, they were suddenly downregulated at day 7 during byssogenesis before they were significantly upregulated in byssogenesis at day 21. Along the time course of the byssogenesis, the expression levels of the genes in cluster D,

E, F, and H were downregulated at the time points 4-, 2-, 3-, and 21-days, respectively. The genes in cluster G appeared to have downregulated expression levels during the byssogenesis of the whole experiment. The downregulation became more and more significant from the beginning of the byssogenesis until day 7, where the downregulation of the genes in cluster G reached maximum level. The three genes in cluster G included an exocrine gland peptide-like gene (BG16_H05), a gene with anaphylatoxin domain and an epidermal growth factor region (BG16_G06), and a gene homologous to neuropeptide-like protein 31 (MF030105_E03). On the contrary, the cluster I contained three genes whose expression patterns all appeared to be upregulated during the byssogenesis with the maximum expression levels at 7 days post-treatment. Two of the genes in cluster I, BG18_G06 and BG16_D10, had no homologues found in the database. The other one in this cluster, BG29_D08, was a homologue to another exocrine gland peptide, which was structurally different than that in cluster G. Eleven genes were included in cluster J and were upregulated at days 4 and 7 during byssogenesis. Expression patterns of the six genes in cluster K were first upregulated at day 7, and subsequently downregulated at day 21 byssogenesis. The last cluster (cluster L) had all the genes with the expression profiles downregulated at the early stage of byssogenesis (12-hour) but the expression levels were continuously increased until day 7 (Figure 4-1).

Validation of the results obtained from microarray

The expression patterns of seven genes identified by microarray were validated by qRT-PCR. According to the qRT-PCR results, the expressions of BG20G04, BG97/192_E09, and BG15_H03 were all upregulated at time points 12 hours, 1 day and 4 days post the initiation of byssogenesis, respectively (Figure 4-2B), which is consistent

with the microarray results (Figure 4-2A). The other four genes were all differentially expressed at two time points during the byssogenesis based on the microarray results. BG16_H05 was downregulated after 4 days and 7 days reattachment (Figure 4-2A). Similarly, the downregulation of the gene expression of BG97/192_E09 during the byssogenesis appeared at 12 hours and 1 day time point, along the time course (Figure 4-2A). On the contrary, BG30_H08 demonstrated upregulated expression profiles 2 days and 3 days post reattachment (Figure 4-2A). Moreover, the downregulation of BG03_H04 expression started after 7 days reattachment while it was significantly upregulated at day 21 post reattachment. The expression profiles of all four of these genes were confirmed by qRT-PCR (Figure 4-2B).

The *in situ* expression of selected genes within the zebra mussel foot

Within the foot tissues, the expression locations of two genes, BG15_F03-DPFP (Figure 4-3A) and BG16_H05-EGP (Figure 4-3B), were observed by using fluorescence *in situ* hybridization (FISH). The yellow signal in Figure 4-3A represented the stem-forming gland area where both beta-actin (red) and BG15_F03-DPFP (green) were expressed. The RNA of BG15_F03-DPFP in the thread-forming gland and plaque-forming gland was not detected. The expression of the gene BG16_H05-EGP (labeled with red dye) can only be observed at the tip of the zebra mussel foot where the plaque-forming gland is located (Figure 4-3B). The strong orange in Figure 4-3B suggests that the expression level of BG16_H05-EGP gene in the plaque-forming gland was very high compared to the background stained by β-actin (green). The other byssus gland areas, the stem-forming gland and thread-forming gland, did not show significant RNA signal of BG16_H05-EGP.

The relative expression levels of selected genes in zebra mussel tissues

The distribution of the gene expression within the zebra mussel was detected also by using RT-PCR. The results demonstrated that the expression levels of all the seven genes selected in this assay were extremely high in the foot tissue. Although the transcripts of Bg91/192_E09, BG04_H04, BG15_F03, BG16_H05, and BG20_G04 were detected in all the other four tissue and hemocyte samples, the expression levels of these genes in the zebra mussel foot are significantly higher than in any of the other tissues (Figure 4-4). The expression area of the gene MF030105_H05 and BG30_H08 was smaller. MF030105_H05 was also slightly expressed in the hemocyte and muscle samples other than the foot, while the BG30_H08 can be only detected within the foot and hemocyte. Similarly, the expression levels of these two genes outside the zebra mussel foot were extremely low compared to that within the foot (Figure 4-4).

Discussion

Byssogenesis is a vital physiological activity of healthy zebra mussels to ensure firm underwater attachment, primarily by a number of strong byssal threads with adhesive pads. Detached mussels try to resume attachment through the production of fresh byssal threads (Rajagopal *et al.* 2002). Two distinct types of byssal threads are produced during byssogenesis: temporary and permanent threads. Within a short time after detachment, the temporary threads are produced, followed by the generation of permanent threads without interrupting the production of additional temporary threads. There is also significant morphological difference between the fresh (< 1 week) and the aged (3 week) permanent byssal threads, due to the crosslink reaction of the foot proteins with 3,4-dihydroxyphenylalanine (DOPA) and the melanization of the exogenous L-DOPA with catechol oxidase (Farsad *et al.* 2009). It suggests that at least three stages exist in byssogenesis: temporary byssal thread generation, permanent byssal thread generation, and permanent byssal thread modification. By comparing the genes differentially expressed at the different stages of zebra mussel byssogenesis, we found some genes that were differentially expressed at one time period only. According to Clark and MacMahon (Clarke & McMahon 1996c, b; Clarke & McMahon 1996a; Clarke 1999), after 12 hours reattachment, most of the byssal threads are of the temporary type. Therefore, the genes with the significantly differentiated expression patterns on or around day 1 post-treatment are associated with the generation of temporary byssal threads. This may explain the drop in the number of upregulated genes following the Day 1 sampling. Similarly, since the three-week permanent byssal threads are morphologically different than the fresh threads of less than one week, the genes identified by microarray after 21

days reattachment may play certain roles in the biochemical and morphological modification noticed in the permanent byssal threads. The other genes that were differentially expressed between day 1 and day 7 are possibly involved in the generation and maintenance of the permanent byssal threads. Unfortunately, a large number of identified genes do not have predictable functions due to the limited availability of homologues in gene bank database. The putative functions and their roles in the byssogenesis require additional analysis.

Among the identified genes with predicted functions, a number of sequences were homologous to the *Dreissena polymorpha* foot protein 1 (Dpfp-1) whose expression seems to have fluctuated over our observation period. This particular protein has been identified as a major foot protein with at least 10-15 variants which constituted a polymorphic protein family specific to the zebra mussel (Rzepecki & Waite 1993b). Therefore, we were not surprised with the fluctuation of Dpfp-1 homologues between up- and downregulation in our study. FISH has demonstrated that the transcription products of a Dpfp-1 homologue had been detected in the byssal thread stem-forming gland, which probably contributes to the protein components forming the stem of the byssal thread. Functionally, the Dpfp-1 serves as a protection or structure supporting material instead of adhesive matrix (Anderson & Waite 1998). Based on the observation of Farsad *et al* (Farsad *et al.* 2009), it is likely that other DOPA-rich foot proteins with adhesive characteristics are deposited between the plaques and the hard surface to thereby accomplish the underwater adhesion.

A large number of EGP-like encoding genes identified at all the time points in this study may also play a role in adhesion. Structurally, the amino acid compositions of these

EGP-like proteins are very similar to zebra mussel Dpfp whose amino acid sequences are dominanted by proline, tyrosine, aspartic acid/asparagine, lysine, threonine, and glycine residues (Anderson & Waite 1998). It structurally allows these molecules to form crosslinks with DOPA which largely exists in the zebra mussel byssal threads and seems to be crucial in their attachment to underwater substrates (Waite 2002). As demonstrated in this study, several EGP-like genes were differentially expressed at all time points in this study, suggesting their heavy involvement in byssal thread production, with up- or downregulation representing a cascade of feedback mechanisms. Our FISH results indicated that one of the EGP homologues was indeed found abundantly expressed in the plaque-forming gland of zebra mussel; hence this protein is involved in zebra mussel thread formation.

Based on the results of this study, it seems that byssal thread formation is also associated with the differential expression of a battery of genes whose transcribed proteins seem to serve multiple purposes. For example, our results suggest that genes encoding neuropeptide-like proteins of *Caenorhabditis elegans*, known to be involved in neuronal signaling, development and antimicrobial activities (Nathoo *et al.* 2001; Couillault *et al.* 2004; McVeigh *et al.* 2008). Whether these molecules have a similar function in the zebra mussels and are, therefore, regulated due to their importance in development, sensation, antimicrobial activities, is currently unknown and deserves additional studies. Likewise, the differentially displayed zebra mussel genes whose transcripts found no homologues in databases are likely important in byssogenesis and related processes. Based only on significant differential expression of genes on day 2 post-treatment, we have previously demonstrated that the zebra mussel byssogenesis

environment is likely affected by surrounding environmental factors such as temperature, dissolved oxygen, and water flow (Rajagopal *et al.* 2005).

The cDNA microarray technique has been proven to be a powerful tool to differentiate the gene expression patterns along the time-course of a biological process (Evans & Somero 2008; Pritchard & Nelson 2008; Suh *et al.* 2008), and in this case, the zebra mussel byssus cDNA microarray was successfully applied to identify the differentially expressed genes during the different stages of zebra mussel byssogenesis. Given that all the genes identified in this study are indeed involved in certain stages of the byssogenesis, the mechanism of zebra mussel underwater adhesion is still an intriguing puzzle. What do these proteins encoded by the up or downregulated genes exactly do during the byssal thread generation? How are they regulated by the other components involved in this process? Are the foot proteins constitutively produced by byssus glands or can they be induced by external stimuli? These questions will not be answered until detailed analyses are performed on the functional genes identified by this study. Further analyses of these genes will be a great progress in understanding the zebra mussel attachment mechanism.

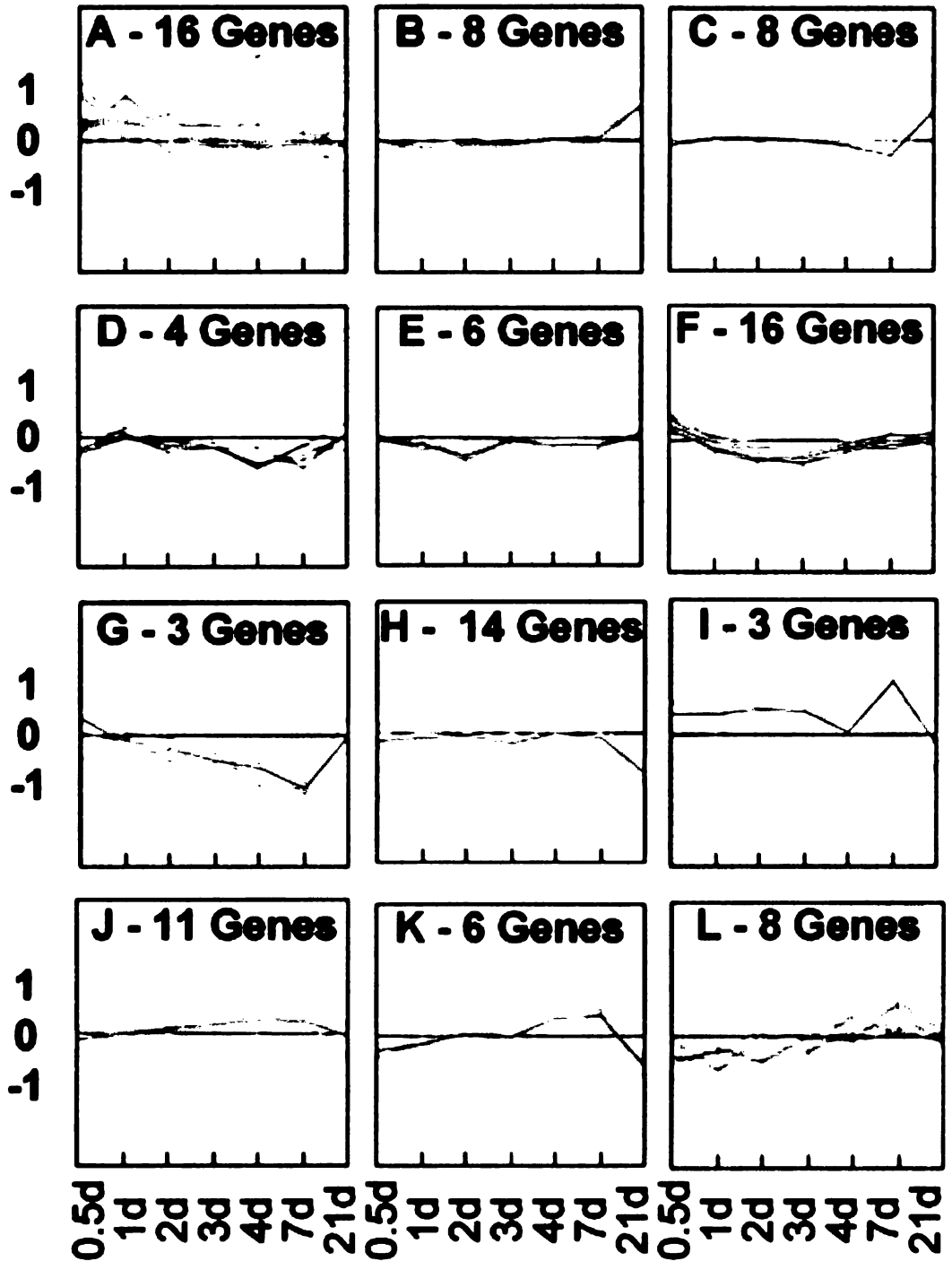
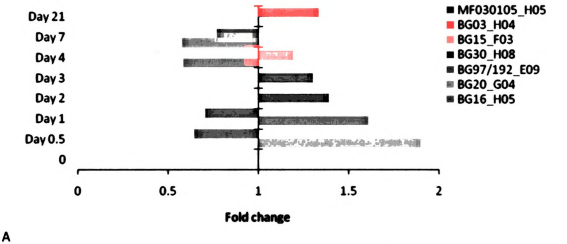
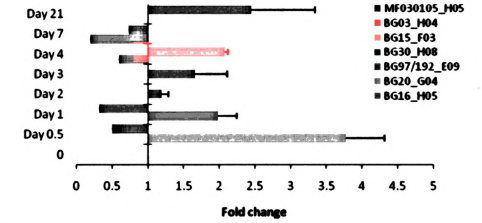


Figure 4-1 The hierarchical cluster with the microarray identified genes.
The hierarchical cluster of the genes differentially expressed during the byssogenesis based on their expression patterns at the different time points along the time-course. Twelve clusters are generated by hierarchical cluster analysis. The pink curve in each cluster figure represents the expression profiles of the genes in this cluster during the byssogenesis.

This figure is in color.



A



B

Figure 4-2 The comparisons of gene expression profiles between the microarray results and qRT-PCR.

A. The differentially expressed genes identified by cDNA microarray and their expression profiles during byssogenesis. B. The qRT-PCR results validating the relative expression levels of the microarray identified genes during the byssogenesis. The byssogenesis samples and non-byssogenesis reference samples were treated as described in Materials and Methods, and the non-byssogenesis sample was used as control with value 1. The amount of the transcripts of the gene was calculated by using $2^{-\Delta\Delta Ct}$ model.

This figure is in color.

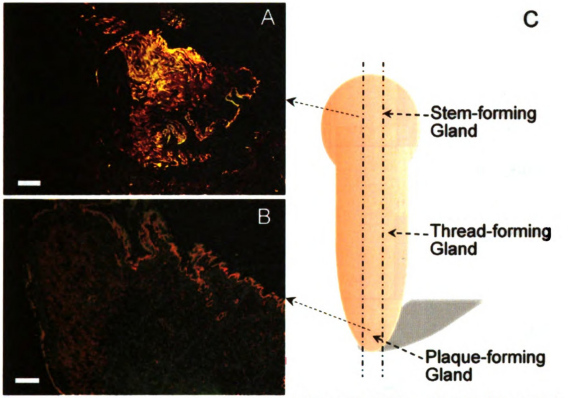


Figure 4-3 The *in situ* expression of the gene BG15_F03-DPFP and BG16_H05-EGP in zebra mussel foot tissue.

A. The synthesized antisense RNA of BG15_F03-DPFP was labeled with Alexa 488 dye (green) while the complementary RNA for beta-actin was labeled with Alexa 594 (red). B. The RNA probe for BG16_H05-EGP was labeled with Alexa 594 (red) and the reference probe, beta-actin in this case was labeled with Alexa 488 (green). The foot was cut longitudinally as shown by the two parallel dashed lines across the foot. The positions of the byssus glands were marked on the foot. The dashed lines connected between slides and foot indicated the position of the three slides on the foot.

This figure is in color.

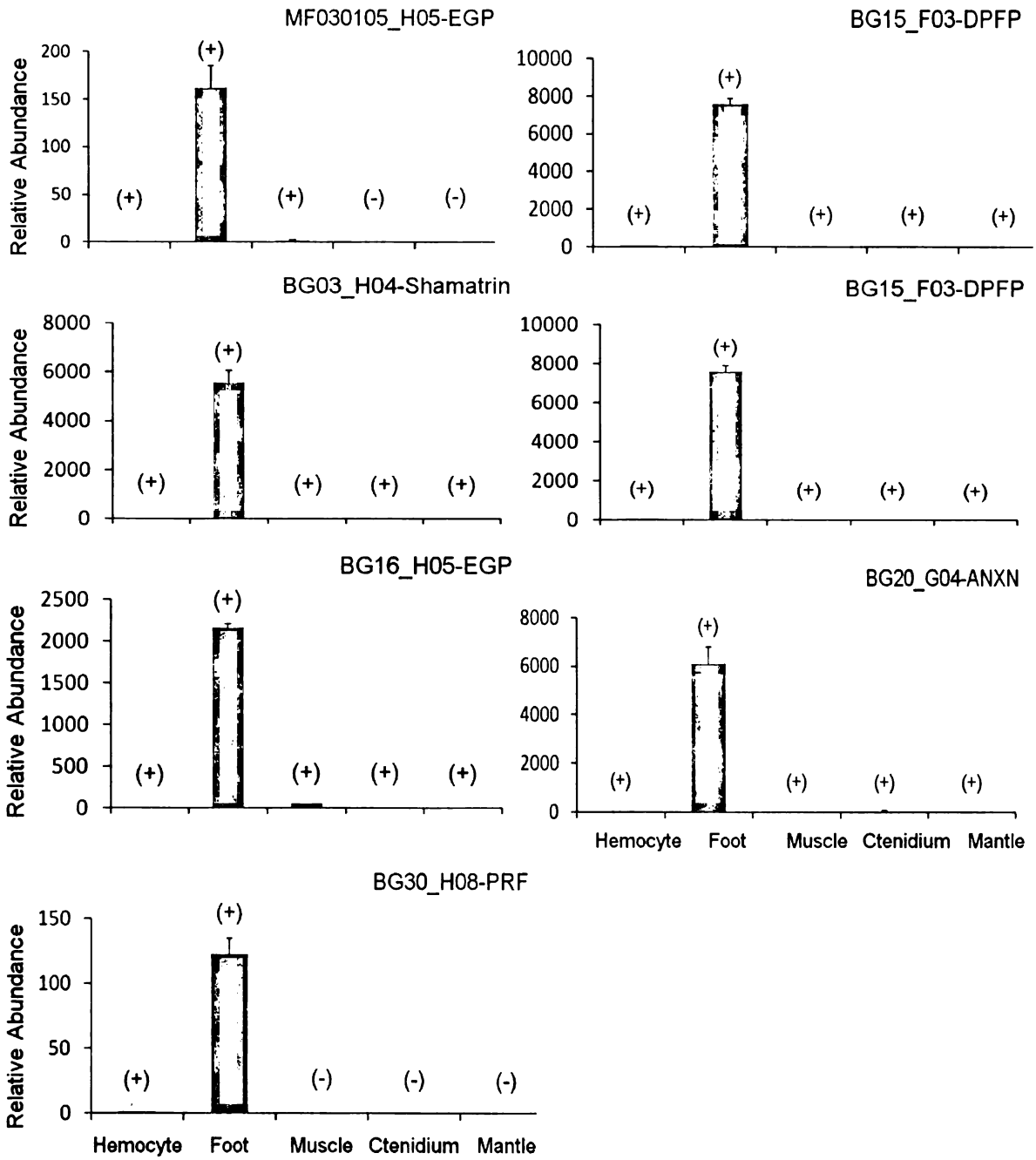


Figure 4-4 The distribution of the mRNA products of the selected genes within zebra mussel tissues.
(+) indicated the existence of the gene in the tissue; (-) suggested no detected transcripts in the tissue. For each gene, the relative expression levels of the gene in other tissues were calculated by using $2^{-\Delta\Delta Ct}$ model.

Table 4-1 The gene specific primers used for qRT-PCR.

Target gene	Name of the primer	Sequence
MF030105_H05	MF030105_H05_EGP-RT-Fw	5' -TCC GCC ACC ATA GTC GTC AT- 3'
	MF030105_H05_EGP-RT-Rv	5' -GGT AGC GAA CGC CCA AAA C- 3'
BG97/192_E09	BG97/192_E09_DPFP-RT-Fw	5' -CCG ATG GGC CAT ATG ATA AGA A- 3'
	BG97/192_E09_DPFP-RT-Rv	5' -ACG TCG ATG CCT TGC TTG TC- 3'
BG03_H04	BG03_H04_Smtrn-RT-Fw	5' -TGG CGG GTA TCC AGG AAA- 3'
	BG03_H04_Smtrn-RT-Rv	5' -CCG TAA CCA CCC CAT TTG C- 3'
BG15_F03	BG15_F03_DPFP-RT-Fw	5' -TCG ATG CCT TGC TTG TCT TG- 3'
	BG15_F03_DPFP-RT-Rv	5' -ACA CCG ATG GGC CAT ATG AT- 3'
BG16_H05	BG16_H05_EGP-RT-Fw	5' -CGA TTG CTC CGG TAG CAA AC- 3'
	BG16_H05_EGP-RT-Rv	5' -TAT AAC CAT CCG CCA CCA TAG TT- 3'
BG20_G04	BG20_G04_IRCI-RT-Fw	5' -CGG CGA CCA AGA AAA ATA CG- 3'
	BG20_G04_IRCI-RT-Rv	5' -GAC CCA CCA ACG GCA TTC- 3'
BG30_H08	BG30_H08_PRF-RT-Fw	5' -GCG GCG TTG CGT TGA G- 3'
	BG30_H08_PRF-RT-Rv	5' -CAT GGT CCA GGC CAA TAG GA- 3'
***18SrRNA**	18SrRNA_Forward	5'- GAC ACG GCT ACC ACA TCC AA -3'
	18SrRNA_Reverse	5'- CTC GAA AGA GTC CCG CAT TG -3'

* The primers used as internal reference for QRT-PCR.

Table 4-2 The genes that were differentially expressed after 12-hour reattachment. Selected by *t*-test $P < 0.05$, when Log (FC) > 0, the gene was upregulated during the byssogenesis; otherwise was downregulated.

Accession #	Gene ID	*P*	Log(FC)	Homologues
AM230194*	BG28_F04	0.00235	-1.03	N/A
AM230436*	BG29_A08	0.00548	0.362	CAA55094| Cytochrome oxidase III [*Kellia laperousii*]
AM230139*	BG20_G04	0.00783	0.925	XP_001134471| Annexin VII [*Dictyostelium discoideum* AX4]
AM229777*	BG31_H08	0.00875	-0.432	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM229723	BG32_C03	0.01453	-0.334	AAC39039| Foot protein 1 precursor [*Dreissena polymorpha*]
AM229947	BG26_E11	0.01684	0.717	NP_504107| Neuropeptide-like protein 31 [*Caenorhabditis elegans*]
AM230020	MF030105_H05	0.02056	-0.637	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM229916	BG13_E10	0.02386	0.429	AAT92111| Exocrine gland peptide NPL-2 [*Ixodes pacificus*]
AM230391	BG17_E04	0.02853	0.275	N/A
AM230219	BG12_B10	0.02952	0.647	N/A
AM230428	BG24_C06	0.0342	0.345	N/A
AM230236	BG15_C08	0.0445	0.33	N/A
AM230086	BG10_C01	0.04607	0.364	BAE93436| Shematrin-4 [*Pinctada fucata*]

* Differentially expressed genes with *t*-test $P < 0.01$

Table 4-3 The genes that were differentially expressed after 1-day reattachment. Selected by the *t*-test $P < 0.05$, when Log (FC) > 0, the gene was upregulated during the byssogenesis; otherwise was downregulated.

Accession #	Gene ID	*P*	Log(FC)	Homologues
AM230212*	BG10_A09	0.00135	-0.511	N/A
AM229999*	BG12_D08	0.00312	0.335	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM230238	BG34_F07	0.01997	0.657	N/A
AM230370	BG17_D01	0.02299	-0.503	N/A
AM229820	BG12_A09	0.02401	0.307	AAT92111| Exocrine gland peptide NPL-2 [*Ixodes pacificus*]
AM230030	BG33_E04	0.02495	0.296	N/A
AM229726	BG97/192_E09	0.02747	0.685	AAF75279| Byssal protein Dpfp1 precursor [*Dreissena polymorpha*]
AM230280	BG97/192_H07	0.03323	0.428	N/A
AM230020	MF030105_H05	0.0338	-0.505	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM230222	BG21_A07	0.03851	-0.196	N/A
AM230177	BG09_C10	0.0396	0.253	P27085| Ribosomal protein S26 [*Octopus vulgaris*]
AM229905	BG15_E06	0.04041	0.432	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM230398	BG18_C01	0.04853	-0.2	N/A

* Differentially expressed genes with *t*-test $P < 0.01$

Table 4-4 The genes that were differentially expressed after 2-day reattachment. Selected by the *t*-test $P < 0.05$, when Log (FC) > 0, the gene was upregulated during the byssogenesis; otherwise was downregulated.

Accession #	Gene ID	*P*	Log(FC)	Homologues
AM230170*	BG14_A11	0.00099	-0.62	ABI52762| 60S ribosomal protein L27 [*Argas monolakensis*]
AM230268*	MF030105_B12	0.00311	-0.862	N/A
AM230436*	BG29_A08	0.00836	-0.291	CAA55094| Cytochrome oxidase III [*Kellia laperousii*]
AM230156	BG30_H08	0.0103	0.477	BAB12683| Polypeptide release factor 3 [*Yarrowia lipolytica*]
AM230135	BG07_E12	0.01653	-0.33	BAE93433| Shematrin-1 [*Pinctada fucata*]
AM229907	BG33_D04	0.01699	-0.295	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM229915	BG10_B03	0.01749	0.403	N/A
AM230168	BG16_D03	0.01875	-0.29	AAN05585| Ribosomal protein L22 [*Argopecten irradians*]
AM230262	BG29_D12	0.02504	-0.626	N/A
AM230201	BG97-192_C04	0.0251	-0.324	N/A
AM230279	BG33_B06	0.02726	-0.392	N/A
AM230213	BG03_E10	0.02799	0.387	N/A
AM230299	BG05_A01	0.02951	-0.382	N/A
AM230428	BG24_C06	0.03067	-0.301	N/A
AM229867	BG16_B01	0.03153	-0.275	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM230089	BG20_F04	0.03511	-0.376	N/A
AM229837	BG97/192_G04	0.03959	-0.339	AAT92111| Exocrine gland peptide NPL-2 [*Ixodes pacificus*]
AM230302	BG05_D06	0.04032	-0.255	N/A
AM230270	MF030105_D06	0.04264	-0.264	N/A
AM229724	BG07_H12	0.04388	-0.269	AAF75279| Byssal protein Dpfp1 precursor [*Dreissena polymorpha*]

* Differentially expressed genes with *t*-test $P < 0.01$

Table 4-5 The genes that were differentially expressed after 3-day reattachment. Selected by the *t*-test $P < 0.05$, when Log (FC) > 0, the gene was upregulated during the byssogenesis; otherwise was downregulated.

Accession #	Gene ID	*P*	Log(FC)	Homologues
AM230436*	BG29_A08	0.00248	-0.341	CAA55094| Cytochrome oxidase III [*Kellia laperousii*]
AM230391*	BG17_E04	0.00254	-0.339	N/A
AM230270*	MF030105_D06	0.00562	-0.375	N/A
AM229957	BG26_B09	0.01178	-0.322	N/A
AM230080	BG23_C05	0.01255	-0.314	CAA10192| Glycine-rich protein 2 [*Cicer arietinum*]
AM230258	BG16_G06	0.01553	-0.518	With anaphylatoxin domain (PS01177) and EGF-like region (PS00022)
AM229881	BG97/192_B12	0.02036	-0.558	N/A
AM230428	BG24_C06	0.02568	-0.311	N/A
AM230440	BG31_F01	0.02893	-0.281	N/A
AM230145	BG28_F03	0.03174	-0.197	N/A
AM230214	BG16_G04	0.03261	0.242	N/A
AM230387	BG17_D06	0.03391	-0.212	N/A
AM230156	BG30_H08	0.03449	0.382	BAB12683| Polypeptide release factor 3 [*Yarrowia lipolytica*]
AM230306	BG28_E01	0.04006	-0.227	N/A
AM230220	BG08_B01	0.04406	-0.285	N/A
AM229950	MF030105_E03	0.04621	-0.404	NP_504107| Neuropeptide-like protein 31 [*Caenorhabditis elegans*]
AM229946	BG11_E06	0.04767	0.339	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]

* Differentially expressed genes with *t*-test $P < 0.01$

Table 4-6 The genes that were differentially expressed after 4-day reattachment. Selected by the *t*-test $P < 0.05$, when Log (FC) > 0, the gene was upregulated during the byssogenesis; otherwise was downregulated.

Accession #	Gene ID	*P*	Log(FC)	Homologues
AM229866*	BG28_E10	0.00716	0.321	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM229864*	BG16_H05	0.00983	-0.775	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM230217	BG11_F12	0.01064	0.311	N/A
AM230230	BG15_A03	0.01095	0.317	N/A
AM230173	BG22_A12	0.0148	0.464	ABR23471| 40S ribosomal protein S11 [*Ornithodoros parkeri*]
AM229737	BG32_C03	0.01774	0.276	AAC39039| Foot protein 1 precursor [*Dreissena polymorpha*]
AM229913	BG31_D01	0.02009	-0.459	NP_505834| Neuropeptide-like protein 33 [*Caenorhabditis elegans*]
AM230212	BG10_A09	0.02227	0.335	N/A
AM230209	BG08_E12	0.02575	0.256	N/A
AM230170	BG14_A11	0.02715	-0.388	ABI52762| 60S ribosomal protein L27 [*Argas monolakensis*]
AM229738	BG15_F03	0.03366	0.254	AAF75279| Byssal protein Dpfp1 precursor [*Dreissena polymorpha*]
AM230148	BG09_D03	0.03713	0.24	N/A
AM230094	BG33_H08	0.03736	-0.433	BAA76901| Choriogenin Hminor [*Oryzias latipes*]
AM230064	BG29_B07	0.04237	-0.439	BAE93436| Shematrin-4 [*Pinctada fucata*]
AM230258	BG16_G06	0.04583	-0.42	With anaphylatoxin domain (PS01177) and EGF-like region (PS00022)
AM229789	BG06_C03	0.04902	0.332	N/A

* Differentially expressed genes with *t*-test $P < 0.01$

Table 4-7 The genes that were differentially expressed after 7-day reattachment. Selected by the *t*-test $P < 0.05$, when Log (FC) > 0, the gene was upregulated during the byssogenesis; otherwise was downregulated.

Accession #	Gene ID	*P*	Log(FC)	Homologues
AM230258*	BG16_G06	0.00027	-0.876	With anaphylatoxin domain (PS01177) and EGF-like region (PS00022)
AM229950*	MF030105_E03	0.00129	-0.721	NP_504107| Neuropeptide-like protein 31 [*Caenorhabditis elegans*]
AM229748*	BG31_E04	0.0031	0.817	N/A
AM229864*	BG16_H05	0.00944	-0.798	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM229810	BG29_D08	0.01067	0.629	AAT92111| Exocrine gland peptide NPL-2 [*Ixodes pacificus*]
AM230230	BG15_A03	0.01075	0.325	N/A
AM230148	BG09_D03	0.01202	0.302	N/A
AM229949	BG05_C08	0.0221	0.523	N/A
AM230173	BG22_A12	0.02354	0.437	ABR23471| 40S ribosomal protein S11 [*Ornithodoros parkeri*]
AM229798	BG06_C03	0.0271	0.385	N/A
AM230036	BG03_H04	0.02725	-0.379	BAE93436.1| Shematrin-4 [*Pinctada fucata*]
AM230106	BG05_C07	0.03377	0.442	N/A
AM230377	BG16_D10	0.03458	0.973	N/A
AM230211	BG05_D02	0.03577	-0.449	N/A
AM230407	BG18_G06	0.04034	0.778	N/A
AM229725	BG12_C01	0.04359	0.175	AAF75279| Byssal protein Dpfp1 precursor [*Dreissena polymorpha*]
AM229736	BG32_C03	0.04386	0.236	AAC39039| Foot protein 1 precursor [*Dreissena polymorpha*]
AM230404	BG18_F03	0.04685	-0.548	N/A
AM230344	BG12_G06	0.04945	-0.44	N/A
AM229727	BG15_H10	0.0499	-0.668	AAF75279| Byssal protein Dpfp1 precursor [*Dreissena polymorpha*]

* Differentially expressed genes with *t*-test $P < 0.01$

Table 4-8 The genes that were differentially expressed after 21-day reattachment. Selected by the *t*-test $P < 0.05$, when Log (FC) > 0, the gene was upregulated during the byssogenesis; otherwise was downregulated.

Accession #	Gene ID	*P*	Log(FC)	Homologues
AM230363*	BG15_D08	0.00051	-0.697	N/A
AM229907*	BG33_D04	0.00192	0.47	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM230179*	BG08_B03	0.00247	0.792	N/A
AM229753*	BG23_D01	0.00559	0.553	AAT92111| Exocrine gland peptide NPL-2 [*Ixodes pacificus*]
AM230251*	BG29_B09	0.00668	-0.823	N/A
AM230111*	BG30_G02	0.00835	-0.749	AAN10061| Kunitz-like protease inhibitor [*Ancylostoma caninum*]
AM230206	BG29_B05	0.01314	-0.735	N/A
AM230302	BG05_D06	0.01414	0.364	N/A
AM230408	BG97/192_E03	0.01813	-0.592	N/A
AM502285	BG09_F07	0.01987	-0.642	AAF75279| Byssal protein Dpfp1 precursor [*Dreissena polymorpha*]
AM230248	BG25_H11	0.01997	0.398	With Antifreezei domain and Thiolase domain
AM230143	BG29_F09	0.02004	0.529	N/A
AM230253	BG34_F09	0.02196	-0.392	N/A
AM229730	BG05_D07	0.02324	0.517	N/A
AM229728	BG21_E12	0.02548	-0.586	AAF75279| Byssal protein Dpfp1 precursor [*Dreissena polymorpha*]
AM230344	BG12_G06	0.02555	0.58	N/A
AM229901	BG15_E03	0.02856	-0.815	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM230390	BG33_C06	0.03041	0.425	N/A
AM230036	BG03_H04	0.03334	0.416	BAE93436| Shematrin-4 [*Pinctada fucata*]
AM229892	BG32_F04	0.03801	0.898	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM230172	BG26_B01	0.03833	-0.394	AAK95191| 40S ribosomal protein S9 [*Ictalurus punctatus*]
AM230048	BG16_E06	0.04033	0.689	N/A
AM229879	BG32_F05	0.04249	-0.427	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM229876	BG23_F02	0.04295	0.363	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM229878	BG34_H02	0.04315	-0.573	AAV80789| Exocrine gland peptide [*Ixodes scapularis*]
AM229802	BG22_H06	0.04316	0.381	N/A
AM230095	BG16_C02	0.04396	-0.408	BAE93436.1| Shematrin-4 [*Pinctada fucata*]
AM230123	BG12_G10	0.04409	-0.545	N/A
AM229752	BG33_B02	0.04517	0.453	AAT92111| Exocrine gland peptide NPL-2 [*Ixodes pacificus*]

* Differentially expressed genes with *t*-test $P < 0.01$

CHAPTER V

Factorial Microarray Analysis of Zebra Mussel (*Dreissena polymorpha*) Adhesion

Abstract

The zebra mussel (*Dreissena polymorpha*) has been well known for its expertise in attaching to substances under the water. Studies in past decades on this underwater adhesion focused on the adhesive protein isolated from the byssogenesis apparatus of the zebra mussel. However, the mechanism of the initiation, maintenance, and determination of the attachment process remains largely unknown. In this study, we used a zebra mussel cDNA microarray previously developed in our lab and a factorial analysis to identify the genes that were involved in response to the changes of four factors: temperature (Factor A), current velocity (Factor B), dissolved oxygen (Factor C), and byssogenesis status (Factor D). One hundred and seventeen probes in the microarray were found to be altered by at least one of the factors, while 18 genes were modulated by two of them. The transcription products of four selected genes, DPFP-BG20_A01, EGP-BG97/192_B06, EGP-BG13_G05, and NH-BG17_C09 were unique to the zebra mussel foot based on the results of quantitative reverse transcription PCR (qRT-PCR). The expression profiles of these four genes under the attachment and non-attachment were also confirmed by qRT-PCR and the result is accordant to that from microarray assay. The *in situ* hybridization with the RNA probes of two identified genes DPFP-BG20_A01 and EGP-BG97/192_B06 indicated that both of them were expressed by a type of exocrine gland cell located in the middle part of the zebra mussel foot. It turns out that the factorial design and analysis of

the microarray experiment is a reliable method to identify the influence of multiple factors on the expression profiles of the probesets in the microarray; therein it provides a powerful tool to reveal the mechanism of zebra mussel underwater attachment.

Introduction

The Eurasian, non-native mollusk, the zebra mussel (*Dreissena polymorpha*), invaded the Laurentian Great Lakes in the 1980s and has expanded since then to other regions in North America (Hebert *et al.* 1989; Benson & Raikow 2009). The spread of the zebra mussel has been followed by ecologic and economic devastation of unprecedented magnitude (Nalepa & Schloesser 1993). Through their unique ability to attach to underwater surfaces, the invading mussels interfered with navigation and flow of industrial effluents, a problem that has caused billions of dollars in losses (New-York-Sea-Grant 1994b, a). Moreover, the selective filtering capacity of the zebra mussel has altered the microbial communities in the Great Lakes to the extent that some resident species have become endangered and others are facing extinction (Nalepa & Schloesser 1993). The problem is compounded by the fact that all control methods used to minimize the zebra mussel invasion impacts have been unsuccessful (Snyder *et al.* 1997; Molloy 1998; Bially & MacIsaac 2000). As a result, a pressing need has emerged to unravel the strategies used by the invading mollusk to survive and mechanisms governing its firm attachment to underwater surfaces.

The zebra mussel is unique among freshwater bivalves in that in the complete life cycle it keeps its byssus, a group of exocrine glands embedded in the mussel's foot that secretes threads with attachment pads through byssogenesis (Rzepecki & Waite 1993a, b). Our previous studies have identified 716 genes that are unique to zebra mussel byssus using the suppression subtractive hybridization (SSH) cDNA library technique (Xu & Faisal 2008). Besides the classic *D. polymorpha* foot protein -1 and -2 (Dpfp-1 and -2), a

number of molecules with different putative functions have also been produced by byssus glands, such as the Exocrine gland peptides homologous to the salivary gland peptides of *Ixodes* spp. (Francischetti *et al.* 2005), host defense related molecules, and the proteins without predictable functions. The subsequently developed cDNA microarray with these 716 unique genes was used to compare the expression profiles of the gland unique genes under different attachment status. The results indicated that the expressions of 52 genes are either up or down regulated in the attachment status (Xu & Faisal 2009a). It is suggested that the microarray technique can be used as a high-throughput tool to study the mechanism of zebra mussel attachment at the molecular level.

Zebra mussel attachment is influenced by several environmental factors including water temperature (Clarke & McMahon 1996c), dissolved oxygen level (Clarke & McMahon 1996b), and current velocity (Clarke & McMahon 1996a). The mechanism(s) by which these factors influence zebra mussel attachment is currently unknown and involves a large number of molecules. Identifying the influence of environmental conditions on gene encoding is rather complicated and may require a novel way of computation and statistical analysis. In this context, the studies of retinal development in zebra fish (*Danio rerio*) have adapted the multifactorial analysis to cDNA microarray data and proved useful in deciphering the complicated developmental process (Leung *et al.* 2008). The factorial microarray analysis allows us to analyze the effects of more than one independent variable (the main factors involved in this study) on a dependent variable (the differentiation of gene expression) (Wu & Hamada 2000; Montgomery 2005). In this study, a non-reference loop experimental design was performed to test the effects of multiple environmental factors on the expression profiles of the byssus unique

genes. This experimental design has been reported as a reliable design that leads to less disparity in precision and power comparisons between any two groups (Tempelman 2005).

The purpose of the factorial design that we are going to describe here is to identify the influence of the temperature, water current velocity, dissolved oxygen, and byssogenesis status on the expression profiles of the molecules unique to zebra mussel byssus glands to thereby to better understand the mechanism of zebra mussel attachment under the complex underwater environment.

Materials and Methods

Zebra mussel collection and maintenance

Zebra mussels used in this study were collected from Vineyard Lake in Brooklyn, MI, USA (Latitude: 42°4'59''N; Longitude: 84°12'34"W). The mussels were thoroughly cleaned and rinsed with deionized water several times before they were allowed to acclimate in aerated, filtered Vineyard Lake water for eight weeks. The mussels were kept in a glass tank and fed weekly with a pure culture of the algus *Ankistrodesmus falcatus*.

Treatments of zebra mussels and experimental design

Three environmental factors were involved in this study including temperature (Factor A), current velocity (Factor B), and dissolved oxygen level (Factor C). The status of byssogenesis was considered as Factor D. The model used in this study is:

$$y_g = \mu_g + A + B + C + D + \varepsilon$$

This one way model includes only the basal expression value of the probe (μ_g), the main effects coefficients (A, B, C, D), and the error term (ε). In each effect, two levels were created. To create two different attachment statuses of the mussels, we severed the byssal threads from randomly selected mussels and put them back into the tanks in different orientations. The mussels lying on the ventral side of their shells for 48 hours were considered to be attached individuals with byssogenic activity (A) while the ones with ventral sides facing up for 48 hours were taken as detached mussels without byssogenesis (D). Two temperature levels were used in this study: room temperature (R) and a low

temperature level at 4 °C. The two dissolved oxygen (D. O.) levels used were the normal D. O. level (N), which was maintained by airstones at about 10 mg/L, and the lower D. O. level, which was about 5 mg/L without airstones. The current of the water was performed by using VWR 371 Hotplate/Stirrer (VWR International LLC, West Chester, PA) and a magnetic stirring bar with a stirring speed of 60 rpm (F). The static water was used as non-current condition of Factor B in the experiment (S).

Sixteen treatment combinations were obtained from this 2×2×2×2 factorial experimental design. To maintain high statistical power for the analyses of each factor effect, a loop experimental design was performed with four biological replicates for each treatment (Figure 5-1) (Tempelman 2005). Every two samples connected with an arrow were used to hybridize with one microarray slide. The dye labeling and hybridizations are indicated by Figure 5-1. Four null hypotheses were tested and the interpretation of the rejection of each null hypothesis is listed in Table 5-1.

RNA extraction and cDNA synthesis

The total RNAs used for microarray hybridization were extracted from pools of zebra mussel feet with six individuals in each pool. The total RNA purification kit, 5-PRIME PerfectPure RNA Tissue Kit (5 PRIME Inc, Gaithersburg, MD) was used for RNA purification and the protocol followed the manufacturer's instructions. The RNAs were then reverse transcribed to single strand cDNAs by using SuperScript *Plus* Indirect cDNA Labeling Kit (Invitrogen, Carlsbad, CA) containing Alexa Fluor 555 & 647 fluorescent dyes. The cDNA syntheses and dye labeling procedures were all followed as described by the kit's instructions.

Microarray hybridization and data analyses

The zebra mussel byssus cDNA microarray was developed and printed in the Center for Animal Functional Genomics, Michigan State University (Xu & Faisal 2009a). The GeneTAC HybStation (Genomic Solutions) was used as the hybridizer for all hybridizations and an 18-hour step-down protocol was applied as described in our previous study (Xu & Faisal 2009a). The hybridized dual-channel array slides were scanned by GenePix 4000B two-laser Scanner (Molecular Devices, Downingtown, PA), and GenePix Pro 6.0 (Molecular Devices) software was then used for image processing and spot intensity file generation. The spot density files were analyzed with the Limma software package (Smyth 2005). The preliminary data processing includes background correction, within array normalization, and between array normalization. The differentially expressed genes were determined by least square regression. The coefficients of the analysis were the logarithmic values of the ratios between two levels in each factor. The homologues of the genes are listed with the accession numbers and putative functions, as well as with the names of the species (Table 2-5). The up- or down-regulation of the factor to probes was decided by the coefficient. The hierarchical clustering was done with the software Genesis (Sturn *et al.* 2002). The dataset of this microarray study was deposited in Gene Expression Omnibus (GEO) with the series accession number GSE16397. Some of the identified genes with similar primary structures were differentiated by multiple sequence alignment with ClustalW (Thompson *et al.* 1994).

Microarray data validation

The one-step quantitative reverse transcription PCR (qRT-PCR) was performed with *Power* SYBR® Green RNA-to-C_T™ *1-Step* Kit (Applied Biosystems Inc., Foster City, CA) to identify the distribution of the candidate genes in zebra mussel tissues. The mussels were also collected from Vineyard Lake in Brooklyn, MI. The shells of randomly selected individuals were swabbed with 70% ethanol, followed by a soak in 150 ml sterilized double-distilled water containing 5,000 U penicillin, 5 mg streptomycin, and 10 mg gentamicin for 30 minutes (Davids & Yoshino 1998). Thereafter, the mussels were transferred to sterilized double distilled water. The RNA samples were extracted from hemocytes, feet, muscles, ctenidia, and mantles of zebra mussels. The protocol for hemocyte collection was described by our previous studies (Xu & Faisal 2007, 2009b) and the other samples were taken by sterilized scissors and scalpels. The primers were designed by Primer Express™ (Applied Biosystem Inc.). The primers used in this study are listed in Table 6.

The qRT-PCR was also applied to validate the expression profiles of the genes that were affected by the attachment status identified by microarray. The zebra mussels were allowed to fully attach themselves on glass Petri-dishes for six weeks. Then byssal threads from all individuals were cut off with scalpels; thereafter, five mussels were dissected and the RNA extracted from their feet was considered as group 0. The rest of the mussels were put back in the Petri-dishes in water, and allowed to regenerate their byssal threads. The RNA samples were extracted from the feet of the mussels at 12 hours, one day, two days, and three days post re-attachment. Two biological replicates were used for each group. Then for each candidate gene, the qRT-PCR was performed with

these five group samples. The information of the gene specific RT Primers are listed in Table 6.

RNA fluorescence *in situ* hybridization (FISH)

In order to locate the transcripts of two selected genes *in situ*, fluorescence *in situ* hybridization (FISH) was used. The zebra mussel foot tissue was fixed in 2% paraformaldehyde phosphate buffer saline (0.02M NaH_2PO_4, 0.0077M Na_2HPO_4, 1.4M NaCl, 2% w/v paraformaldehyde, pH 8.0) for five hours at room temperature. The sample was then conserved in absolute ethanol until needed. The conserved sample was buried in paraffin and longitudinally sectioned at the same level. The five μm thick sectioned sample was obtained and put on an RNase- and DNase-free glass slide. Ten slides were obtained with one stained by H&E method.

The antisense RNA strand of a fragment of each gene was produced. Briefly, the PCR product of the gene fragment was cloned into pGEM®-T Easy Vector Systems (Promega U.S., Madison, WI). Then the positive colony was picked for sequencing with the M13 forward primer to clarify the direction of the insert of the recombinant plasmid. Once the direction of the insert was known, the T7 or Sp6 RNA polymerase was selected to synthesize the antisense RNA with the linearized recombinant plasmid. In this study, T7 polymerase was used on both probes. The probe DPFP-BG20_A01 was labeled with fluorescent dye Alexa 488 (Invitrogen) while the probe EGP-BG97/192_B06 was labeled with Alexa 594 (Invitrogen). The steps of the RNA synthesis and labeling can be found in the instructions of the two kits: FISH Tag^{TM} RNA Green Kit (Invitrogen) and Fish Tag^{TM} RNA Red Kit (Invitrogen). The RNA FISH was performed following the

recommended protocol in the appendix of the Fish Tag^{TM} RNA Kit instruction. The hybridization buffer mentioned in the protocol was replaced by ULTRAhybTM Ultrasensitive Hybridization Buffer (Applied Biosystems/Ambion, Austin, TX). The hybridized slides were visualized by Olympus BX41 (Olympus America Inc., Center Valley, PA) microscope under the excitation of a mercury lamp and the images were captured by Olympus DP25 digital camera (Olympus America Inc.). The image processing software DP2-BSW (Olympus America Inc.) was used to combine the different color channels into a single image.

Results

Overview of the factorial analysis results

Microarray data analyses suggest that each of the factors tested (temperature, stirring, dissolved oxygen level and accessibility to attachment surfaces) had altered the expression of zebra mussel foot genes. The genes whose expressions had been altered by each of the four factors are listed in Tables 2-5. While the gene expression profiles had been altered, however, the type and numbers of differentially expressed genes differed with each experimental factor used. At a 95% confidence level ($P < 0.05$), 73 genes (10% of the total genes on the slide) were differentially expressed by the accessibility of the mussel to attachment surfaces (Table 2). At the same confidence level, 27 genes (4% of the total number of genes) were found either up- or down-regulated with the temperature increased from 4 °C to the room temperature of 22 °C (Table 3). Twenty-six of the genes (4% of the total genes) were modulated by the level of dissolved oxygen (D.O.), which was close to that in the temperature group (Table 4). Surprisingly, only nine genes, 1% of the whole probesets on the array, were identified with their expression levels influenced by the current velocity (Table 5). With the *P* values set at 0.01, fewer probes were identified. Eleven probes had their expression profiles altered by the status of attachment (Table 2). Five genes demonstrated significantly different expression levels with the temperature change (Table 3). With the level change, each of the other two environmental factors, D.O. and current velocity, modulated the transcription profiles of two genes only (Table 4 and 5).

Ten major clusters were generated using the hierarchical clustering method on the probesets whose expression profiles were modulated by at least one of the factors (Figure 5-2). The genes in each cluster had similar expression patterns under the effects of the four factors. For instance, three genes had been grouped into cluster A, suggesting that the expression profiles of the three genes were significantly upregulated when the temperature was decreased or when the mussel was undergoing byssogenesis. Similarly, cluster B contained seven genes whose expression patterns were upregulated by water flow, through were downregulated by byssogenesis (Figure 5-2).

Among the genes identified by microarray analysis, some genes demonstrated dramatic expression level changes in response to the combination of more than one factor (Figure 5-3). For example, from the 59 genes affected by the the attachment status (Factor D), six genes responded to changes in temperature (Factor A); another set of six genes were altered by dissolved oxygen levels (Factor C); and two other genes were affected by water current velocity (Factor B). Similarly, two genes were regulated by both Factor C and Factor D, while only one gene was affected by the combination of Factor B and C or Factor B and D (Figure 5-3).

Among the genes that were altered by byssogenesis at the $P < 0.01$, five were homologous to invertebrate exocrine gland peptides (GenBank accession number AAV80789 and AAS92593), while three genes from those modulated by temperature were homologous to two other invertebrate exocrine gland peptides (Accession numbers AAV80789 and AAT92111). The multiple sequences alignment suggested that the five genes in Table 2 whose expression was altered at $P < 0.01$ were not identical. BG31_H01 and BG10_C05 shared some similarities while BG13_F10 and BG23_B03 were

structurally close to each other. The sequence BG97/192_B06 shared the least similarity to the other four genes in nucleotide sequences (Figure 5-4). Similarly, among the three genes from the temperature modulated gene group, BG23_D02 and BG03_B01 were closest in their primary structure and both of them were downregulated in response to the lower temperature. On the contrary, EST, BG27_B08 had its expression profile upregulated at low temperature (Figure 5-5).

Validation

Four genes that were significantly altered by byssogenesis status were selected for validation (Table 6). The qRT-PCR with RNA samples from five tissues of zebra mussels showed that the abundance of all four selected genes in the zebra mussel foot were extremely high as compared to other tissues. The RNA product of an excocrine gland peptide-like gene (EGP-BG97/192_B06) was detected in all selected samples; however, the expression level in the foot was significantly higher than in any of the others. The lowest expression level of EGP-BG97/192_B06 was in the ctenidium, which was 17399 ± 844 times less than that in the foot. Although the quantity of the gene EGP-BG97/192_B06 transcripts in the muscle was also higher than that in the ctenidium, the foot tissue produced 15 ± 1 times more gene product than the ctenidium cells. Another EGP gene, EGP-BG13_G05, was exclusively expressed in the foot and hemocyte while the ratio between expression levels in foot and hemocyte was 954 ± 56. The product of a Dpfp-like protein encoded gene, DPFP-BG20_A01 was detected in foot, muscle, and hemocyte samples. The abundance of the gene product in the foot was the highest, at 383 ± 63 times that in the hemocyte and 49 ± 8 times more than that in the muscle. Another gene that was expressed in all five tissues was a gene without known homologues in

GenBank database, NH-BG17_C09. Compared to the least amount of RNA in the ctenidium, the foot provided 87 ± 3 times more NH-BG17_C09 product (Figure 5-6). The comparisons of the expression levels of four selected genes between attached and detached status were also performed by qRT-PCR. As shown in Figure 5-7, compared to the detached mussels (D), the gene DPFP-BG20_A01 had significantly higher expression levels in attached groups (A) at 12 hours, 1 day, 2 days, and 3 days post-treatment. The most significant difference between A and D was shown 2 days post-treatment. The two EGP-like protein encoding genes, EGP-BG97/192_B06 and EGP-BG13_G05, demonstrated similar expression profiles during the byssogenesis; both were upregulated in attachment status at 1 day and 2 days post-treatment, while down-regulated at 12 hours and 3 days post-treatment. However, the most obvious difference between A and D appeared at Day 1 in gene EGP-BG97/192_B06 and one day later in gene EGP-BG13_G05. At most time points, gene NH-BG17_C09 acted as downregulated in group A. Only at Day 2, the expression level of this gene demonstrated higher level in the attached status than the detached. All four candidate genes demonstrated the same expression profiles to those identified by microarray assay.

RNA *in situ* hybridization

Three major byssus gland cells were observed from longitudinally sectioned zebra mussel foot with H&E stain; namely, stem-forming gland (SFG) cell, thread-forming gland (TFG) cell, and plaque-forming gland (PFG) cell (Figure 5-8). The SFG was located at the root of the foot containing the cells with the smallest size among the three types of gland cells (Figure 5-8a). The TFG cells were distributed in the middle section of the foot, along the epithelial cells on the surface. The color of the TFG cells was lighter

than the other two types of gland cells, and the size of TFG cells was larger than SFG cells but smaller than PFG cells (Figure 5-8b). The cells of PFG had the largest size but the amount of this type of cells was the least and all the PFG cells were embedded in a small area of the foot tip (Figure 5-8c).

The *in situ* hybridization results indicated that the transcription of both genes DPFP-BG20_A01 (Alexa 488 labeled) and EGP-BG97/192_B06 (Alexa 594 labeled) were not detectable at the foot root or tip sections (Figure 5-9 a and c); however, strong signals of DPFP-BG20_A01 mRNA and EGP-BG97/192_B06 mRNA were detected in the middle area of the foot containing the TFG cells. The cells that expressed DPFP-BG20_A01 and those that contain EGP-BG97/192_B06 mRNAs were largely overlapped, demonstrating yellow signal. Some of the TFG cells that exclusively expressed either DPFP-BG20_A01 (green) or EGP-BG97/192_B06 (red) were also observed (Figure 5-9b).

Discussion

The factorial analysis used in this study was designed as three enclosed loops encompassing four biological replicates for each treatment. This design allowed the estimation of the effects of the experimental factors with a minimal number of hybridized slides and a high statistical power. In contrast to conventional microarray analysis, the factorial analysis is unique as it can properly estimate the potential effects of an experimental condition that was not directly included in the hybridization (Leung *et al.* 2008). Most importantly, the factorial analysis of microarrays employed in this study is optimal for studies involving aquatic animals, as they are exposed simultaneously to a multitude of environmental factors that are in continuous fluctuation. Similar loop designs of microarray data were successfully applied by Wit *et al* (Wit *et al.* 2005) and Zou *et al* (Zou *et al.* 2005) and proved efficient in comparing gene expression alteration due to multiple treatments. Despite the clear advantages of the factorial analysis of microarray data, the false discovery rate (FDR) continues to be a problem that cannot be ignored (Pawitan *et al.* 2005). To minimize FDR, several statistical methods have been integrated into the microarray data analysis (28-32). In this study, we presented the data using both $P < 0.05$ and $P < 0.01$ as the cutoff for the microarray results. Subsequent validation demonstrated that many of the differential expression, of several genes with either P values in our study were indeed positive.

Among the complete probesets identified by this array study, 62% were altered in response to adhesion status (Factor D), therefore, subsequent study of individual probes focused on this group of genes that were most probably associated with the zebra mussel's ability to attach. The four candidate genes selected for qRT-PCR were from this

category. It was confirmed by qRT-PCR that after 48 hour treatment, the expression levels of these four genes were all upregulated during byssogenesis; however, curves of the relative expression levels of the four genes during the three-day byssognesis were not consistent. It is likely that these genes are involved at different stages during byssogenesis and are not always up- or downregulated. This periodical expression or depression may reflect the sequence of events of byssogenesis.

One of the identified genes, DPFP-BG20_A01, that had been validated by qRT-PCR is homologous to the *Dreissena polymorpha* foot protein-1 (Dpfp1), which was previously identified as an important protein for byssal thread structure (Rzepecki & Waite 1993b, a) and was located by immunohistochemistry along the ventral groove of the zebra mussel foot (Anderson & Waite 2000). The *in situ* hybridization performed in this study located DPFP-BG20_A01 in an area along the groove of the mussel's foot that is dominated by thread forming glands (Bonner & Rockhill 1994a). It is noticeable in this study that the expression profiles of many Dpfp-1 homologues were found in response to the change of adhesion status. This suggests that multiple foot proteins are involved in the byssogenesis of the zebra mussel in addition to the Dpfp-1. A number of foot proteins produced by byssus glands and playing different roles in the byssogenesis of Mytilids, another genus of mussels with a lifelong byssus apparatus, have been found over past decades (Benedict & Waite 1986; Papov *et al.* 1995; Waite *et al.* 1998; Waite & Qin 2001; Zhao & Waite 2006; Lin *et al.* 2007; Cha *et al.* 2008). It is very likely that, in the zebra mussel, there is also a group of foot proteins serving in different stages of byssogenesis.

Not only were foot proteins identified, but also the homologues of exocrine gland peptides (EGP) (Narasimhan *et al.* 2007) were commonly found to be differentially expressed during byssogenesis. The transcription of one of the EGP-like molecules, EGP-BG97/192_B06, was also observed in the TFG cells area. This observation underscores the importance of these molecules in zebra mussel adhesion and that byssogenesis in the zebra mussel involves a myriad of proteins. The primary sequence of the EGP-BG97/192_B06 suggests its distinction from Dpfp-1, despite the domination of its amino acids with proline, tyrosine, aspartic acid, lysine, threonine, and glycine residues, which is similar to Dpfp-1 (Anderson & Waite 1998). However, according to Francischetti *et al*, the homologous genes encode for salivary gland peptides with a variety of functions in ticks, such as anti-coagulant, anti-microbial, and oxidant metabolism (Francischetti *et al.* 2005). In our analysis, the ClastalW with zebra mussel EGP-like genes in response to the byssogenic activities suggested that these molecules were structurally similar but not identical to each other. Therefore, their potential functions during byssogenesis are most likely different. It suggested that during the adhesion a number of molecules with various functions were involved in this process. Some of the molecules may be directly involved in the generation of the byssal threads and adhesive proteins, while others may play alternate roles in the complicated process such as protecting the byssal proteins from degeneration by microbes, releasing the adhesive proteins, and modification of byssal thread components.

The results of the study demonstrated that that very few genes had their expression profiles altered by more than one experimental factor. This suggests that each of the experimental factors affected only a few zebra mussel genes involved in

byssogenesis independently; i.e., no uniformal response to all factors exists. A series of experiments was done by Clark and McMahon in 1996 to analyze the effects of temperature (Clarke & McMahon 1996c), hypoxia (Clarke & McMahon 1996b), and current velocity (Clarke & McMahon 1996a) on byssogenesis rate under laboratory conditions. The change of the water temperature seems to have the most significant effect on byssogenesis rate. From 5 °C to 30 °C, every 10 °C shift caused dramatic difference in the rate of byssal thread regeneration. Even 5 °C increase led to a significant increase of the byssogenesis rate when the temperature was above 25 °C (Clarke & McMahon 1996c). Our results demonstrated that the number of differentially expressed genes in response to temperature changes was the highest of the four factors tested (27 genes). On the contrary, the byssogenesis rate of the zebra mussel showed the least sensitivity to the change of current velocity. Only in a certain flow speed (0.2 m/s) was the byssogenesis rate of zebra mussel significantly higher than that in the other speeds (0.1 m/s, 0.15 m/s, 0.27 m/s) (Clarke & McMahon 1996a). Consistently, the least number of genes (9 genes) identified by microarray showed the significantly differential expressions during the change of the current velocity. Comparing the genes significantly altered by the four experimental factors, we can easily see that adhesion status, temperature, and dissolved oxygen level tends to affect the byssogenesis by modulating similar genes. For example, the homologues of Dpfp-1, EGP, *C. elegans* neuropeptide-like proteins (NP_504109 and NP_505834), and oyster shematrins (BAE93436) (Yano *et al.* 2006) were widely identified. It is possible that the temperature and oxygen level along with the adhesion status, affect the byssogenesis through the similar pathway, while the current velocity seems to control this process through a different pathway.

For the genes that were altered by multiple experimental factors, it is also very interesting to observe how environmental factors can alter gene expression in different directions. For example, BG23_B03 is downregulated in attached status, while is upregulated with the decrease of the temperature; while BG97/192_B06 can be upregulated in the attached status and downregulated by the dissolved oxygen level. What these gene-environmental factor interactions mean to zebra mussel attachment and associated mechanisms, cannot be easily answered, primarily due to the fact that the putative functions of these genes remain largely unknown. In conclusion, the factorial analysis of the microarray results provides important information on how their expression is influenced by important environmental factors. This newly generated knowledge will help to better understand the molecular mechanism of zebra mussel underwater adhesion. The factorial analysis of zebra mussel adhesion mechanism has provided for us a great tool to learn about the role of the four experimental factors during the zebra mussel attachment under controlled laboratory conditions. One can easily speculate that in the mussel's complex aquatic habitat more genes are likely to be involved in the attachment process. Additional studies, also using cDNA microarray factorial analysis, deem necessary to decipher the mechanisms governing zebra mussel byssogenesis, a process that has devastated the fragile ecosystem in infested waters of the Laurentian Great Lakes basin.

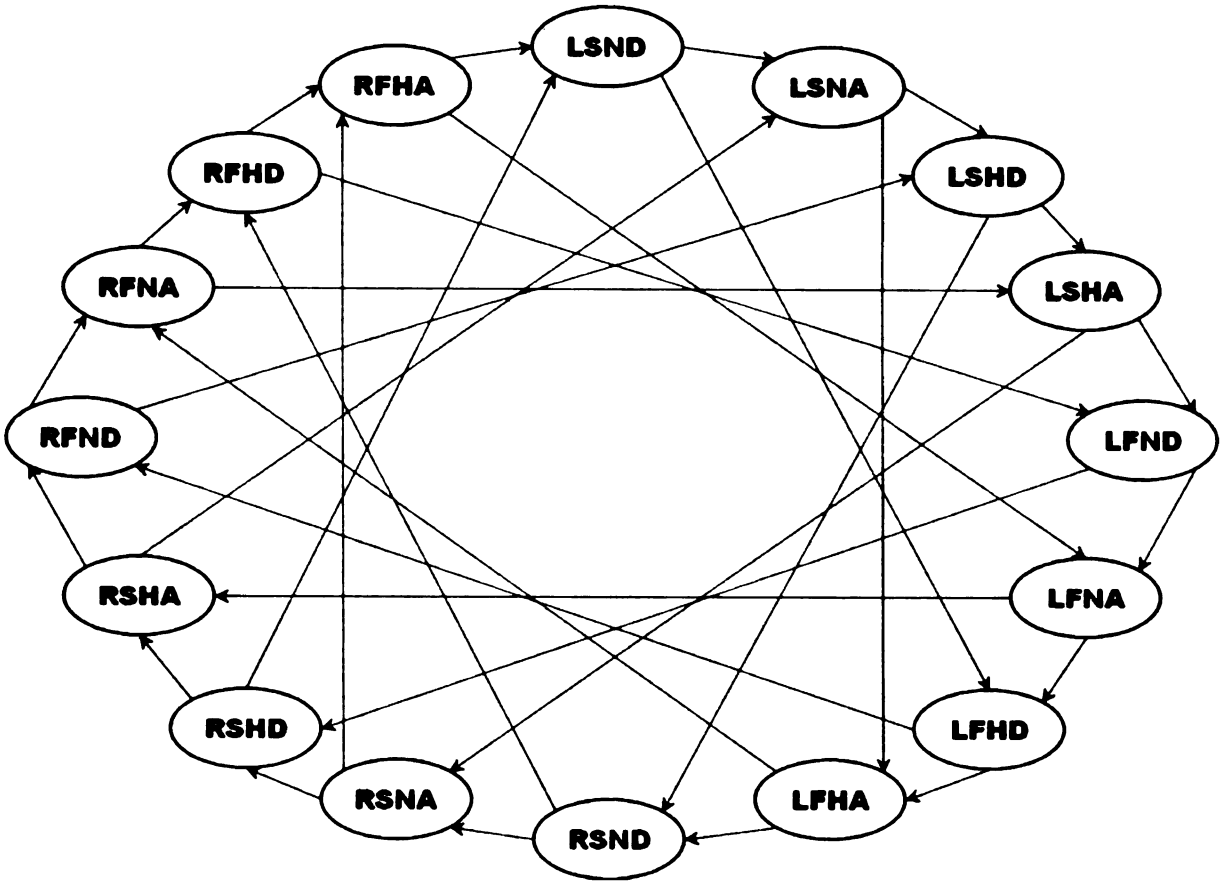


Figure 5-1 The non-reference interwoven loop design for microarray analysis. Effects of temperature, dissolved oxygen, current velocity, and byssogenesis status on gene expression in the zebra mussel foot using cDNA microarray analysis. The figure displays the different treatment combination selected as per the loop design approach (Churchill 2002). Each arrow represents one microarray hybridization. The start of the arrow stands for the sample labeled with dye Alexa 647 while the end point of an arrow represents the sample labeled with dye Alexa 555. L stands for low temperature (4 °C) while R stands for a higher temperature (22 °C); S stands for static water while F means flow water stirred by magnetic stirring bars; H represents the low dissolved oxygen level (5mg/L) while the N represents the normal dissolved oxygen level (10 mg/L).

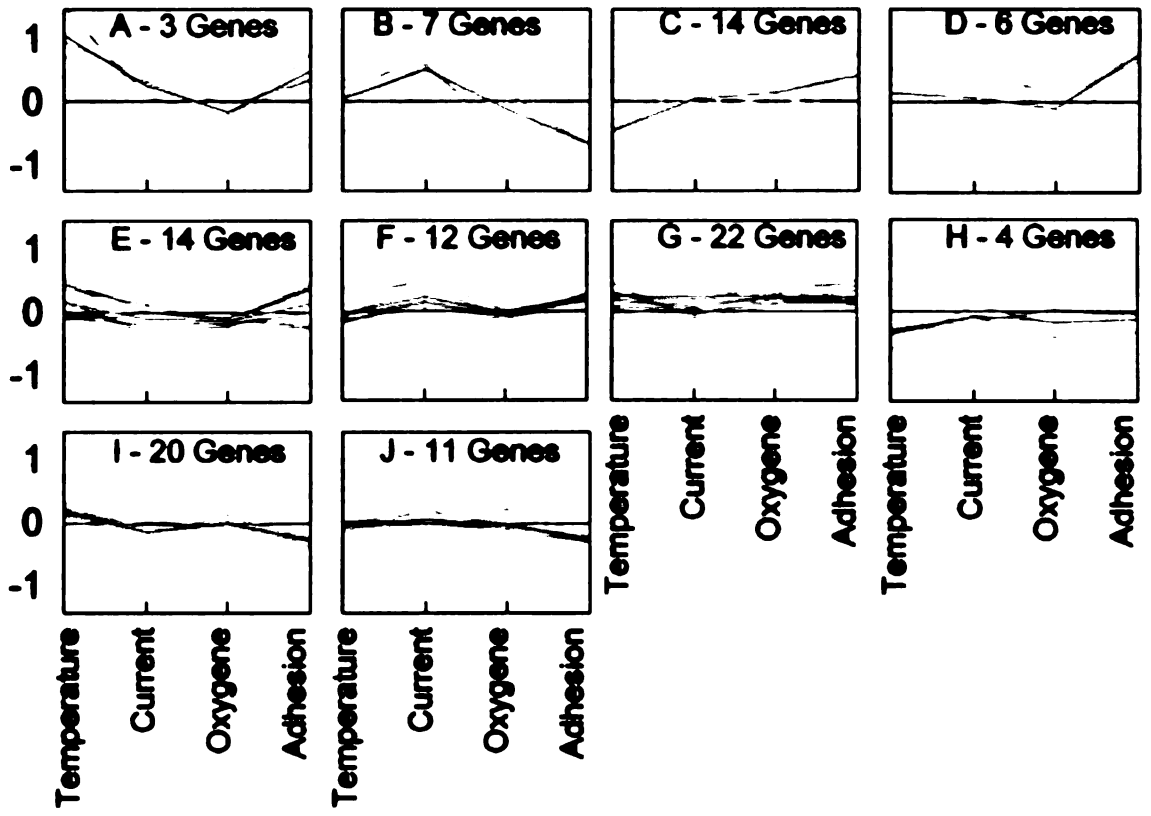


Figure 5-2 The hierarchial cluster of microarray identified genes.
The identified probesets with $P < 0.05$ were hierarchically clustered based on their expression profiles under the effects of the four factors. Ten clusters are generated representing the genes with the expressions induced by the change of the four factors.

This figure is in color.

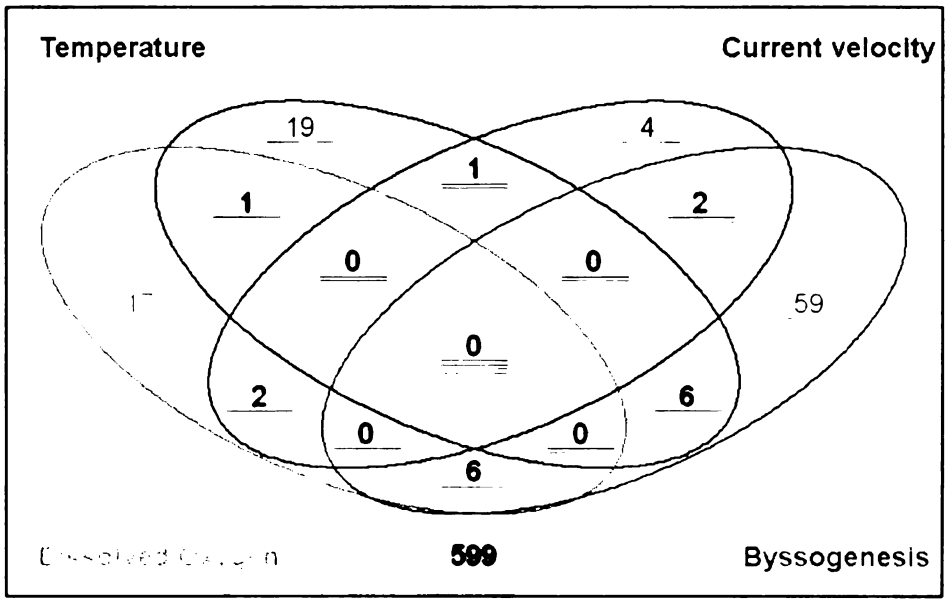


Figure 5-3 The numbers of genes whose expression profiles are altered by single factor or multiple factors.
The values in the ellipse were the numbers of the probes which were altered by the factor with the same color. For example, the sum of the numbers in red ellipse is 19+1+1+6=27 and that means 27 probes were found modulated by the temperature. The underline(s) under each number suggested the amount and types of factors could regulate its expression profiles. For instance, the number 19 with a read underline means the expression profiles of the 19 genes can only be altered by factor A while the value 1 with a red and blue lines suggests that one gene was regulated by both temperature and current velocity.

This figure is in color.

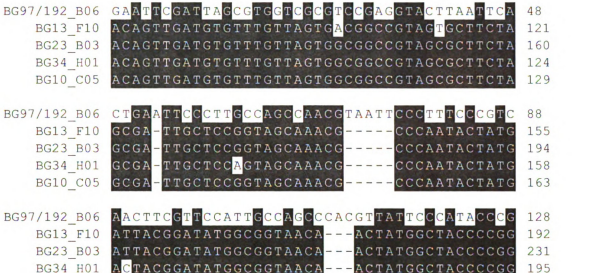


Figure 5-4 Multiple alignment with the identified EGP encoding genes whose expression profiles can be altered by byssogenic activity.

The comparison was performed by using ClustalW multiple sequences alignment. The protein product encoded by the EST BG97/192_B06 is homologous to an exocrine protein isolated from a filarial nematode *L. sigmodontis* while the rest of the EGPs encoded by selected ESTs have a salivary gland peptide from blacklegged tick (*I. scapularis*) as their homologue. The nucleotides were colored with black when more than 60% nucleotides in this locus were identical.

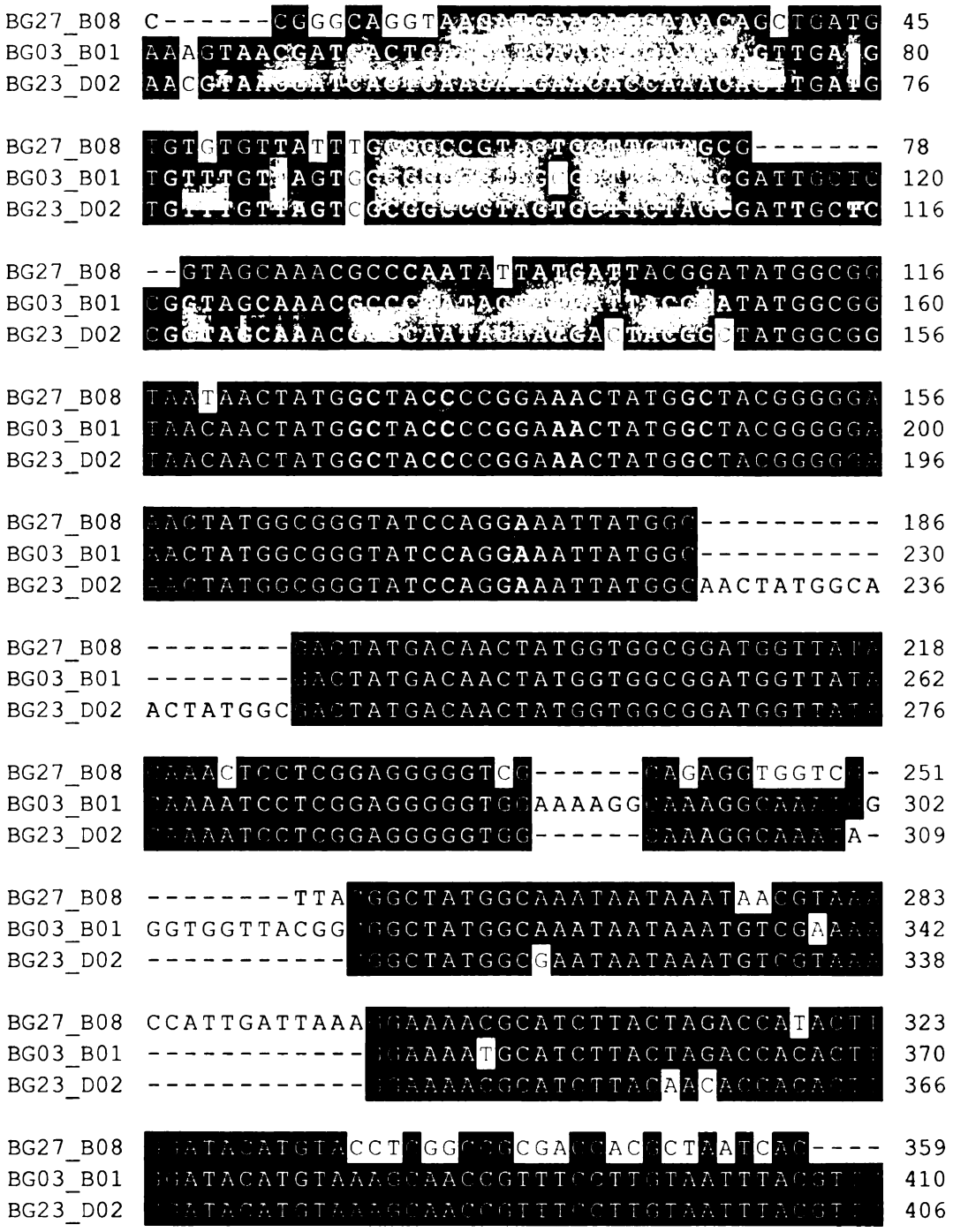


Figure 5-5 Comparison of primeray structures of the EGP encoding genes whose expression profiles can be altered by water temperatures.
The three EGP genes are all differentially expressed under the change of water temperature. The EST BG27_B08 is homologous to salivary gland peptide identified from the western black-legged tick (*I. pacificus*). The other two probesets are all homologous to the blacklegged tick (*I. scapularis*). The nucleotides were colored with black when more than 60% nucleotides in this locus were identical. The gap between the nucleotides were labeled as "-".

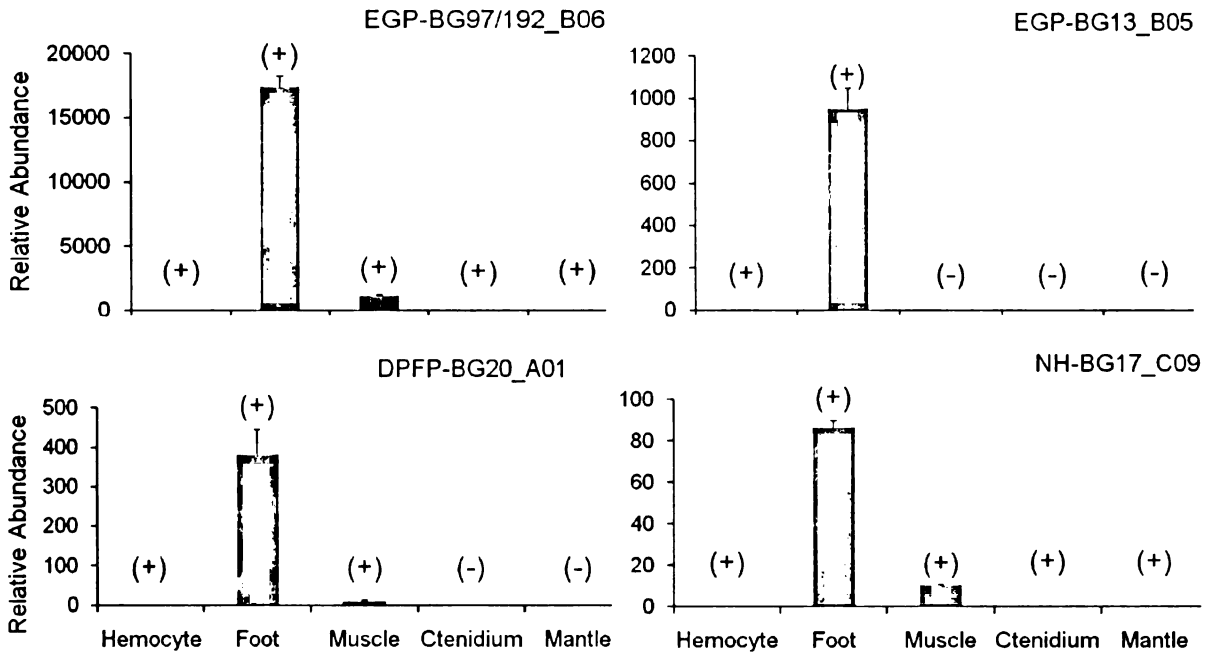


Figure 5-6 The distribution of the mRNA products of the selected genes within zebra mussel tissues.
(+) indicated the existence of the gene in the tissue; (-) suggested no detected transcripts in the tissue. For each gene, the lowest (+) sample was used as control and the relative expression levels of the gene in other tissues were calculated by using $2^{-\Delta\Delta Ct}$ model.

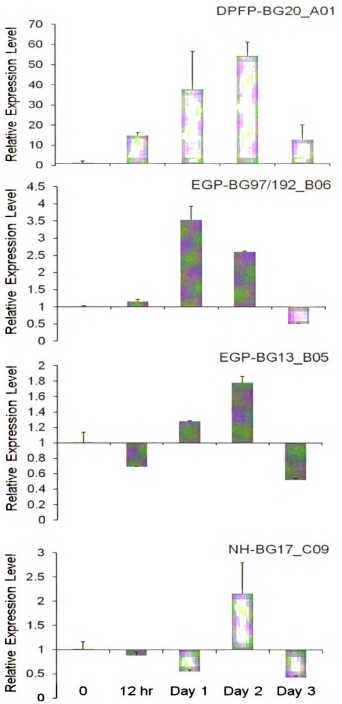


Figure 5-7 The qRT-PCR results demonstrated the relatively expression levels of the gene during the byssiogenesis.

The byssogenesis and non-byssogenesis samples were treated as described in materials and methods, and the non-byssogenesis sample was used as control with value 1. The amount of the transcripts of the gene was calculated by using $2^{-\Delta\Delta Ct}$ model. The comparisons of the gene expression levels between byssogenesis and non-byssogenesis and detached sample were made at 0, 12 hours, 1 day, 2 days, and 3 days post treatment.

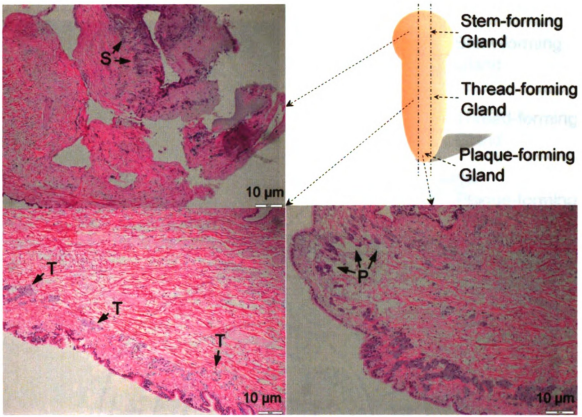


Figure 5-8 The distribution of the zebra mussel byssus gland cells in mussel foot.
Figure 5-8. The distribution of the zebra mussel byssus gland cells in mussel foot. The zebra mussel foot sections are stained by H&E method. The sections are made along the longitude axis as demonstrated by the paralleled dash lines across the foot. Arrows with dashed line indicate the position of the three major byssal glands embeded in the mussel's foot.
a: A longitudinal section in the root of the mussel's foot. Arrows point to the stem-forming glandular cells (S).
b: The section of middle area of the foot. The light purple cells alsong the epethelial surface of the foot were thread-forming cells which were demostrated by arrows and the letter T.
c: The tip of the foot containing the deep purple stained plaque-forming gland cells that were labled as P.

This figure is in color.

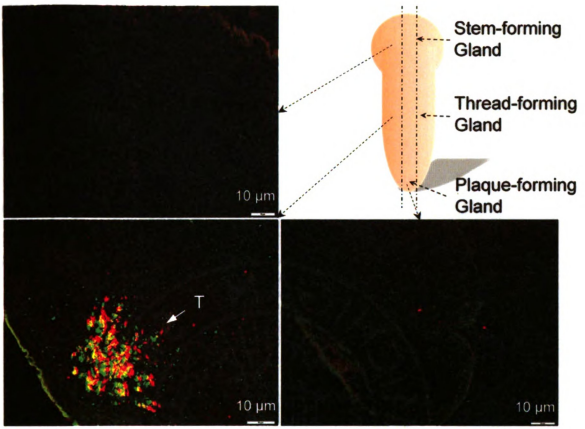


Figure 5-9 The *in situ* expression of DPFP-BG20_A01 and EGP-BG97/192_B06 in foot.
Figure 5-9. The *in situ* expression of the gene DPFP-BG20_A01 and EGP-BG97/192_B06 in zebra mussel foot tissue. The synthesized antisense RNA of DPFP-BG20_A01 was labeled with Alexa 488 dye (green) while the complementary RNA for EGP-BG97/192_B06 was labeled with Alexa 594 (red). The foot was cut along longitude as shown by the two paralleled dash lines across the foot. The positions of the byssus glands were labeled on the foot. The dash lines connected between slides and foot indicated the position of the three slides on the foot.
a: The root area of zebra mussel foot including stem-forming gland;
b: The middle region of the mussel foot containing thread-forming gland;
c: The tip section of the foot including plaque-forming gland.
T: Thread-forming gland cells.

This figure is in color.

Table 5-1 The null hypothesis and the interpretation of tested probesets.

Null hypothesis	**Interpretation when null hypothesis is rejected**
H_0: $A = 0$	The expression profile of certain genes is altered in response to the surrounding temperature.
H_0: $B = 0$	The expression profile of the gene is altered by the change of water movement.
H_0: $C = 0$	The expression profile of the gene is altered by dissolved oxygen level.
H_0: $D = 0$	The expression profile of the gene is altered by byssogenic activity.

Table 5-2 The gene specific primers used for qRT-PCR.

Target EST	Primer Name	Primer Sequence
AM229726	BG20A01_DPFP_RT_Fw	5' -ATG GGC CAT ATG ATA AGA AAC CA- 3'
	BG20A01_DPFP_RT_Rv	5' -TCC AGG AGG TTC CAA TGG AA- 3'
AM229885	BG13B05_EGP_RT_Fw	5' -CGG TTG CTA TAC ATG TGT CCA AGT- 3'
	BG13B05_EGP_RT_Rv	5' -GGG AGG TTA CGG CGG CTA T- 3'
AM230073	BG97/192B06_EGP_RT_Fw	5' -CAT CCC CGT ATG GGA TCC A- 3'
	BG97/192B06_EGP_RT_Rv	5' -GGT GCA ACG GCC AAG TTT AT- 3'
AM230384	BG17C09_NH_RT_Fw	5' -TCC GGA TAT TGG TTG TCC TCA T- 3'
	BG17C09_NH_RT_Rv	5' -TTC TCC GTA GCC ACA CCA TTT- 3'
***18SrRNA**	*18SrRNA_Forward	5'- GAC ACG GCT ACC ACA TCC AA -3'
	*18SrRNA_Reverse	5'- CTC GAA AGA GTC CCG CAT TG -3'

* The primers used as internal reference for QRT-PCR.

Table 5-3 The genes whose expression profiles have been significantly altered by the status of byssogenesis.
When Log(FC) > 0, the gene is upregulated in attachment status.

Gene ID	Accession #	p.value	Log(FC)	Homologue
BG17_G02* **B**	AM230231	0.00011	-0.3	N/A
BG10_C05*	AM229883	0.00114	0.073	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
BG33_H03*	AM230185	0.0013	-0.084	N/A
BG97/192_B06* **C**	AM230076	0.00317	0.107	AAS92593.1| Excretory/secretory protein [*Litomosoides sigmodontis*]
BG20_F04*	AM230089	0.00339	-0.061	N/A
BG13_F10*	AM230013	0.00374	-0.065	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
BG23_B03* **A**	AM229897	0.00498	-0.092	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
BG14_A11*	AM230170	0.00523	0.136	ABI52762.1| 60S ribosomal protein L27 [*Argas monolakensis*]
BG16_D03* **C**	AM230168	0.00801	0.084	AAN05585.1| Ribosomal protein L22 [*Argopecten irradians*]
BG13_C11*	AM230205	0.00946	-0.062	N/A
BG34_H01*	AM229894	0.01029	0.087	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
BG20_A01	AM229724	0.01149	0.185	AF265353_1| Byssal protein Dpfp1 [Dreissena polymorpha]
BG26_C06	AM230431	0.0119	-0.079	N/A
BG05_D06	AM230302	0.01191	0.076	N/A
BG17_C09 **A**	AM230384	0.01206	0.158	N/A
BG08_B10	AM230189	0.01278	0.048	N/A
BG10_F10	AM230328	0.01321	-0.076	N/A
BG32_C03	AM229736	0.01371	-0.141	AAC39039.1| Foot protein 1 precursor [*Dreissena polymorpha*]
BG32_D04	AM230156	0.01374	0.058	BAB12683.1| Polypeptide release factor 3 [*Yarrowia lipolytica*]
BG15_D07	AM230362	0.01379	-0.055	N/A
BG23_E03	AM229779	0.01382	-0.057	BAE93436.1| Shematrin-4 [*Pinctada fucata*]
BG13_B05 **C**	AM229816	0.0139	0.15	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
BG23_C05	AM230080	0.01572	-0.058	N/A
BG14_C07 **C**	AM230353	0.01679	0.096	N/A
BG22_C12	AM229893	0.01706	-0.072	N/A
BG07_F07	AM230102	0.01764	-0.057	NP_504109.1| Neuropeptide-Like protein 29 [*Caenorhabditis elegans*]
BG06_E09	AM230224	0.01777	-0.073	AAF75279.1| Byssal protein Dpfp1 [*Dreissena polymorpha*]
BG04_A09	AM230090	0.02014	-0.054	N/A
BG12_D10	AM230249	0.02035	0.146	N/A
BG31_E11 **C**	AM230254	0.02171	0.147	N/A
BG25_H08	AM230042	0.02211	0.168	N/A

Table 5-3 (cont'd-1)

Gene ID	Accession #	p.value	Log(FC)	Homologue
BG04_G04	AM230066	0.02267	-0.05	ABD62888.1| Serine proteinase-like protein [*Penaeus monodon*]
BG04_F03	AM230188	0.02308	-0.161	N/A
MF030105_B09	AM230050	0.02341	-0.076	N/A
BG28_H05[A]	AM229934	0.02359	0.114	BAE93436.1| Shematrin-4 [*Pinctada fucata*]
BG34_F03	AM230174	0.02459	0.07	ABG81984.1| Ribosomal protein S14e [*Diaphorina citri*]
BG33_E04	AM229997	0.02496	-0.067	N/A
BG18_B07	AM230216	0.0271	0.061	N/A
BG10_D04	AM230104	0.02748	-0.143	AAK68690.1| Hemicentin [*Homo sapiens*]
BG20_B09	AM230000	0.02829	-0.057	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
BG30_H12	AM229764	0.02853	-0.053	N/A
BG25_F07	AM229895	0.0296	0.058	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
MF030105_G10	AM229731	0.02971	0.163	AAF75279.1| Byssal protein Dpfp1 [*Dreissena polymorpha*]
BG08_B08	AM230321	0.0341	0.109	N/A
MF030105_H09	AM229864	0.03422	-0.041	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
BG06_C03	AM229866	0.03508	-0.142	N/A
BG23_B06	AM229738	0.03609	0.046	AF265353_1 Byssal protein Dpfp1 precursor [*Dreissena polymorpha*]
BG05_E08	AM230198	0.03709	0.057	N/A
MF030105_F07	AM230273	0.03711	-0.047	N/A
BG26_B09[A]	AM229855	0.03759	-0.048	N/A
MF030105_C07[B]	AM229749	0.03772	-0.145	N/A
BG34_B04	AM230256	0.03796	0.043	N/A
MF030105_F02	AM230272	0.03899	-0.11	N/A
BG31_A11	AM230107	0.04079	0.047	EAY59350.1| Serine protease pepD [*Mycobacterium tuberculosis* C]
BG03_B01[A]	AM229901	0.04086	0.064	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
BG05_B04[A]	AM230070	0.04134	0.055	N/A
BG10_B09	AM230213	0.0415	0.105	N/A
BG97/192_A06	AM230225	0.04155	0.07	N/A
BG08_F02	AM229725	0.04254	0.052	AAF75279.1| Byssal protein Dpfp1 [*Dreissena polymorpha*]
BG16_A07	AM230120	0.04305	-0.06	ABE03741.1| Prophenoloxidase activating factor [*Penaeus monodon*]
BG26_G12	AM229957	0.04376	-0.049	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
BG04_A11	AM230155	0.04379	0.05	BAB12683.1| Polypeptide release factor 3 [*Yarrowia lipolytica*]
BG29_E08	AM229745	0.04467	-0.062	AAF75279.1|Byssal protein Dpfp1 precursor [*Dreissena polymorpha*]
BG14_D01	AM230121	0.04495	-0.043	N/A
BG13_A02	AM230345	0.04566	0.087	N/A

Table 5-3 (cont'd-2)

Gene ID	Accession #	p.value	Log(FC)	Homologue
BG04_H02	AM229815	0.04732	-0.05	AAT92111.1| Exocrine gland peptide NPL-2 [*Ixodes pacificus*]
BG14_D12	AM229862	0.04786	0.045	BAE93436.1| Shematrin-4 [*Pinctada fucata*]
BG15_F03	AM229727	0.04786	0.048	AAF75279.1|Byssal protein Dpfp1 precursor [*Dreissena polymorpha*]
BG17_F03	AM230143	0.04871	-0.129	N/A
BG10_H04	AM230118	0.04908	0.069	N/A
BG31_D01	AM230039	0.04942	0.133	NP_505834.1| Neuropeptide-Like protein 33 [*Caenorhabditis elegans*]
BG15_D03	AM229780	0.04982	0.097	N/A
BG30_D12 [C]	AM229805	0.04991	-0.045	BAE93436.1| Shematrin-4 [*Pinctada fucata*]

* The differentially expressed ESTs with $P < 0.01$.

[A] Also altered by Factor A (Temperature); [B] Also altered by Factor B (Agitation);

[C] Also altered by Factor C (D.O.)

Table 5-4 The genes whose expression profiles are significantly altered due to the change of temperature.
Log(FC) > 0 means the gene is upregulated with the decrease of temperature.

Genes ID	Accession #	p.value	Log(FC)	Homologues
BG17_C09* [D]	AM230384	7.00E-05	0.314	N/A
BG27_B08*	AM230019	0.00043	0.116	AAT92111.1| Exocrine gland peptide NPL-2 [*Ixodes pacificus*]
BG28_H05* [D]	AM229934	0.0052	-0.161	BAE93436.1| Shematrin-4 [*Pinctada fucata*]
BG03_B01*	AM229901	0.00653	-0.098	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
BG23_D02*	AM229752	0.0068	-0.095	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
BG28_C09	AM229790	0.01034	-0.127	BAE93436.1| shematrin-4[*Pinctada fucata*]
BG17_F07	AM229867	0.01123	-0.057	BAE93436.1| shematrin-4 [*Pinctada fucata*]
BG27_H05	AM230068	0.01216	0.091	AAC05725.1| RNA helicase A [*Mus musculus*]
BG27_E09	AM229903	0.01427	0.1	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
BG26_D08	AM230432	0.01818	0.054	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
BG13_B06	AM229798	0.01919	0.152	AAT92111.1| Exocrine gland peptide NPL-2 [*Ixodes pacificus*]
BG28_B04	AM230435	0.02361	-0.19	N/A
BG29_E12	AM230154	0.02449	-0.065	N/A
BG05_B04	AM230070	0.02598	-0.067	N/A
BG16_F01	AM230081	0.02693	0.096	AAS92593.1| Excretory/secretory protein [*Litomosoides sigmodontis*]
BG23_B03 [D]	AM229897	0.02749	0.077	NP_505834.1|Neuropeptide-Like protein 33 [*Caenorhabditis elegans*]
BG26_C05	AM229917	0.02753	0.08	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
BG18_D01	AM229730	0.02863	0.201	N/A
BG16_G06	AM230258	0.03025	-0.212	N/A
BG05_A11	AM229750	0.03037	0.069	Q25460| Adhesive plaque matrix protein [*Mytilus edulis*]
BG27_B10	AM230153	0.03739	0.065	BAB12683.1| Polypeptide release factor 3 [*Yarrowia lipolytica*]
BG33_A08	AM229799	0.03749	0.093	AAT92111.1| Exocrine gland peptide NPL-2 [*Ixodes pacificus*]
BG97/192_D02	AM229740	0.03991	-0.096	N/A
BG07_H06 [B]	AM230138	0.04305	0.078	ABN13415.1| Choriogenin H [*Oryzias melastigma*]
BG28_E01 [C]	AM229772	0.04418	0.053	N/A
BG28_F03	AM230145	0.04425	0.069	N/A
BG26_B09 [D]	AM229879	0.0449	0.051	N/A

* The differentially expressed ESTs with $P < 0.01$.

[B] Also altered by Factor B (Agitation); [C] Also altered by Factor C (D.O.); [D] Also altered by Factor D (Adhesion).

Table 5-5 The genes whose expression profiles are significantly altered due to the change of D.O.
Log (FC) > 0 indicates the gene is upregulated when the dissolved oxygen in water is at lower level.

Gene ID	Accession #	p.value	Log (FC)	Homologue
BG23_D06*	AM230109	0.0026	-0.135	N/A
BG14_F09*	AM230247	0.00639	0.08	N/A
BG12_B10	AM230219	0.01343	0.098	N/A
BG30_D12 [D]	AM229805	0.0138	0.048	BAE93436.1| Shematrin-4 [*Pinctada fucata*]
MF030105_D04	AM230269	0.01606	-0.125	N/A
BG18_B03	AM230397	0.02033	0.117	N/A
BG16_D03 [D]	AM230168	0.02521	0.056	AAN05585.1| Ribosomal protein L22 [*Argopecten irradians*]
BG05_D02	AM230211	0.02647	0.103	N/A
BG14_C07 [D]	AM230353	0.02743	0.071	N/A
BG97/192_C05	AM230277	0.02815	0.063	N/A
BG13_E01	AM229847	0.02927	0.044	BAE93436.1| Shematrin-4 [*Pinctada fucata*]
BG25_H11 [B]	AM230248	0.02979	-0.047	N/A
BG16_B01	AM229753	0.03044	0.052	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
MF030105_H01	AM229821	0.03044	0.134	AAT92111.1| Exocrine gland peptide NPL-2 [*Ixodes pacificus*]
BG07_F03	AM229726	0.03069	-0.083	AAF75279.1| Byssal protein Dpfp1 precursor [*Dreissena polymorpha*]
BG12_G10	AM230123	0.03203	0.054	N/A
BG97/192_B06 [D]	AM230076	0.03528	0.059	AAS92593.1| Exocrine/secretory protein [*Litomosoides sigmodontis*]
BG27_F10	AM230146	0.04161	0.047	N/A
BG24_B11	AM230213	0.04259	0.041	N/A
BG28_E01 [A]	AM229772	0.04306	0.039	N/A
BG26_B01	AM230172	0.04342	-0.041	AAK95191.1| 40S ribosomal protein S9 [*Ictalurus punctatus*]
BG31_E11 [D]	AM230254	0.04511	-0.103	N/A
BG13_B05 [D]	AM229885	0.04647	0.096	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]
BG22_H06 [B]	AM229789	0.04694	-0.039	N/A
BG20_F01	AM229782	0.04702	-0.107	N/A
BG25_B01	AM229884	0.04732	0.089	AAV80789.1| Exocrine gland peptide [*Ixodes scapularis*]

* The differentially expressed ESTs with P <0.01.

[A] Also altered by Factor A (Temperature); [B] Also altered by Factor B (Agitation);

[D] Also altered by Factor D (Adhesion).

Table 5-6 The genes whose expression profiles are significantly altered due to the current velocity.
Log(FC) > 0 indicates the gene is upregulated when the mussel is in flow water.

Gene ID	Accession #	p.value	Log (FC)	Homologue
BG29_B09*	AM230251	0.00415	0.116	N/A
BG22_H06*[C]	AM229789	0.00952	-0.078	N/A
BG10_F01	AM230379	0.01195	0.074	N/A
BG29_A05	AM230250	0.02185	0.053	N/A
MF030105_C07[D]	AM229749	0.02546	-0.188	N/A
BG07_H06[A]	AM230138	0.0265	0.094	ABN13415.1| Choriogenin H [*Oryzias melastigma*]
BG17_G02[D]	AM230231	0.03856	0.166	N/A
BG25_H11[C]	AM230248	0.03927	-0.065	N/A
BG12_H05	AM230222	0.0491	0.076	N/A

* The differentially expressed ESTs with P <0.01.
[A] Also altered by Factor A (Temperature); [C] Also altered by Factor C (D.O.);
[D] Also altered by Factor D (Adhesion)

CHAPTER VI

Defensin of the Zebra Mussel (Dreissena polymorpha): Molecular Structure, in vitro Expression, Antimicrobial Activity, and Potential Functions

Abstract

A 409 bp full length defensin cDNA was cloned and sequenced based on an expressed sequence tag (EST) obtained from a normalized zebra mussel (*Dreissena polymorpha*) foot cDNA library developed in our laboratory. The *D. polymorpha* defensin (Dpd) gene encoded a peptide with 76 amino acid residues. The mature Dpd contains 54 amino acids with a fully functional insect defensin A domain. Homologue searching against GenBank database suggested that this Dpd was phylogenetically close to defensins from a group of insects with six conserved Cysteine residues. Predicted with homology modeling method, the three dimensional structure of Dpd also demonstrated a significant similarity with insect defensin A. The recombinant Dpd was *in vitro* expressed through an *Escherichia coli* expression system. The antimicrobial activities of the re-folded recombinant Dpd were found against the growth of *Morganella* sp., *Plesiomonas shigelloides*, *Edwardsiella tarda*, *Escherichia coli*, and *Staphylococcus aureus aureus* (ATCC12598) with the minimal inhibition concentration (MIC) 0.35, 0.43, 1.16, 6.46, and 30.39 μM, respectively. However, with less than 50 μM no detectable inhibition activities were observed against the other four Gram-negative bacteria (*Aeromonas salmonicida salmonicida*, *Motile Aeromonas*, *Flavobacterium* sp., *Pseudomonas fluorescens*,

Shewanella putrifaciens), as well as the other Gram-positive bacteria (*Bacillus megaterium* ATCC14581, *Carnobacterium maltaromaticum* ATCC27865, *Enterococcus faecalis* ATCC19433, and *Micrococcus leteus* ATCC4698), and the yeast *Saccharomyces cerevisiae*. The RNA fluorescent in situ hybridization (FISH) of Dpd in the zebra mussel suggested that the Dpd was expressed in a variety of tissues, such as foot, retractor muscle, ctenidium, mantle, hemocytes, gonad, digestive gland, and intestine. By using the quantitative PCR, the expression level of Dpd in the zebra mussel foot was the highest, followed by the muscle. By comparing the amount of Dpd transcripts in the zebra mussel foot under two different byssogenesis conditions (byssogenesis and non-byssogenesis) with qPCR, we found that the higher expression levels of Dpd were always associated with the byssogenesis status from two days to four days after the start of byssogenesis. Findings of this study suggest that the expression of the Dpd gene was upregulated in vivo by byssogeneis, and by hemocytes *in vitro* upon stimulation with microbial antigens.

Introduction

The invasion of the zebra mussel (*Dreissena polymorpha*) in North America has raised several questions pertaining to the ability of a non-native mollusk to adapt and reproduce in a new environment. While the lack of natural predators and ability to attach firmly to underwater substrates are logical explanations, scientists believe that zebra mussels may have extraordinary host defense mechanisms that allow them to survive microbial attacks in the pathogen-rich aquatic environment (Frischer *et al.* 2000). While the zebra mussels could be experimentally infected with pathogens such as *Aeromonas* spp. (Maki *et al.* 1998) and *Pseudomonas fluorescens* (Molloy 2002), no spontaneous infections by either pathogens were observed since the zebra mussel was first introduced to North America in the 1980s (Hebert *et al.* 1989).

Like other invertebrates, bivalves, to which zebra mussels belong, have both cellular and humoral innate host defense mechanisms (Cheng 1983; Hine 1999). The former includes two main pathways, phagocytosis or encapsulation, followed by the destruction of pathogens via enzymes or the release of oxygen metabolites (Pipe 1990, 1992), while the latter contains two families of molecules, opsonins (Renwrantz & Stahmer 1983) and antimicrobial peptides (AMPs) (Hubert *et al.* 1996; Mitta *et al.* 2000a). Most of the knowledge about bivalve innate immunity comes from marine mussels or oysters, such as *Mytilus edulis* (Pipe *et al.* 1997), *M. galloprovincialis* (Carballal *et al.* 1997), *Crassostrea virginica*, and *C. gigas* (Oliver & Fisher 1995). The studies on the host defense mechanism of the zebra mussel are very limited (Giamberini *et al.* 1996).

Previous study of zebra mussel hemocyte-mediated host defense mechanisms demonstrated that exposure of hemocytes to microbial-derived proteins causes a sharp rise in the expression of genes with potential defense and cell protection functions such as matrilin and heat shock proteins (Xu & Faisal 2007, 2009b). Interestingly, when a cDNA library was constructed with zebra mussel foot and normalized by suppression subtractive hybridization technique, 45 immune-related genes were identified (Xu & Faisal 2008). Considering that the foot cDNA library was developed through subtraction of the mussels' retractor muscle, which contains a sinus full of hemocytes, the immune-related genes found in the subtracted foot cDNA were either overly expressed in foot cells as compared to hemocytes or unique to the foot. Located in the zebra mussel foot, there are three major glands that produce foot proteins, which build unique structures called byssal threads to maintain underwater attachment. The generation of byssal threads, also known as byssogenesis, is affected by many environmental factors such as temperature (Clarke & McMahon 1996c), current velocity (Clarke & McMahon 1996a), dissolved oxygen (Clarke & McMahon 1996b), and microbial communities (Kavouras & Maki 2003b; Kavouras & Maki 2003a; Kavouras & Maki 2004). In particular, the interference of microbial communities during the zebra mussel byssogenesis has been observed in both the larval and adult stages of zebra mussels. The host defense related genes identified from our previous study are probably involved in the self protection employed by zebra mussel byssus. Among these, an AMP molecule with high similarity to defensin was found in the zebra mussel foot SSH cDNA library.

AMPs play important roles in the innate immune system of a variety of organisms (Boman 2003; Bulet *et al.* 2004). Over 1,000 antimicrobial peptides have been reported

(Bulet *et al.* 2004), including three basic types based on their secondary structure of sequences, linear peptides with α-helices, highly disulphide-bonded (cysteine-rich) β-sheets, and those with proline- or glycine-rich character (Douglas *et al.* 2003; Gueguen *et al.* 2006). The amphipathic character of most members in the AMP family enables them to be inserted into biological membranes. Although, in most cases, the primary action of AMP is to lyse the membranes of pathogens, they are also involved in a number of other reactions depending on their tissue distributions, such as mediating inflammation (Hancock & Lehrer 1998).

Among AMPs, the defensin represents a family of peptides with the antimicrobial activities against a wide range of bacteria and fungi (Boulanger *et al.* 2004). Structurally, the animal defensin molecules contain three or four disulfide bonds and can be classified into four major groups: α-, β-, θ-, and invertebrate defensin. In many vertebrates, the α-defensin is constitutively expressed in neutrophils and acts intracellularly within the phagolysozom (Cunliffe 2003). The β-denfensin is inducible by a variety of pathogens in epithelial cells and neutrophils, and also functions to initiate the adaptive immune system of vertebrates (Yang *et al.* 1999). The θ-defensin, expressed in the neutrophils and monocytes of non-human primates, was most recently was found to have lectin-like activities that can mediate the protection of lymphocytes from infection by viruses (Cole *et al.* 2002; Munk *et al.* 2003; Wang *et al.* 2003).

The first insect defensin, first purified from the larvae of the flesh fly *Phormia terranova* (Hoffmann & Hetru 1992), has been isolated from a broad range of insect species which includes six insect orders (Cociancich *et al.* 1993). As the most diverse

animal group, the insects have the most remarkable diversity not only reflected by their morphology and lifecycles but also by various biomolecules including defensins.

A common insect defensin contains six cysteine residues engaged in three disulfide bonds. Three distinct domains can be found in a mature insect defensin including an amino-terminal loop, an amphipathic α-helix, and a carboxyl-terminal anti-parallel β-sheet (Bonmatin *et al.* 1992). Like the vertebrate defensins, the insect defensins also inhibit the growth of a wide range of microorganisms (Hoffmann & Hetru 1992). Interestingly, the defensins identified from mollusks were similar to insect defensins. Thus far, two defensins, *Cg-def* and MGD-1, have been isolated from *Crassostrea gigas* (Gueguen *et al.* 2006) and *Mytilus galloprovincialis* (Li *et al.* 2000), respectively. Antimicrobial assays demonstrated that insect and molluscan defensins share functional similarities. To this end, this study was designed to identify the structure of zebra mussel defensin, and to determine its tissue distribution, inducibility, antimicrobial activities, and potential involvement in byssogenesis.

Materials and Methods

The collection and maintenance of zebra mussels

The zebra mussels used in all the assays of this study were collected from Vineyard Lake in Brooklyn, MI, USA (Latitude: 42°4'59''N; Longitude: 84°12'34"W). The mussel shells were thoroughly cleaned with a brush and rinsed with deionized water. Then the mussels were maintained in aerated, sterile Vineyard Lake water in a glass tank and fed weekly with a pure culture of the algus *Ankistrodesmus falcatus*. The zebra mussels used in this study were allowed to acclimate to the laboratory environment for six weeks before use.

Gene cloning of zebra mussel defensin

The expressed sequence taq (EST) of Dpd was obtained from a normalized cDNA library constructed with zebra mussel foot RNA (Xu & Faisal 2008). A pair of gene specific primers (BG11E08-RT- Fw: 5' - CGA AGG CGG ATA TTG CTA CTG - 3' and BG11E08-RT- Rv: 5' - TGA CCA TAA CAC TCC ATC GGT G - 3') was designed based on the Dpd EST with PimerExpress 2.0 (Applied Biosystem, Foster City, CA). The full length cDNA library was constructed with SMART™ RACE cDNA Amplification Kit (Clontech Laboratories, Mountain View, CA). The 3' and 5' ends of the Dpd gene were amplified following the instructions of the kit. Purified PCR products were sequenced at the Research Technology Support Facility of Michigan State University. The full sequence of Dpd was deposited in GenBank with the accession number GU139954.

Primary and tertiary structure analyses of zebra mussel defensin

The homologue searching of the full length Dpd gene was performed with the Basic Local Alignment Search Tool (BLAST) through NCBI server. The Dpd gene was translated into amino acid sequence by the DNA-Protein translation tool provided by Expert Protein Analysis System (ExPASy) (Gasteiger *et al.* 2003). The primary structure of Dpd protein was compared to defensin molecules from other species by using the multiple sequence alignment program CLUSTAL W (Thompson *et al.* 1994).

The tertiary structure of zebra mussel defensin was predicted with a homology-modeling method via ESyPred3D using neural networks (Lambert *et al.* 2002). The visualization and characterization of the three dimensional structure of the zebra mussel defensin was performed with software PyMOL (DeLano 2009). The tertiary structure of zebra mussel defensin was compared to its homologue, the insect defensin A (PDB entry: 1ICA) isolated from *Phormia terraenovae* (Cornet *et al.* 1995), as well as two defensin molecules, MGD-1 (1FJN) of *M. galloprovincialis* (Li *et al.* 2000) and *Cg-def* (2B68) of *C. gigas* (Gueguen *et al.* 2006).

The expression of zebra mussel defensin in *E. coli in vitro* expression system

The recombinant zebra mussel defensin was produced *in vitro* with prokaryote expression system, QIAexpression® System (Qiagen, Valencia, CA). The total RNA was extracted from the zebra mussel foot followed by cDNA synthesis with SuperScript® III Reverse Transcriptase (Invitrogen). A pair of primers (Forward: 5' - GCG CGC ATG CGG CAC CCC AGA AGC GTA TTA C - 3'; Reverse: 5' - ATA TAT GTC GAC TTA ACC AAG GAT TTC CGA GAA GG - 3') was designed to amplify the whole functional domain of the zebra mussel defensin as well as to introduce the recognition site of

restricted endonuclease SphI and SalI to the 5' and 3' ends of the sequence, respectively. The PCR product was obtained through RT-PCR with zebra mussel cDNA template and was purified with Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI). The purified PCR product was sequentially digested with endonuclease SphI and SalI (New England Biolabs, Ipswich, MA). Meanwhile, the expression vector PQE-1 from the QIAexpression® System (Qiagen) was also digested by SphI and SalI in sequence. The digested PCR product and PQE-1 vector were purified, and then connected to each other with T4 DNA ligase from pGEM®-T Easy Vector System I (Promega). Thereafter, the recombinant plasmid was transformed to competent M15 *E. coli* cells. The preparation of competent cells, transformation, and the bacteria culture were based on the protocol in the manufacturer's instruction of the QIAexpressionist™ (Qiagen). The Isopropyl β-D-1-thiogalactopyranoside (IPTG) with a final concentration 1mM was used to induce the expression of the defensin gene in *E. coli* cells. The matured zebra mussel defensin with poly-histidines at its carboxyl-terminus (C-terminus) were abundantly produced after six-hour of IPTG induction. The 6×His-tagged defensin was then purified with the Ni-NTA Agarose filled polypropylene column (Qiagen) under the naïve condition. The detailed purification steps were also available in the manual of QIAexpressionist™ (Qiagen). The purified 6×His-tagged defensin was overnight desalted with Slide-A-Lyzer® Dialysis Cassette (Thermo Scientific, Rockford, IL) immersed in 1× TAGZyme™ Buffer (20 mM NaH_2PO_4, 0.15 M NaCl, pH 7.0) at 4 °C. The 6×His-tag at the C-terminus was then cleaved by DAPase™ from TAGZyme™ Kit (Qiagen). The cleavage and enzyme removal were described in the TAGZyme™ Handbook (Qiagen). The cleaved product

was detected by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) that was prepared according to the protocol written by Walker (Walker 2002).

The antimicrobial activity of recombinant Dpd

To have the recombinant defensin recovering the maximum activities, the mature defensin was refolded under laboratory conditions (Gueguen *et al.* 2006). First of all, the zebra mussel defensin dissolved in 1× TAGZyme™ buffer was overnight dialyzed at 4 °C with refolding buffer containing 100 mM Tris-HCl at pH 8.0. Then the cleaved defensin protein was reduced in the refolding buffer in the presence of 100 mM dithiothreitol (DTT). Finally, the defensin solution was dialyzed again at 4 °C with pure refolding buffer overnight. The defensin in refolding buffer was allowed to refold at room temperature for 48 hours. The concentration of the refolded protein was detected by Qubit™ fluorometer with the Quant-iT™ Protein Assay Kit (Invitrogen). Then a series of dilutions were made to create a set of Dpd solutions with different concentrations. The minimal inhibition concentrations (MICs) of the Dpd were tested with radial diffusion assay following the steps described by Lehrer *et al* (Lehrer *et al.* 1991). The microorganisms we used included nine Gram-negative bacteria (*Aeromonas salmonicida salmonicida*, *Motile Aeromonas*, *Morganella* sp., *Edwardsiella tarda*, *E. coli* DH5α, *Flavobacterium* sp., *Plesiomonas shigelloides*, *Pseudomonas fluorescens*, and *Shewanella putrifaciens*), four Gram-positive bacteria (*Bacillus megaterium* ATCC14581, *Carnobacterium maltaromaticum* ATCC27865, *Enterococcus faecalis* ATCC19433, *Micrococcus leteus* ATCC4698, and *Staphylococcus aureus aureus* ATCC12598), and one yeast (*Saccharomyces cerevisiae*). Except for *E. coli* DH5α and the Gram-positive bacteria, all the other microorganisms were isolated from fish disease cases by the

Aquatic Animal Health Laboratory at Michigan State University and identified by biochemical and/or molecular testing. The bacteria and yeast isolates were inoculated and incubated in 3% trypticase soy broth (TSB) at room temperature overnight. One milliliter aliquot of each culture was transferred into a 1.5 ml microcentrifuge tube. The microorganisms were centrifuged at 900 × g for 10 minutes at 4 °C. After the supernatant was discarded, the pellet of each microorganism was re-suspended with 1 ml ice cold phosphate buffered saline (PBS, pH 7.4). The optical density of the suspension was measured at 620 nm. The 1 ml suspended solution of each microorganism (4×10^6 CFU) was mixed with 20 ml previously autoclaved warm (42°C) PBT (PBS with 0.05% Tween 20, pH7.4) containing 3 mg TSB powder and 1% low-electroendosmosis-type (LE) agarose (Applied Biosystems /Ambion, Austin, TX). The mixture was poured into a 150 × 15 mm Petri-dish after the microorganism was well dispersed. After the medium was solidified, a ~10 mm deep layer was formed at the bottom of the Petri-dish. A sterile disposable transfer pipet (VWR International, West Chester, PA) was then used to make wells with ~3 mm diameter evenly spaced on the surface of the medium plate. Five microliters of each Dpd dilution was added in each well and incubated at room temperature for at least three hours. On the top of each plate, sterilized overlay medium containing 6% TSB and 1% LE agarose was poured at 42 °C. The double layer medium plates were kept in room temperature overnight to 24 hours.

After the inhibition rings were visualized on each plate, the diameter of the ring was measured with 10 × Hastings Measuring Magnifier (Bausch & Lomb Inc., Rochester, NY). The diameters of the inhibition rings (y-axis) were plotted against the $\log_{10}$ values of peptide concentrations (x-axis). The linear regression of the diameter / concentration

was estimated and the x-axis intercept of the regression line corresponds to log (MICs) (Ciornei *et al.* 2005).

Determination of Dpd expression in zebra mussel tissues

The RNA fluorescent *in situ* hybridization (FISH) was performed to identify the expression levels of defensin in different zebra mussel tissues. The zebra mussel was fixed in Davidson's fixative solution. One liter solution includes 115 ml glacial acetic acid, 330 ml 95% ethanol, 220 ml 100% formalin and 335 ml double distilled water with pH 3-4 (Andrade *et al.* 2008). The mussel was conserved in the fixative solution for 24 hours until the mussel shell is dissolved. The fixed mussel was buried in paraffin and sectioned along the longitude axis of the mussel. The 5 μm thick sectioned sample was put on an RNase- and DNase- free glass slide. The probes used for the FISH were RNA probes which were complementary strands of the Dpd mRNA. The PCR product (amplified with primer BG11E08-RT-Fw and BG11E08-RT-Rv) of the gene fragment was cloned into pGEM®-T Easy Vector Systems (Promega U.S., Madison, WI). Then the positive colony was picked for sequencing with M13 forward primer to find the direction of the insert of the recombinant plasmid. The PCR was applied on the selected colonies with both M13 forward and reverse primers. The PCR products were purified with Wizard SV Gel and PCR Clean-Up Sysyem (Promega) and used as the templates of RNA transcription. The T7 RNA polymerase was selected based on the insert direction of the recombinant plasmid to synthesize the antisense RNA. An RNA probe of zebra mussel cytoplasmic actin (AF082863) gene was also produced in the same way as a reference. Meanwhile, the sense-strand of Dpd probe was also synthesized as negative control (NC).

The probe for Dpd (Dpd-R) was labeled with fluorescent dye Alexa 594 (Invitrogen) while the probe for actin (Act-G) was labeled with Alexa 488 (Invitrogen). The steps of the RNA synthesis and labeling can be found in the manual of either FISH Tag™ RNA Red Kit (Invitrogen) or Fish Tag™ RNA Green Kit (Invitrogen). The probe Dpd-R and ACT-G were used to hybridize with the same zebra mussel whole body section slide. The RNA FISH was performed following the recommended protocol in the appendix of the Fish Tag™ RNA Kit instruction. The hybridized slides were visualized by Olympus BX41 (Olympus America Inc., Center Valley, PA) microscope under the excitation of mercury lamp and the images were captured by Olympus DP25 digital camera (Olympus America Inc.). The image processing software DP2-BSW (Olympus America Inc.) was used to combine the different color channels in a single image. To enhance the contrast of signals given by Dpd and actin, the blue color instead of green was assigned to actin by the software. Similarly, the probe NC was also labled with Alexa 594 and hybridized to a slide with Alexa 488 labeled actin as background.

The quantitative RT-PCR (qRT-PCR) was used to detect the relative abundance of defensin mRNA within the different tissues of zebra mussels. The shells of the zebra mussels used in this assay were swabbed with 70% ethanol, followed by a soak in 150 ml sterilized double distilled water containing 5,000 U penicillin, 5 mg streptomycin, and 10 mg gentamicin for thirty minutes (Davids & Yoshino 1998). Thereafter, the mussels were transferred to sterilized double distilled water. The RNA samples were extracted from hemocytes, feet, muscles, ctenidia, and mantles of zebra mussels. The hemolymph was extracted from the adductor muscle of 10 zebra mussels as and mixed with equal volume of a buffer that consisted of 0.05 M TRIS/HCl, 2% glucose, 2% NaCl (pH 7.4), as

described by Pipe *et al* (Pipe *et al.* 1997). The mixture was centrifuged at 900 g for 10 minutes at 4 °C. The other tissues were collected under a dissecting microscope with sterile scalpels and scissors. The total RNA of each sample was extracted with 5-PRIME PerfectPure RNA Cell & Tissue Kit (5 PRIME Inc, Gaithersburg, MD). The relative expression level of zebra mussel defensin in each sample was estimated by a one-step qRT-PCR performed with Power SYBR® Green RNA-to-CT^{TM} 1-Step Kit (Applied Biosystems). The paired primers designed for full length Dpd cloning (BG11E08-RT-Fw and BG11E08-RT-Rv) were used in qRT-PCR. The primers for internal reference of the qRT-PCR were also designed by PrimerExpress^{TM} based on the sequence of zebra mussel 18S ribosomal RNA (18S-Fw: 5'- GAC ACG GCT ACC ACA TCC AA -3' and 18S-Rv: 5'- CTC GAA AGA GTC CCG CAT TG -3'). The thermocycler used for qRT-PCR was a Mastercycler ep realplex S (Eppendorf, Westbury, NY). The thermocycler program used for this one-step RT-PCR was 30 minutes at 48 °C, 10 minutes at 95 °C and 40 repeats of a 2-step temperature cycle (15 seconds at 95 °C and 1 minute at 60 °C). The calculation of gene expression level was accomplished with $2^{-\Delta\Delta Ct}$ algorithm using cycle threshold (Ct) of PCR described by Livak and Schmittgen (Livak & Schmittgen 2001). This experiment was repeated once to have two biological replicates.

The expression profiles of Dpd in the different statuses of byssogenesis

In this assay, the zebra mussels were divided into two different groups. The first group contained mussels fully attached underwater to Petri-dishes for at least 30 days. In the second group, the mussels were detached by removal of the byssal threads with scalpels at their stems; thereafter, they were put back into water. The mussels were

randomly taken from both groups at day 1, day 2, day 3 and day 4 after the detachment of the mussels in the second group. The total RNA samples were extracted from the feet of the mussels collected every day. The qRT-PCR was performed with the RNA samples from both groups. Comparisons were made between the two groups at the same day or between two time points in the same group. The steps of qRT-PCR were previously mentioned in the sections above and a two-way ANOVA test has been performed to analyze the effects of the byssogenesis status and time to the expression level of Dpd.

Results

The full length gene sequence and primary protein structure of Dpd

The full sequence of the zebra mussel Dpd gene obtained from our study consisted of 409 bp, including an open reading frame which encoded Dpd protein with 76 amino acid residues. The BLAST result showed that Dpd was homologous to a number of defensins identified from a variety of species in non-redundant protein sequence database (nr), such as the Japanese disc abalone (*Haliotis discus hannai*, ABF69125), the parasitic wasp (*Nasonia vitripennis*, NP_001159944), the bumblebee (*Bombus ignitus*, AAQ90412), the western honey bee (*Apis mellifera*, NP_001011616), and the scarab beetle (*Anomala cuprea*, BAD77966), etc. Phylogenetically, the Dpd is closest to the defensin of Japanese disc abalone. The defensins of these two molluscan species were closely related to two types of defensins identified from Apocrita (wasps, ants and bees) (Figure 6-1). Aligned with another 15 defensin molecules from 14 species, the Dpd was found to have a conserved motif of common defensins, six conserved cysteine residues. Two continuous glycines in all the defensin molecules were also observed (Figure 6-2).

The three dimensional structure of Dpd

The tertiary structure of Dpd was predicted to be the closest to an insect defensin A model (1ICA) in protein database. Both of these two molecules share a number of structural similarities. First, both had one α-helix and two β-sheets. Second, both have six cysteine residues in their amino acid sequences with three disulfide bonds. Last, the disulfide bonds in Dpd (Cys^{28}-Cys^{57}, Cys^{43}-Cys^{62}, and Cys^{47}-Cys^{54}) and in insect

defensin A (Cys^{57}-Cys^{84}, Cys^{70}-Cys^{90}, Cys^{74}-Cys^{82}) existed in similar positions (Figure 6-3). On the other hand, there was some minor difference between these two defensin molecules. The α-helix and β-sheets in Dpd are shorter than the corresponding secondary structures in the insect defensin A. The α-helix in Dpd consists of seven amino acid residues, while in the insect defensin A, there are nine amino acid residues. Similarly, both β-sheets in Dpd had four amino acid residues, which was one residue shorter than those in the insect defensin A. Moreover, the 3D structure of the insect defensin A was more compactable than that of Dpd (Figure 6-3). The other two molluscan defensin molecules, MGD-1 and *Cg-def*, were structurally close to each other. Both of them have four disulfide bonds instead of three: Cys^{4}-Cys^{25}, Cys^{10}-Cys^{33}, Cys^{14}-Cys^{35}, and Cys^{21}-Cys^{38} in MGD-1 (Figure 6-3D); Cys^{4}-Cys^{25}, Cys^{11}-Cys^{34}, Cys^{15}-Cys^{36}, and Cys^{20}-Cys^{39} in *Cg-def* (Figure 6-3D).

Besides the conserved cysteine residues in both defensin molecules, there were also some other amino acid residues with hydrophobic or positively-charged characters conservatively distributed. In Dpd, His^{46} and Arg^{65} were positively charged when at pH 7.0 (Figure 6-4A) while the His^{73} and Arg^{93} were also observed in insect defensin A (Figure 6-4B). They were all considered as conserved amino acid residues based on their primary protein structures (Figure 6-4C). Similarly, there were five conserved hydrophobic amino acid residues found in both Dpd (Leu^{30}, Leu^{31}, Ile^{36}, Ala^{42}, and Ala^{44}) and insect defensin A (Leu^{59}, Leu^{60}, Ile^{65}, Ala^{69}, and Ala^{71}).

The minimal inhibition concentrations (MICs) of recombinant Dpd

The *in vitro* expressed recombinant Dpd contained the complete functional domain including 58 amino acid residues (from Ala^{22} to Gly^{75}). Among all the microorganisms used in MICs test, the growth inhibition was observed in four Gram-negative bacteria strains. The lowest MIC was found in the inhibition test against *Morganella* sp. (0.35 μM) while the *P. shigelloides* showed the second highest sensitivity to the Dpd with the MIC equal to 0.43 μM. *E. tarda* and *E. coli* showed stronger resistance to Dpd with the MICs 1.16 μM and 6.46 μM, respectively. Only one Gram-postive bacterium strain *Staphylococcus aureus aureus* was inhibited by Dpd with MIC 30.39 μM. With the concentration of Dpd less than 50 μM, there was no obvious inhibition observed in the growth of the other bacteria, nor in the yeast *S. cerevisiae.*

The *in situ* expression of Dpd in zebra mussel

The fluorescent *in situ* hybridization with Dpd RNA probe demonstrated the constitutive expression of Dpd in multiple zebra mussel tissues. The Dpd was mildly expressed by the epithelial cells covering the surface of the mussel's foot (Figure 6-5A) while it was abundantly expressed in a number of granular cells in the middle section of the foot (Figure 6-5B). A strong positive signal was noted at the position of thread-forming glands in zebra mussel foot (Figure 6-5C). The negative control of the FISH in zebra mussel foot was shown in Figure 6-6 where only the actin probe was observed (blue). The expression level of Dpd in the retractor muscle was also present, albeit the signal was weaker than in the foot (Figure 6-7A). Hemocytes that were distributed within the muscle sinus exhibited a relatively strong signal (Figure 6-7B). In the section of zebra mussel ctenidium (gill), the signal of Dpd mRNA was detected in the epithelial lamellae

but not in the water channel or connected tissue between the two epithelial lamellae (Figure 6-7C). The expression of the Dpd was also found in zebra mussel mantle, exclusively in the mantle epithelium (Figure 6-7D). In general, all hemocytes, regardless of their location, exhibited strong signals (Figure 6-7E). Germ cells in the gonads exhibited strong Dpd signals as well (Figure 6-7F). The significant level of Dpd expressions was also observed in the digestive gland cells (Figure 6-7G) and in the intestinal mucosal lining (Figure 6-7H). The negative control for the FISH in the zebra mussel tissues other than foot was displayed in Figure 6-8.

Determination of relative expression level of Dpd in zebra mussel tissues by qPCR

The expression levels of Dpd within zebra mussel hemocytes, foot, retractor muscle, ctenidium, and mantle were different, and the order from highest to lowest was foot, retractor muscle, ctenidium, mantle, and hemocytes (Figure 6-9), sequentially. Out of the five tissues tested in this assay, the zebra mussel foot had the most abundant Dpd expression compared to the other four tissues ($P < 0.01$). The muscle tissue had the second highest expression level of Dpd. The statistical analysis indicated that the expression level of Dpd in muscle is dramatically higher than in hemocyte, ctenidium, and mantle, but significantly lower than in the foot ($P < 0.01$). The Dpd was also expressed in the other tissues tested, hemocyte, ctenidium, and mantle; however, the expression levels were fairly low compared to that in foot or muscle (Figure 6-9). The comparisons were also made between any two of the rest of the tissues and it turned out that the expression levels of Dpd in hemocyte, ctenidium, and mantle of zebra mussel were not significantly different ($P > 0.05$)

The expression of Dpd in the zebra mussel foot during the early stages of byssogenesis

To understand how the expression of Dpd is associated with the zebra mussel byssogenesis, quantitative PCR was used to compare the expression levels in two different byssogenesis status, non-byssogenesis and byssogenesis status. After one-day adhesion, the zebra mussel did not show significant difference of Dpd expression in the feet of non-byssogenesis and byssogenesis individuals. Starting from Day 2 post initiation of byssogenesis, the expression levels of Dpd in the individuals with byssogenesis were significantly higher than the non-byssogenesis ($P < 0.05$ at Day2, $P < 0.01$ at Day 3, 4, and 5). It suggested that the higher expression levels were intended to associate with the mussels under the byssogenesis status (Figure 6-10).

Discussion

Compared to the studies on oyster defensin (*Cg-def*) (Gueguen *et al.* 2006), the Dpd seems to have a broader range of tissue distribution in the zebra mussel. The expression level of *Cg-def* was found significantly higher in the mantle than any of the other tissues of *C. gagas* including hemocytes, adductor muscle, digestive gland, gills, heart, and labial palps. The FISH analysis in our study suggested that the Dpd were constitutively expressed in all the tissues or cells examined, especially the epithelial lining, such as ctenidium (gill), mantle, intestine, and foot. This is similar to what has been reported on vertebrate β-defensin, which was overly expressed in epithelium (Yang *et al.* 1999). That defensin was also expressed in the zebra mussel ovary is also consistent with the previous report (Radhakrishnan *et al.* 2007), which demonstrated that vertebrate β-defensin was expressed in the murine ovary. The authors suggested that the presence of defensin in murine ovaries has the function of protection of fertilization. It was noticeable that the relative abundance of Dpd is extremely high in the zebra mussel foot. The FISH results demonstrated that the Dpd expressed by the epithelial lining of the foot was much less than Dpd expression in byssal glands. Such high level of Dpd within the byssus gland cells suggests that some of the zebra mussels attempt to protect the relatively young byssal threads from degradation in the bacteria-rich aquatic environment. In a previous study, it was demonstrated that these same exocrine glands produce a number of foot proteins among other secretory proteins. The Dpd, which had extremely high expression levels within the foot, and was continuously expressed in high levels during the byssogenesis (Figure 6-10), may be the ideal antimicrobial molecule to provide sufficient protection during the byssogenesis. Whether Dpd is the only

antimicrobial peptide employed by the zebra mussel to protect its byssus or not is currently unknown and deserves further investigation.

The Dpd showed certain growth inhibition activity to four Gram-negative bacteria including *Morganella* sp., *P. shigelloides*, *E. tarda*, and *E. coli*, as well as a Gram-positive bacterium *Staphylococcus aureus aureus*, but no obvious inhibition to the only one Gram-positive bacterium or the yeast species used in this study. Both *Cg-def* and MGD-1 were demonstrated to mostly inhibit the growth of many Gram-positive bacteria with only limited inhibition to some *E.coli* strains. The difference in the activities of these three antimicrobial peptides can be possibly attributed to the difference in their three-dimensional structures. Both *Cg-def* and MGD-1 have eight cysteines in their amino acid sequences, which help proteins to form four disulfide bonds, instead of the three that exist in most insect defensin molecules. Moreover, the homologues searching analysis indicated that the Dpd is closer to the defensins from insects, such as that of the bee (*N. vitripennis*, *B. ignitus*, and *A. mellifera*) and beetle (*A. cuprea*), instead of their molluscan relatives. In insects, the antimicrobial activities of defensins have broader spectrum against the pathogens compared to *Cg-def* and MGD-1, including gram-positive / negative bacteria and fungi (Hoffmann & Hetru 1992; Cociancich *et al.* 1993; Hetru *et al.* 1998).

The recombinant Dpd peptide used in the MIC study was produced in *E. coli* and refolded *in vitro*. There are concerns about the correct *in vitro* refolding of eukaryote proteins with multiple disulfide bonds when expressed in prokaryote cell system (Harder *et al.* 2001). However, the same system has been successfully applied to study several defensin species, such as human α- and β-defensins (Ouellette *et al.* 2000; Harder *et al.*

2001), as well as the oyster defensin (Gueguen *et al.* 2006). The same method we used in this study produced considerable amount of proteins with the MICs analysis. It was noticeable that the recombinant proteins purified under the denatured conditions did not show any inhibition activity against the growth of bacteria and fungus. This suggests that the refolded zebra mussel defensin retains some of its biological activities. Due to the lack of the similar study with freshwater bivalves, the molecules we can use to compare to the Dpd do not exist to the best of our knowledge.

The known functions of defensin have always centered on its microbicidal effects. However, in this study, it is the first time that defensin was found associatd with a physiological process such as byssal thread formation. Given that the Dpd is largely produced by zebra mussel byssus gland cells, it has to be transported through the byssus apparatus to the outside of the mussel where the microbial communities exist. Whether the role of defensin in byssogenesis is limited to biological protection of its microbial degradation or is an integral part of a cascade that leads to the production, secretion, or activation of another proteins remains to be determined. As displayed above (Figure 6-11), Dpd level was elevated as early as Day 2 of byssogenesis denoting to its importance to the process. The data obtained in this study shed more light on the process of byssogenesis in general, and the host defense mechanisms of this nuisant species in particular.

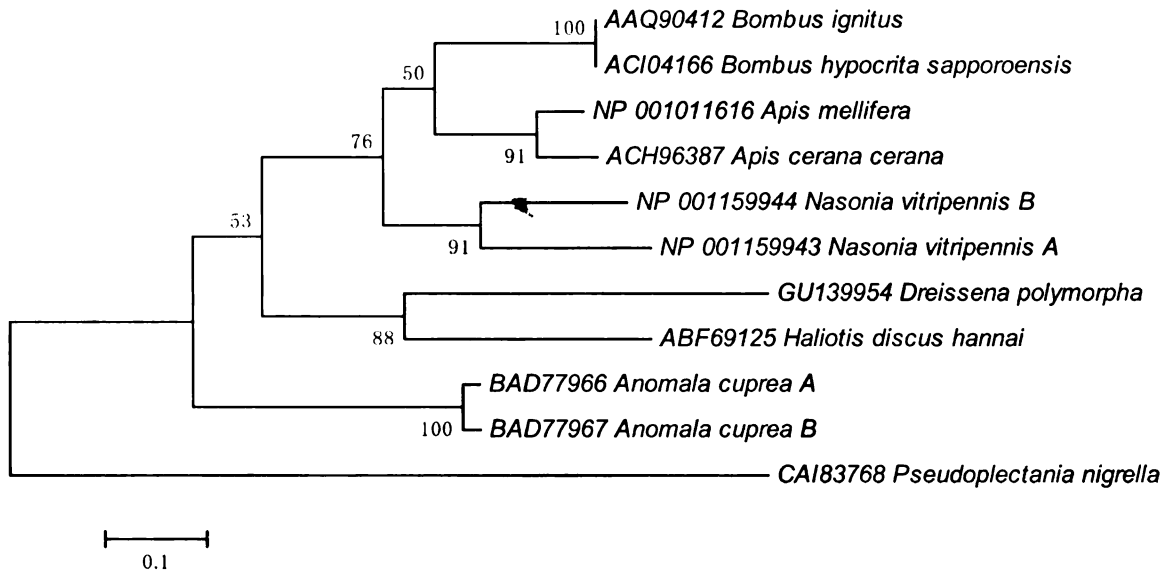


Figure 6-1 Phylogenetic analyses of zebra mussel defensin (Dpd) and the homologues. Neighbor-Joining bootstrap test was used for this analysis. The value above each node stands for the percentage of bootstrapping after 1000 replications. The Dpd was mostly close to a defensin molecule identified from the Japanese disc abalone (*Haliotis discus hannai*). The clade of Dpd and abalone defensin was phylogenetically related to a group of defensins from Apocrita (wasps, ants and bees). The defensin A and B of the scarab beetle (*Anomala cuprea*) had less identity to the defensin of Mollusca and Apocrita. The fungus (*Pseudoplectania nigrella*) defensin was used as an outgroup.

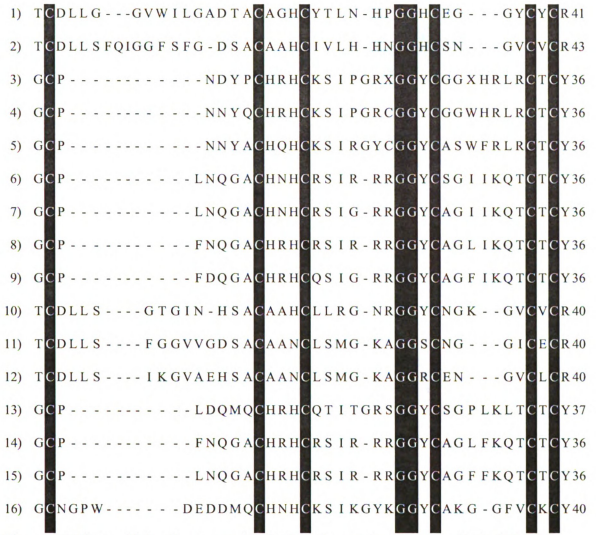


Figure 6-2 The multiple alignment with the homologous insect and mollusk defensins. The conserved amino acid residues and their positions in defensin molecules are highlighted in dark. Typically, the defensin molecules have at least six conserved cysteines which form three disulfide bonds. Also, there are two glycines that are conserved in mollusks, insects, arthropods and fungus. Besides the Dpd (1), another 15 defensins identified from 14 different species were selected in the multiple sequences alignment, including the Japanese disc abalone (2, *Haliotis discus hannai*, ABF69125), the blue mussel (3, *Mytilus edulis*, P81610), the Mediterranean mussel (*Mytilus galloprovincialis*, 4, P80571 and 5, AAD52660), the American dog tick (6, *Dermacentor variabilis*, AAO24323), the bush tick (7, *Haemaphysalis longicornis*, BAD93183), the cattle tick (8, *Boophilus microplus*, AAO48943), the deer tick (9, *Ixodes scapularis*, AAV74387), the Northern blowfly (10, *Protophormia terraenovae*, P10891), the parasitic wasp (11, *Nasonia vitripennis*, NP_001159944), the bumblebee (12, *Bombus ignitus*, AAQ90412), the dragonfly (13, *Aeschna cyanea*, P80154), the fat-tailed scorpion (14, *Androctonus australis*, P56686), the yellow scorpion (15, *Leiurus quinquestriatus*, P41965), and a fungus (16, *Pseudoplectania nigrella*, CAI83768).

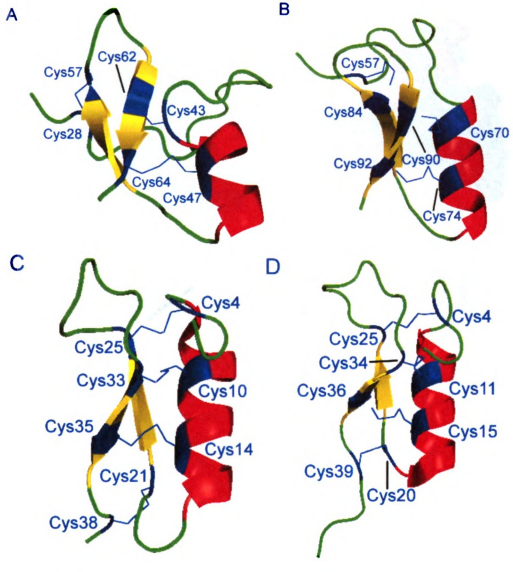


Figure 6-3 The tertiary structures of Dpd and disulfide bonds.
A. the 3D structure of Dpd; B. the 3D structure of insect defensin A (1ICA). Both Dpd and insect defensin A had similar secondary structures, one α-helix (red) and two β-sheets (yellow). The tertiary structures of Dpd and insect defensin A are also very similar. The relative positions of six Cysteins (labeled with blue) and the three disulfide bonds (blue lines between Cysteines) formed by Cysteines in the molecules were conserved. The two defensin molecules from another two mollusks, *Mytilus galloprovincialis* and were also compared to Dpd. C. defensin MGD-1 of *M. galloprovincialis* (1FJN); D. defensin *Cg-def* of *Crassostrea gigas* (2B68). Both MGD-1 and *Cg-def* have four disulfide bonds instead of three in Dpd and insect defensin A.

This figure is in color.

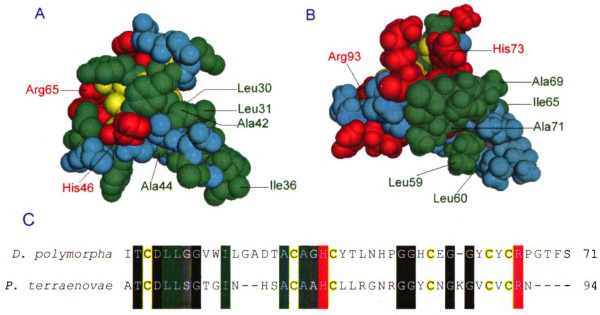


Figure 6-4 The conserved amino acid residues in Dpd.
The amino acid residues of both Dpd and insect defensin A are characterized with different colors. Both hydrophobic (green) and positively charged (red) residues at pH7.0 are demonstrated. The two positively charged residues His^{46} and Arg^{65} in Dpd (A) were also observed in insect defensin A (B) at conserved positions (His^{73} and Arg^{93}). Five hydrophobic residues were found in both Dpd (Leu^{30}, Leu^{31}, Ile^{36}, Ala^{42}, and Ala^{44}) and insect defensin A (Leu^{59}, Leu^{60}, Ile^{65}, Ala^{69}, and Ala^{71}). The six conserved Cysteines are also labeled (yellow) in both tertiary structures (A and B) and primary structures (C) of both defensin proteins. The other conserved amino acid residues are highlighted with dark and the residues with similar biochemical characters at same positions are colored with grey.

This figure is in color.

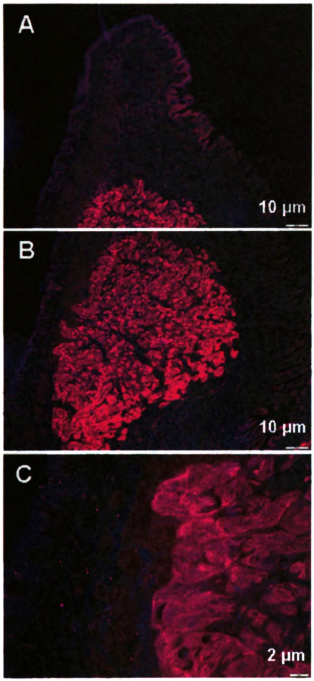


Figure 6- 5 FISH result of Dpd within zebra mussel foot.
The Dpd was observed abundantly expressed in the zebra mussel foot. The Dpd was slightly expressed in the epithelial cells on the surface of zebra mussel foot (A). The main resource of expressed Dpd was observed in the middle of zebra mussel foot (B), filled with byssus gland cells (C). The blue background was the signal given by the stained cytoplasmic actin probe. The pink fluorescent color was the combination of the Dpd probe signal (red) and the cytoplasmic actin probe (blue).

This figure is in color.

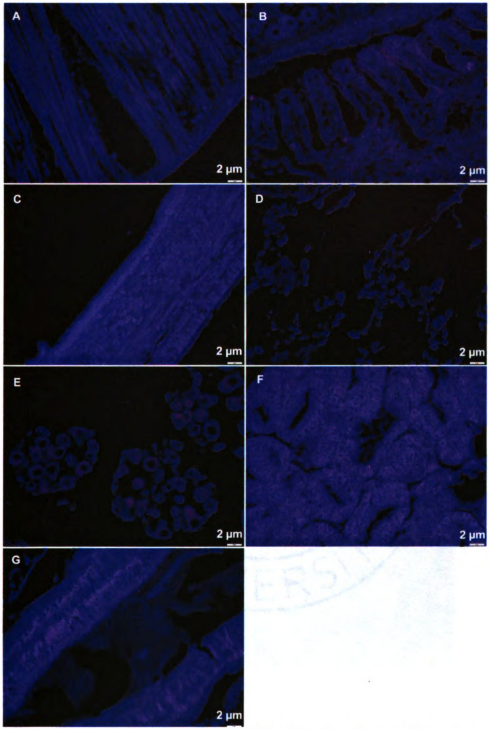


Figure 6-8 The negative control of FISH Dpd in other tissues of the zebra mussel. The in situ hybridization with sense-strand of Dpd RNA probe (red) and anti-sense-strand actin RNA probe (blue). The signal for actin probe can be clearly observed in all tissues but there is no significant fluorescent signal of sense-strand of Dpd. (A) Retractor muscle; (B) Ctenidium; (C) Mantle; (D) Hemocytes; (E) Gonad; (F) Digestive gland; (G) Intestine.

This figure is in color.

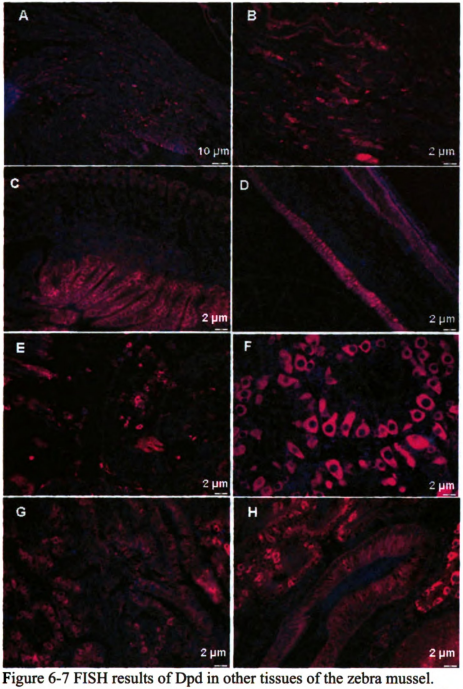


Figure 6-7 FISH results of Dpd in other tissues of the zebra mussel.
The Dpd was also found also expressed in other tissues besides that the zebra mussel foot (red). The expression level of Dpd in the retractor muscle was not significant (A). The sporadically distributed Dpd mRNA was likely from the hemocytes scattered in the muscle (B) since the existence of hemocyte sinus in it. The transcription of Dpd was also detected from ctenidia (C), out layers of mantles (D), hemocytes (E), gonads (F), digestive gland cells (G), and epithelial cells of intestines (H). The cytoplasmic actin was stained as background (blue).

This figure is in color.

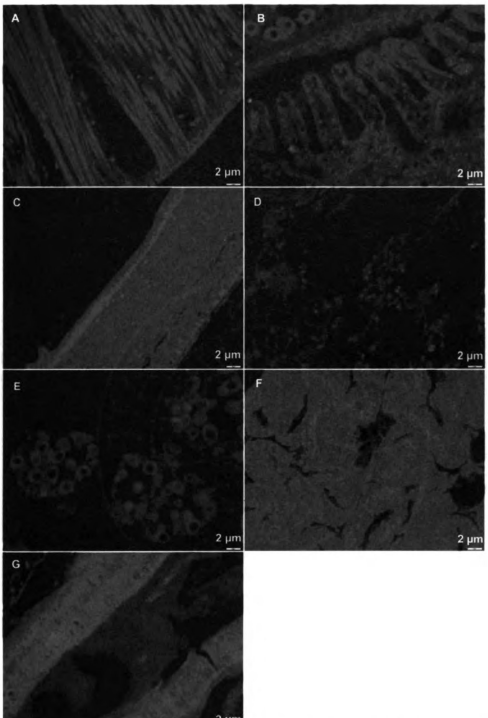


Figure 6-8 The negative control of FISH Dpd in other tissues of the zebra mussel. The in situ hybridization with sense-strand of Dpd RNA probe (red) and anti-sense-strand actin RNA probe (blue). The signal for actin probe can be clearly observed in all tissues but there is no significant fluorescent signal of sense-strand of Dpd. (A) Retractor muscle; (B) Ctenidium; (C) Mantle; (D) Hemocytes; (E) Gonad; (F) Digestive gland; (G) Intestine.

This figure is in color.

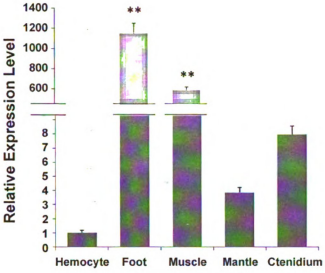


Figure 6- 9 The relative abundance of Dpd in different zebra mussel tissues with qPCR. The expression levels of Dpd in zebra mussel foot and muscle are significantly higher than other tissues ($P < 0.01$). The abundance of Dpd in the zebra foot is also significantly higher than that in muscle. The ANOVA with Tukey's test was used for the multiple comparisons. ** $P < 0.01$.

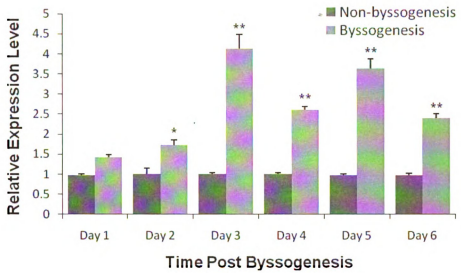


Figure 6-10 The relative expression levels of Dpd under the condition of non-byssogenesis and byssogenesis.
Compared to the non-byssogenesis status, the expression levels of Dpd in zebra mussel with adhesion are significantly higher at most time point post the start of byssogenesis except for Day 1. The ANOVA with Tukey's test was used for the multiple comparisons. * $P < 0.05$; ** $P < 0.01$.

Table 6-1 The minimal inhibition concentrations (MICs) of Dpd against the bacteria and yeast.

Microorganisms	**MIC (μM)**
Bacteria	
Gram-negative	
Morganella sp.	0.35
Plesiomonas shigelloides	0.43
Edwardsiella tarda	1.16
Escherichia coli DH5α	6.46
Aeromonas salmonicida salmonicida	>50
Motile Aeromonas	>50
Flavobacterium sp.	>50
Pseudomonas fluorescens	>50
Shewanella putrifaciens	>50
Gram-positive	
Staphylococcus aureus aureus ATCC12598	30.39
Bacillus megaterium	>50
Carnobacterium maltaromaticum ATCC27865	>50
Enterococcus faecalis ATCC19433	>50
Micrococcus luteus ATCC4698	>50
Fungus	
Saccharomyces cerevisiae	>50

CHAPTER VII

Conclusions and Future Studies

Conclusions

With the results obtained from this project, identification of the genes associated with the byssus activities, we were able to gradually understand the regulation mechanism of zebra mussel attachment. The methodology of this study has been proven sufficient and reliable.

The application of suppression subtractive hybridization (SSH) cDNA technique dramatically decreased the population of the genes of interest. In eukaryotes, the mRNAs of a typical somatic cell are distributed in three frequency classes (Bishop *et al.* 1974; Davidson & Britten 1979). The two classes with the most abundance (> 90% of total mRNA population) consist of the genes involved in the common metabolism of the cells. On average, the most prevalent class consists of about 10 mRNA species, each represented by 5000 copies per cell, whereas the class of high complexity comprises 15,000 different species each represented by 1-15 copies only (Soares *et al.* 1994). The SSH cDNA library technique enriched the cDNAs unique to zebra mussel byssus by subtracting the muscle RNA from the foot RNA. Compared to the 3000 genes obtained from non-normalized hemocyte cDNA library of scallop *Argopecten irradians* (Song *et al.* 2006) and *Chlamys farreri* (Wang *et al.* 2009), respectively, the 750 ESTs generated from normalized zebra mussel foot cDNA library with each of the rare genes represented that is related to the function of the foot. By quantitative PCR, the genes selected from

the microarray were all confirmed to be significantly expressed in zebra mussel foot compared to the other tissues. The decrease of the interference caused by the gene constitutively expressed in all the tissues dramatically improved the efficiency of the cDNA library. However, to identify the genes related to byssus activities of zebra mussel from more than 700 ESTs still seems unpromising until the application of the cDNA microarray technology.

Our newly constructed *Dreissena polymorpha* microarray is the first cDNA microarray designed for the study of the freshwater bivalves and also the first array to in the study of bivalve underwater adhesion mechanism. The cDNA microarray developed in this study contains 716 genes selected from the SSH cDNA library and represents the genes unique to the byssus of zebra mussel. Based on the low false discovery rate (FDR) of this microarray in identifying the differentially expressed genes (Chapter III), the zebra mussel byssus cDNA microarray was proven to be efficient and reliable.

With the applications of zebra mussel cDNA microarray, multiple differentially expressed genes were found from different stages of byssogenesis. Furthermore, the genes with the expression patterns varied in response to the change of environmental factors and byssogenesis status were also identified. Among all these identified genes, not only are the genes encoding the Dpfp-1 variants, but also some other genes with different putative functions found to be involved. The protection to the byssus components provided by zebra mussel has not been widely studied. However, the abundant expression of *D. polymorpha* defensin (Dpd) in zebra mussel byssus gland makes us believe that the byssus gland cells of zebra mussel, primarily as the byssal threads producer, also provide essential protection to the proteins produced by themselves. Combined with the current

knowledge of zebra mussel attachment that was obtained from previous studies, the results of our study partially explained basic questions of the molecular mechanism of zebra mussel underwater adhesion, such as what kind of molecules are produced by zebra mussel byssus glands, which genes were involved in byssogenesis At the same time, so many questions have been raised that the follow-up investigation in this delicate mechanism is necessary in particular.

Future Studies

For as long as history has been recorded, people have been fascinated with underwater adhesion. Water can weaken many of the chemical bonds mediating adhesion. There are many examples of both permanent and temporary underwater attachment that are mediated by complex polymer glues (Waite 1987; Flammang 1996; Smith & Callow 2006). For some aquatic animals, adhesion is vital for survival and completion of their life cycle. One such organism is the zebra mussel (*Dreissena polymorpha*), the bivalve mollusk I have chosen to elucidate with regards to the molecular aspects of its underwater adhesion. The first stumbling block I faced in my studies was the severe lack of background information on the details of underwater attachment processes employed by the zebra mussel. Overwhelmed by the economic losses caused by the zebra mussel in the Laurentian Great Lakes basin, managers and scientists focused their efforts on finding creative ways to clean pipes from blocking by the invading mollusk, with relatively minimal efforts directed towards studying its biology. It is therefore that I find the study of the physiology, defense mechanisms, and attachment mechanisms of the two nuisant dreissenids (zebra and quagga mussels) a first priority for future research and funding.

Studies addressing underwater adhesion by marine invertebrates, which live in dynamic ocean environments by adhering tightly to underwater substratum using their holdfasts, were the only sources available upon starting my research project. Unfortunately, they were of little help as they focused on the biochemical nature of the proteinaceous glue produced by the marine mussels. The approach used in my studies fills a void in our knowledge of the zebra mussel physiology and underwater attachment. Due to the successful application of cDNA microarray technique, many zebra mussel

novel genes were identified and their kinetics of expression followed along the byssogenic cycle.

While the nature and functions of these genes have been identified, whether through experiments performed in my studies or through the functions of their homologues, there is a great need to define these genes, complete sequencing of them, and deduce the nature of the proteins they encode and their potential functions. In this respect, I would argue that the EGP genes constitute a first priority in future research as they may represent important components of a cascade leading to the formation of adhesive particles, alone or combined with other secreted foot proteins. Good candidates are the shematrin-like and neuropeptide (nlp) encoding genes; both were found differentially displayed in most of the experiments performed in this study. The shematrins were first isolated from the mantle tissue of the pearl oyster (*Pinctada fucata*) and defined as a glycine-rich protein family believed to form a scaffold for various structures (Yano *et al.* 2006). Whether zebra mussel shematrins are also essential for the byssus structure warrants further studies. The functions of nlp family cover a broad range including neuronal signaling, development, and antimicrobial activities (Nathoo *et al.* 2001; Couillault *et al.* 2004; McVeigh *et al.* 2008). Whether the nlp could be essential for a kind of "sensation" by which the zebra mussel communicates with the newly formed byssal thread, and thereby determine the surrounding environmental factors, including the suitability of the substrates for adhesion, is definitely important to study in the future.

Shematrin and nlp are examples of those zebra mussel proteins that have hits in the gene databases, however, almost 40% of the zebra mussel foot genes have had no hits, and based on the experiments presented in Chapter III-V, these functionally-unknown

genes seem to play a vital role in zebra mussel adhesions. A study to elucidate their function relative to byssogenesis is deemed necessary for future studies. Indeed, some of the genes without putative functions in the library encode amino acid residues that share great similarities with the zebra and blue mussel foot proteins.

The expressions of many genes in the zebra mussel byssus cDNA library seem to be modulated by the fluctuation in environmental factors and status of byssogenesis (Chapter V). It was surprising to determine that most of the differentially expressed genes are responsive to changes in one or two factors, and not to all four of the factors studied. This observation sheds light on the fact that the expression of these genes may be very specific to one particular factor and not to another. For example, the expression of EGP homologues was not affected by the current velocity. Since EGP seems to play a central role in underwater adhesion, it is empirical that future studies address the role of the surrounding environmental factors in the initiation and magnitude of EGP gene expression. Some other examples were also found specific to certain factors, such as a polypeptide release factor encoding gene whose expression profiles were only altered by water temperature; and choriogenin H-like protein encoding gene with the expression pattern only affected by temperature and current velocity. This suggests that zebra mussel underwater adhesion is not a universal mechanism, rather a number of either independent or sequential responses to various stimuli.

In our study, only two different levels were created for each factor, which is not enough to determine a solid correlation between the expression of certain zebra mussel genes identified in this study for their role in byssogenesis and surrounding environmental factors. In previous morphological studies, multiple levels have been

created for each environmental factor to study the regression of the byssogenesis rate and each experimental factor (Clarke & McMahon 1996b, c; Clarke & McMahon 1996a; Clarke 1999). Applying multiple levels in each factor will allow further study of how the expression levels of certain genes are associated with changes within the environment. This will help us to unravel the regulation mechanism of the byssus unique genes during the byssogenesis in natural conditions. Additionally, the genes found to be involved in byssogenesis in the present study should be further analyzed, their functions deciphered, and factors affecting their expression determined.

The time course study involving the developed cDNA microarray successfully identified genes whose differential expression patterns coincided with morphological changes occurring in byssal threads (Chapter IV). For example, homologous to annexin VII encoding gene, the gene BG20_G04 was only upregulated 12 hours after the start of byssogenesis. Within the first 12 hours, the temporary threads are dominant in freshly produced byssal threads, and the protein product of BG20_G04 gene probably provides material for temporary threads. This particular gene is an excellent candidate for future studies as it may provide details regarding the early stages of byssogenesis.

Besides the zebra mussel models used in our microarray analysis, adhesion vs. non-adhesion (Chapters III and V) and byssogenesis vs. non-byssogenesis (Chapter IV), there is another model that may also be studied with our cDNA microarray: attachment vs. detachment. According to the observations in earlier studies, both adult and juvenile zebra mussels can detach themselves and move using their feet. The crawl rate of juvenile zebra mussel is about 7 cm/night while the adults can move at a rate up to 36 cm/h (Ackerman *et al.* 1994; Toomey *et al.* 2002). This suggests that zebra mussels can

detach themselves, probably by dissolving their previously secreted glue with other molecules of lytic nature (Claudi & Mackie 1994). Finding the nature of the zebra mussel lytic molecule may be a cornerstone toward developing control agents against the nuisant organism. If a laboratory model can be developed to force attached mussels to detach, then the cDNA microarray can be used to identify the gene(s) encoding the glue-dissolving molecule(s) employed by the zebra mussel.

In order to settle down in a particular area, zebra mussels sense the substance surfaces and surrounding environment using their feet during movement (Eckroat *et al.* 1993). This cannot be accomplished without the presence of certain sensory elements. The conduction through the pedal ganglion of *Mytilus* has been reported (Richards 1929), but to date, this phenomenon is very under-appreciated and little understood. It remains unclear whether the zebra mussel foot sensing process is accomplished by the nerves and pedal ganglion like the blue mussel, or if the zebra mussel employs other yet to be discovered, mechanisms. As stated previously, the genes homologous to neuropeptide encoding genes can be potentially involved in the neuronal signaling (Nathoo *et al.* 2001; Couillault *et al.* 2004; McVeigh *et al.* 2008). The analysis of this family of genes may contribute to the understanding of surface recognizing mechanisms employed by zebra mussel, which is a primary step of underwater attachment.

The Dpfp-1 encoding gene has been identified in most microarray experiments performed in this study. The Dpfp-1 contains tandemly repeated and segregated motifs: a heptapeptide and a tridecapeptide consensus motif. At least ten Dpfp-1 variants exist in the protein mass purified from zebra mussel byssus. The multiple protein variants encoded by a single Dpfp gene is probably caused by an alternative splicing mechanism

on gene transcription level which is similar to the other collagen-like proteins (Pihlajaniemi & Tamminen 1990; Rzepecki & Waite 1993). The transcripts of the Dpfp-1 identified in our data can be directly used to study the Dpfp-1 expression regulation in transcription level, which will be essential to obtain more information on byssal thread formation in the zebra mussel.

Although the microarray assays greatly narrowed down the genes of our interests, there are still a large number of genes that need to be studied. The difference of the gene expression on mRNA level may not lead to the change of protein level due to the decay of mRNA in post-transcriptional regulation (Jacobson & Peltz 1996; Ross 1996). Among the three previously studied *D. polymorpha* foot proteins, only two of them were purified directly from the zebra mussel foot (Rzepecki & Waite 1993). It is also likely that, among the EGP genes and nlp-like genes identified by microarray, there are only a few whose differential transcriptional patterns will be also reflected by the differential translation profiles, which means most of the identified genes may not be as heavily involved in zebra mussel byssus activities, as one may think. At this point, analyses of the proteins encoded by the genes identified by microarray, such as Western blot, spectrophotometry, and enzyme-linked immunosorbent assay (ELISA), will help us validate the results on protein level, thereupon discover the genes that are indeed involved in zebra mussel adhesion.

On the protein level, it is hard to locate the zebra mussel byssal thread adhesive proteins due to the lack of commercially available antibodies. The difficulties in obtaining the pure adhesive proteins from zebra mussel byssal thread plaques are the main obstacle in producing the antibodies against adhesive proteins (Farsad *et al.* 2009).

With the genes identified from this study, we can easily produce a large amount of pure foot proteins through *in vitro* expression systems; whereafter, the antibodies against many types of foot proteins can be developed with the recombinant foot proteins. The availability of the foot protein antibodies will contribute to the discovery of novel proteins produced by zebra mussel byssus glands, as well as the structure of the bulk proteins in the plaques of zebra mussel byssal threads.

REFERENCE

Ackerman J.D., Smith B., Nichols S.J. & Claudi R. (1994) A review of the early life history of zebra mussels (*Dreissena polymorpha*): comparisons with marine bivalves. *Can J Zool* **72**, 1169-79.

Allen J.A. (1985) *The recent Bivalvia: their form and function.* Academic Press, New York.

Altschul S.F., Gish W., Miller W., Myers E.W. & Lipman D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**, 403-10.

Anderson K.E. & Waite J.H. (1998) A major protein precursor of zebra mussel (*Dreissena polymorpha*) byssus: deduced sequence and significance. *Biol Bull* **194**, 150-60.

Anderson K.E. & Waite J.H. (2000) Immunolocalization of Dpfp1, a byssal protein of the zebra mussel *Dreissena polymorpha. J Exp Biol* **203**, 3065-76.

Andrade T.P.D., Redmana R.M. & Lightner D.V. (2008) Evaluation of the preservation of shrimp samples with Davidson's AFA fixative for infectious myonecrosis virus (IMNV) *in situ* hybridization. *Aquaculture* **278**, 179-83.

Angarano M.B., McMahon R.F. & Schetz J.A. (2009) Cannabinoids inhibit zebra mussel (*Dreissena polymorpha*) byssal attachment: a potentially green antifouling technology. *Biofouling* **25**, 127-38.

Arumugam M., Romestand B., Torreilles J. & Roch P. (2000) In vitro production of superoxide and nitric oxide (as nitrite and nitrate) by *Mytilus galloprovincialis* haemocytes upon incubation with PMA or laminarin or during yeast phagocytosis. *Eur J Cell Biol* **79**, 513-9.

Auffret M. (2005) Bivalves as models for marine immunotoxicology. In: *Investigative Immunotoxicology* (eds. by Tryphonas H, Fournier M & Blakley BR). CRC Press, Boca Raton, FL.

Baker S.M. & Hornbach D.J. (1997) Acute physiological effects of zebra mussel (*Dreissena polymorpha*) infestation on two unionid mussels, *Actinonaias ligmentina* and *Amblema plicata*. *Can J Fish Aquat Sci* **54**, 512-9.

Benedict C.V. & Waite J.H. (1986) Location and analysis of byssal structural proteins of Mytilus edulis. *J Morphol* **189**, 171-81.

Benson A.J. & Raikow D. (2009) *Dreissena polymorpha*. USGS nonindigenous aquatic species database. URL http://nas.er.usgs.gov/queries/FactSheet.asp?speciesID=5.

Bially A. & MacIsaac H.G. (2000) Fouling mussels (*Dreissena* spp.) colonize soft sediments in Lake Erie and facilitate benthic invertebrates. *Freshwater Biol* **43**, 85-97.

Bishop J.O., Morton J.G., Rosbash M. & Richardson M. (1974) Three abundance classes in HeLa cell messenger RNA. *Nature* **250**, 199-204.

Boman H.G. (2003) Antibacterial peptides: basic facts and emerging concepts. *J Intern Med* **254**, 197-215.

Bonmatin J.M., Bonnat J.L., Gallet X., Vovelle F., Ptak M., Reichhart J.M., Hoffmann J.A., Keppi E., Legrain M. & Achstetter T. (1992) Two-dimensional 1H NMR study of recombinant insect defensin A in water: resonance assignments, secondary structure and global folding. *J Biomol NMR* **2**, 235-56.

Bonner T.P. & Rockhill R. (1994a) Functional morphology of the zebra mussel byssus. *Dreissena polymorpha. Info Rev.* **5**, 4-7.

Bonner T.P. & Rockhill R.L. (1994b) Ultrastructure of the byssus of the zebra mussel (*Dreissena polymorpha*, Mollusca: Bivalvia). *Trans Am Microsc Soc* **113**, 302-15.

Boulanger N., Lowenberger C., Volf P., Ursic R., Sigutova L., Sabatier L., Svobodova M., Beverley S.M., Spath G., Brun R., Pesson B. & Bulet P. (2004) Characterization of a defensin from the sand fly Phlebotomus duboscqi induced by challenge with bacteria or the protozoan parasite Leishmania major. *Infect Immun* **72**, 7140-6.

Bowtell D. & Sambrook J. (2003) *DNA microarrays. A molecular cloning manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Brown C.H. (1952) Some structural proteins of *Mytilus edulis* L. *Q J Microsc Sci* **93**, 487-502.

Bulet P., Stocklin R. & Menin L. (2004) Anti-microbial peptides: from invertebrates to vertebrates. *Immunol Rev* **198**, 169-84.

Carballal M.J., Lopez M.C., Azevedo C. & Villalba A. (1997) Hemolymph cell types of the mussel *Mytilus galloprovincialis*. *Dis Aquat Org* **29**, 127-35.

Cayless R.A. (1991) Interfacial chemistry and adhesion: the role of surface analysis in the design of strong stable interfaces for improved adhesion and durability. *Sur Inter Anal* **17**, 430-8.

Cha H.J., Hwang D.S. & Lim S. (2008) Development of bioadhesives from marine mussels. *Biotechnol J.*

Chandrasekharappa S., Holloway A., Iyer V., Monte D., Murphy M. & Nowak N.J. (2003) Generation of probes for spotted microarrays. In: *DNA microarrays: a molecular cloning manual* (ed. by Bowtell DaS, J.), pp. 1-60. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Cheng T.C. (1983) Bivalves. In: *Invertebrate blood cells* (eds. by Ratcliffe NA & Rowley AF), pp. 234-300. Academic Press, London.

Chu F.L.E. (1988) Humoral defense factors in marine bivalves. In: *Disease Processes in Marine Bivalve Molluscs* (ed. by Fisher WS), pp. 178-88. American Fisheries Society Special publication, Bethesda, MD.

Chu S., DeRisi J., Eisen M., Mulholland J., Botstein D., Brown P.O. & Herskowitz I. (1998) The transcriptional program of sporulation in budding yeast. *Science* **282**, 699-705.

Churchill G.A. (2002) Fundamentals of experimental design for cDNA microarrays. *Nat Genet* **32 Suppl**, 490-5.

Ciornei C.D., Sigurdardottir T., Schmidtchen A. & Bodelsson M. (2005) Antimicrobial and chemoattractant activity, lipopolysaccharide neutralization, cytotoxicity, and inhibition by serum of analogs of human cathelicidin LL-37. *Antimicrob Agents Chemother* **49**, 2845-50.

Clarke M. (1999) The effect of food availability on byssogenesis by the zebra mussel (*Dreissena polymorpha* Pallas). *J Moll Stud* **65**, 327-33.

Clarke M. & McMahon R.F. (1996a) Effects of current velocity on byssal-thread production in the zebra mussel (*Dreissena polymorpha*). *Can J Zool* **74**, 63-71.

Clarke M. & McMahon R.F. (1996b) Effects of hypoxia and low-frequency agitation on byssogenesis in the freshwater mussel *Dreissena polymorpha* (Pallas) *Biol Bull* **191**, 413-9.

Clarke M. & McMahon R.F. (1996c) Effects of temperature on byssal thread production by the freshwater mussel, *Dreissena polymorpha* (Pallas). *Am Malacol Bull* **13**, 105-10.

Claudi R. & Mackie G. (1994) *Practical Manual for Zebra Mussel Monitoring and Control*. Lewis Publishers, Boca Raton, FL.

Claudi R. & Mackie G.L. (1993) *Practical manual for zebra mussel monitoring and control*. Lewis Publishers, Boca Raton, FL.

Cociancich S., Ghazi A., Hetru C., Hoffmann J.A. & Letellier L. (1993) Insect defensin, an inducible antibacterial peptide, forms voltage-dependent channels in *Micrococcus luteus*. *J Biol Chem* **268**, 19239-45.

Cole A.M., Hong T., Boo L.M., Nguyen T., Zhao C., Bristol G., Zack J.A., Waring A.J., Yang O.O. & Lehrer R.I. (2002) Retrocyclin: a primate peptide that protects cells from infection by T- and M-tropic strains of HIV-1. *Proc Natl Acad Sci U S A* **99**, 1813-8.

Conesa A., Gotz S., Garcia-Gomez J.M., Terol J., Talon M. & Robles M. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674-6.

Cope W.G., Bartsch M.R. & Marking L.L. (1997) Efficacy of candidate chemicals for preventing attachment of zebra mussels (*Dreissena Polymorpha*). *Environ Toxicol Chem* **16**, 1930-4.

Cornet B., Bonmatin J.M., Hetru C., Hoffmann J.A., Ptak M. & Vovelle F. (1995) Refined three-dimensional solution structure of insect defensin A. *Structure* **3**, 435-48.

Couillault C., Pujol N., Reboul J., Sabatier L., Guichou J.F., Kohara Y. & Ewbank J.J. (2004) TLR-independent control of innate immunity in *Caenorhabditis elegans* by the TIR domain adaptor protein TIR-1, an ortholog of human SARM. *Nat Immunol* **5**, 488-94.

Coyne K.J. & Waite J.H. (2000) In search of molecular dovetails in mussel byssus: from the threads to the stem. *J Exp Biol* **203**, 1425-31.

Cunliffe R.N. (2003) Alpha-defensins in the gastrointestinal tract. *Mol Immunol* **40**, 463-7.

Davids B.J. & Yoshino T.P. (1998) Integrin-like RGD-dependent binding mechanism involved in the spreading response of circulating molluscan phagocytes. *Dev Comp Immunol* **22**, 39-53.

Davidson E.H. & Britten R.J. (1979) Regulation of gene expression: possible role of repetitive sequences. *Science* **204**, 1052-9.

DeLano W.L. (2009) The PyMOL molecular graphics system. URL http://www.pymol.org.

Deming T.J. (1999) Mussel byssus and biomolecular materials. *Curr Opin Chem Biol* **3**, 100-5.

DeRisi J.L., Iyer V.R. & Brown P.O. (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**, 680-6.

Diatchenko L., Lau Y.F., Campbell A.P., Chenchik A., Moqadam F., Huang B., Lukyanov S., Lukyanov K., Gurskaya N., Sverdlov E.D. & Siebert P.D. (1996) Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci U S A* **93**, 6025-30.

Diatchenko L., Lukyanov S., Lau Y.F. & Siebert P.D. (1999) Suppression subtractive hybridization: a versatile method for identifying differentially expressed genes. *Methods Enzymol* **303**, 349-80.

Dobbin K., Shih J.H. & Simon R. (2003) Questions and answers on design of dual-label microarrays for identifying differentially expressed genes. *J Natl Cancer Inst* **95**, 1362-9.

Dondero F., Piacentini L., Marsano F., Rebelo M., Vergani L., Venier P. & Viarengo A. (2006) Gene transcription profiling in pollutant exposed mussels (*Mytilus* spp.) using a new low-density oligonucleotide microarray. *Gene* **376**, 24-36.

Dormon J.M., Coish C., Cottrell C., Allen D.G. & Spelt J.K. (1997) Modes of byssal failure in forced detachment of zebra mussels. *J Environ Eng* **123**, 933-8.

Douglas S.E., Gallant J.W., Liebscher R.S., Dacanay A. & Tsoi S.C. (2003) Identification and expression analysis of hepcidin-like antimicrobial peptides in bony fish. *Dev Comp Immunol* **27**, 589-601.

Draghici S., Kuklin A., Hoff B. & Shams S. (2001) Experimental design, analysis of variance and slide quality assessment in gene expression arrays. *Curr Opin Drug Discov Devel* **4**, 332-7.

Dufva M. (2009) Introduction to microarray technology. *Methods Mol Biol* **529**, 1-22.

Eckroat L.R., Masteller E.C., Shaffer J.C. & Steele L.M. (1993) The byssus of the zebra mussel (*Dreissena polymorpha*): morphology, byssal thread formation and detachment. In: *Zebra mussels: biology, impacts, and control* (eds. by Nalepa TF & Schloesser DW), pp. 239-63. Lewis, Boca Raton, FL, USA.

Eleveld-Trancikova D., Kudela P., Majerciak V., Regendova M., Zelnik V., Pastorek J., Pastorekova S. & Bizik J. (2002) Suppression subtractive hybridisation to isolate differentially expressed genes involved in invasiveness of melanoma cell line cultured under different conditions. *Int J Oncol* **20**, 501-8.

Evans T.G. & Somero G.N. (2008) A microarray-based transcriptomic time-course of hyper- and hypo-osmotic stress signaling events in the euryhaline fish *Gillichthys mirabilis*: osmosensors to effectors. *J Exp Biol* **211**, 3636-49.

Ewing B. & Green P. (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* **8**, 186-94.

Ewing B., Hillier L., Wendl M.C. & Green P. (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* **8**, 175-85.

Farsad N., Gilbert T.W. & Sone E.D. (2009) Adhesive structure of the freshwater zebra mussel, *Dreissena polymorpha*. In: *Mater. Res. Soc. Symp. Proc.*, pp. 1187-KK02-03 Warrendale, PA.

Filpula D.R., Lee S.M., Link R.P., Strausberg S.L. & Strausberg R.L. (1990) Structural and functional repetition in a marine mussel adhesive protein. *Biotechnol Prog* **6**, 171-7.

Flammang P. (1996) Adhesion in echinoderms. In: *Echinoderm studies* (eds. by Jangoux M & Lawrence JM), pp. 1-60. A. A. Balkema, Rotterdam, Netherlands.

Francischetti I.M., My Pham V., Mans B.J., Andersen J.F., Mather T.N., Lane R.S. & Ribeiro J.M. (2005) The transcriptome of the salivary glands of the female western black-legged tick *Ixodes pacificus* (Acari: Ixodidae). *Insect Biochem Mol Biol* **35**, 1142-61.

Frischer M.E., Nierzwicki-Bauer S.A., Parsons R.H., Vathanodorn K. & Waitkus K.R. (2000) Interactions between zebra mussels (*Dreissena polymorpha*) and microbial communities. *Can. J. Fish. Aquat. Sci.* **57**, 591-9.

Frisina A.C. & Eckroat L.R. (1992) Histological and morphological attributes of the byssus of the zebra mussel *Dreissena polymorpha* (Pallas). *J PA Acad Sci* **66**, 63-7.

Fritz B., Schubert F., Wrobel G., Schwaenen C., Wessendorf S., Nessling M., Korz C., Rieker R.J., Montgomery K., Kucherlapati R., Mechtersheimer G., Eils R., Joos S. & Lichter P. (2002) Microarray-based copy number and expression profiling in dedifferentiated and pleomorphic liposarcoma. *Cancer Res* **62**, 2993-8.

Gasteiger E., Gattiker A., Hoogland C., Ivanyi I., Appel R.D. & Bairoch A. (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* **31**, 3784-8.

Giamberini L., Auffret M. & Pihan J.C. (1996) Haemocytes of the freshwater mussel, *Dreissena polymorpha* (Pallas): cytology, cytochemistry, and E-ray microanalysis. *J Moll Stud* **62**, 367-79.

Gueguen Y., Herpin A., Aumelas A., Garnier J., Fievet J., Escoubas J.M., Bulet P., Gonzalez M., Lelong C., Favrel P. & Bachere E. (2006) Characterization of a defensin from the oyster Crassostrea gigas. Recombinant production, folding, solution structure, antimicrobial activities, and gene expression. *J Biol Chem* **281**, 313-23.

Hancock R.E. & Lehrer R. (1998) Cationic peptides: a new source of antibiotics. *Trends Biotechnol* **16**, 82-8.

Harder J., Bartels J., Christophers E. & Schroder J.M. (2001) Isolation and characterization of human beta -defensin-3, a novel human inducible peptide antibiotic. *J Biol Chem* **276**, 5707-13.

Hart R.A., Miller A.C. & Davis M. (2001) Empirically derived survival rates of a native mussel, *Amblema plicata*, in the Mississippi and Otter tail rivers, Minnesota. *Am Midland Naturalist* **146**, 254-63.

Hebert P.D.N., Muncaster B.W. & Mackie G.L. (1989) Ecological and genetic studies on *Dreissena polymorpha* (Pallas): a new mollusc in the Great Lakes. *Can J Fish Aquat Sci* **46**, 1587-91.

Hetru C., Hoffmann D. & Bulet P. (1998) Antimicrobial peptides from insects. In: *Molecular Mechanisms of Immune Responses in Insects* (eds. by Brey PT & Hultmark D), pp. 40-66. Chapman & Hall, London.

Hine P.M. (1999) The inter-relationships of bivalve haemocytes. *Fish Shellfish Immunol* **9**, 367-85.

Hodgson G., Hager J.H., Volik S., Hariono S., Wernick M., Moore D., Nowak N., Albertson D.G., Pinkel D., Collins C., Hanahan D. & Gray J.W. (2001) Genome scanning with array CGH delineates regional alterations in mouse islet carcinomas. *Nat Genet* **29**, 459-64.

Hoffmann J.A. & Hetru C. (1992) Insect defensins: inducible antibacterial peptides. *Immunol Today* **13**, 411-5.

Huang X. & Madan A. (1999) CAP3: A DNA sequence assembly program. *Genome Res* **9**, 868-77.

Hubert F., Noel T. & Roch P. (1996) A member of the arthropod defensin family from edible Mediterranean mussels (*Mytilus galloprovincialis*). *Eur J Biochem* **240**, 302-6.

Inoue K. & Odo S. (1994) The adhesive protein cDNA of *Mytilus galloprovincialis* encodes decapeptide repeats but no hexapeptide motif. *Biol Bull* **186**, 349-55.

Inoue K., Takeuchi Y., Miki D. & Odo S. (1995) Mussel adhesive plaque protein gene is a novel member of epidermal growth factor-like gene family. *J Biol Chem* **270**, 6698-701.

Inoue K., Takeuchi Y., Takeyama S., Yamaha E., Yamazaki F., Odo S. & Harayama S. (1996) Adhesive protein cDNA sequence of the mussel *Mytilus coruscus* and its evolutionary implications. *J Mol Evol* **43**, 348-56.

Jacobson A. & Peltz S.W. (1996) Interrelationships of the pathways of mRNA decay and translation in eukaryotic cells. *Annu Rev Biochem* **65**, 693-739.

Jenny M.J., Chapman R.W., Mancia A., Chen Y.A., McKillen D.J., Trent H., Lang P., Escoubas J.M., Bachere E., Boulo V., Liu Z.J., Gross P.S., Cunningham C., Cupit P.M., Tanguy A., Guo X., Moraga D., Boutet I., Huvet A., De Guise S., Almeida J.S. & Warr G.W. (2007) A cDNA microarray for *Crassostrea virginica* and *C. gigas*. *Mar Biotechnol (NY)* **9**, 577-91.

Ji W., Wright M.B., Cai L., Flament A. & Lindpaintner K. (2002) Efficacy of SSH PCR in isolating differentially expressed genes. *BMC Genomics* **3**, 12.

Johnson L.E. & Padilla D.K. (1996) Geographic spread of exotic species: ecological lessons and opportunities from the invasion of the zebra mussel (*Dreissena polymorpha*). *Biol Conserv* **78**, 23-33.

Kavouras J.H. & Maki J.S. (2003a) Effects of biofilms on zebra mussel postveliger attachment to artificial surfaces. *Invert Biol* **122**, 138-51.

Kavouras J.H. & Maki J.S. (2003b) The effects of natural biofilms on the reattachment of young adult zebra mussels to artificial substrata. *Biofouling* **19**, 247-56.

Kavouras J.H. & Maki J.S. (2004) Inhibition of the reattachment of young adult zebra mussels by single-species biofilms and associated exopolymers. *J Appl Microbiol* **97**, 1236-46.

Kerr M.K. (2003) Design considerations for efficient and effective microarray studies. *Biometrics* **59**, 822-8.

Kerr M.K. & Churchill G.A. (2001) Experimental design for gene expression microarrays. *Biostatistics* **2**, 183-201.

Kulesh D.A., Clive D.R., Zarlenga D.S. & Greene J.J. (1987) Identification of interferon-modulated proliferation-related cDNA sequences. *Proc Natl Acad Sci U S A* **84**, 8453-7.

Lambert C., Leonard N., De Bolle X. & Depiereux E. (2002) ESyPred3D: Prediction of proteins 3D structures. *Bioinformatics* **18**, 1250-6.

Langley R.R., Ramirez K.M., Tsan R.Z., Van Arsdall M., Nilsson M.B. & Fidler I.J. (2003) Tissue-specific microvascular endothelial cell lines from H-2K(b)-tsA58 mice for studies of angiogenesis and metastasis. *Cancer Res* **63**, 2971-6.

Lee B.P., Chao C., Nunalee F.N., Motan E., Shull K.R. & Messersmith P.B. (2006) Rapid gel formation and adhesion in photocurable and biodegradable block copolymers with high DOPA content. *Macromolecules* **39**, 1740-8.

Lehrer R.I., Rosenman M., Harwig S.S., Jackson R. & Eisenhauer P. (1991) Ultrasensitive assays for endogenous antimicrobial polypeptides. *J Immunol Methods* **137**, 167-73.

Leung Y.F., Ma P., Link B.A. & Dowling J.E. (2008) Factorial microarray analysis of zebrafish retinal development. *Proc Natl Acad Sci U S A* **105**, 12909-14.

Li Y.J., Tian F., Chen Z.C., Guan Y.J., He C.M., Yang X.M. & Xie D.H. (2000) Isolation and Identification of cDNA Sequences Differentially Expressed in Laryngeal Carcinoma. *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)* **32**, 153-7.

Lin Q., Gourdon D., Sun C., Holten-Andersen N., Anderson T.H., Waite J.H. & Israelachvili J.N. (2007) Adhesion mechanisms of the mussel foot proteins mfp-1 and mfp-3. *Proc Natl Acad Sci U S A* **104**, 3782-6.

Livak K.J. & Schmittgen T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-8.

Maki J.S., Patel G. & Mitchell R. (1998) Experimental pathogenicity of aeromonas spp. for the zebra mussel, *Dreissena polymorpha*. *Curr Microbiol* **36**, 19-23.

McDowell L.M., Burzio L.A., Waite J.H. & Schaefer J. (1999) Rotational echo double resonance detection of cross-links formed in mussel byssus under high-flow stress. *J Biol Chem* **274**, 20293-5.

McMahon R.F. (1996) The physiological ecology of the zebra mussel, *Dreissena polymorpha*, in North America and Europe. *Am Zoologist* **36**, 339-63.

McVeigh P., Alexander-Bowman S., Veal E., Mousley A., Marks N.J. & Maule A.G. (2008) Neuropeptide-like protein diversity in phylum Nematoda. *Int J Parasitol.*

Miki R., Kadota K., Bono H., Mizuno Y., Tomaru Y., Carninci P., Itoh M., Shibata K., Kawai J., Konno H., Watanabe S., Sato K., Tokusumi Y., Kikuchi N., Ishii Y., Hamaguchi Y., Nishizuka I., Goto H., Nitanda H., Satomi S., Yoshiki A., Kusakabe M., DeRisi J.L., Eisen M.B., Iyer V.R., Brown P.O., Muramatsu M., Shimada H., Okazaki Y. & Hayashizaki Y. (2001) Delineating developmental and metabolic pathways in vivo by expression profiling using the RIKEN set of 18,816 full-length enriched mouse cDNA arrays. *Proc Natl Acad Sci U S A* **98**, 2199-204.

Mitta G., Vandenbulcke F., Hubert F., Salzet M. & Roch P. (2000a) Involvement of mytilins in mussel antimicrobial defense. *J Biol Chem* **275**, 12954-62.

Mitta G., Vandenbulcke F., Noel T., Romestand B., Beauvillain J.C., Salzet M. & Roch P. (2000b) Differential distribution and defence involvement of antimicrobial peptides in mussel. *J Cell Sci* **113 (Pt 15)**, 2759-69.

Molloy D.P. (1998) The potential of using biological control technologies in the management of *Dreissena*. In: *the Eighth International Zebra Mussel and Other Nuisance Species Conference*. California Sea Grant, Sacramento, California.

Molloy D.P. (2002) Biological control of zebra mussels. In: *Third California Conference on Biological Control*, pp. 86-94, University of California, Davis, CA.

Montgomery D.C. (2005) *Design and Analysis of Experiments*. Wiley, Hoboken, NJ.

Morton B. (1969) Studies on the biology of *Dreissena polymorpha* pall: I. General anatomy and morphology. *J. Mollus. Stud.* **38**, 301-21.

Morton B. (1993) The anatomy of *Dreissena polymorpha* and the evolution and success of the heteromyarian form in the Dreissenoidea. In: *Zebra Mussels: Biology, Impact, and Control* (eds. by Nalepa TF & Schloesser DW), pp. 185-215. Lewis Press, Boca Raton, FL.

Munk C., Wei G., Yang O.O., Waring A.J., Wang W., Hong T., Lehrer R.I., Landau N.R. & Cole A.M. (2003) The theta-defensin, retrocyclin, inhibits HIV-1 entry. *AIDS Res Hum Retroviruses* **19**, 875-81.

Nalepa T.F. & Schloesser D.W. (1993) *Zebra mussels: biology, impact and control.* Lewis Publisher, Boca Raton, Florida.

Narasimhan S., Deponte K., Marcantonio N., Liang X., Royce T.E., Nelson K.F., Booth C.J., Koski B., Anderson J.F., Kantor F. & Fikrig E. (2007) Immunity against *Ixodes scapularis* salivary proteins expressed within 24 hours of attachment thwarts tick feeding and impairs Borrelia transmission. *PLoS ONE* **2**, e451.

Nathoo A.N., Moeller R.A., Westlund B.A. & Hart A.C. (2001) Identification of neuropeptide-like protein gene families in Caenorhabditiselegans and other species. *Proc Natl Acad Sci U S A* **98**, 14000-5.

New-York-Sea-Grant (1994a) "Agency Activities." *Dreissena*! , pp. 5(3), 1-2. New York Sea Grant, Zebra Mussel Clearinghouse, 250 Hartwell Hall, SUNY College at Brockport, Brockport, NY 14420-2928.

New-York-Sea-Grant (1994b) *Dreissena polymorpha* information review. pp. 5(1), 14-5. New York Sea Grant, Zebra Mussel Clearinghouse, 250 Hartwell Hall, SUNY College at Brockport, Brockport, NY 14420-2928.

Nguyen D.V., Arpat A.B., Wang N. & Carroll R.J. (2002) DNA microarray experiments: biological and technological aspects. *Biometrics* **58**, 701-17.

Oliver L.M. & Fisher W.S. (1995) Comparative form and function of oyster *Crassostrea virginica* hemocytes from Chesapeake bay (Virginia) and Apalachicola Bay (Florida). *Dis Aquat Org* **22**, 217-25.

Ouellette A.J., Satchell D.P., Hsieh M.M., Hagen S.J. & Selsted M.E. (2000) Characterization of luminal paneth cell alpha-defensins in mouse small intestine. Attenuated antimicrobial activities of peptides with truncated amino termini. *J Biol Chem* **275**, 33969-73.

Papov V.V., Diamond T.V., Biemann K. & Waite J.H. (1995) Hydroxyarginine-containing polyphenolic proteins in the adhesive plaques of the marine mussel *Mytilus edulis*. *J Biol Chem* **270**, 20183-92.

Pathy D.A. & Mackie G.L. (1993) Comparative shell morphology of *Dreissena polymorpha*, *Mytilopsis leucophyta*, and the "quagga" mussel (Bivalvia: Dreissenidae) in North America. *Can J Zool* **71**, 1012-23.

Pawitan Y., Michiels S., Koscielny S., Gusnanto A. & Ploner A. (2005) False discovery rate, sensitivity and sample size for microarray studies. *Bioinformatics* **21**, 3017-24.

Pihlajaniemi T. & Tamminen M. (1990) The alpha 1 chain of type XIII collagen consists of three collagenous and four noncollagenous domains, and its primary transcript undergoes complex alternative splicing. *J Biol Chem* **265**, 16922-8.

Pipe R.K. (1990) Hydrolytic enzymes associated with the granular haemocytes of the marine mussel *Mytilus edulis*. *Histochem J* **22**, 595-603.

Pipe R.K. (1992) Generation of reactive oxygen metabolites by the haemocytes of the mussel Mytilus edulis. *Dev Comp Immunol* **16**, 111-22.

Pipe R.K., Farley S.R. & Coles J.A. (1997) The separation and characterisation of haemocytes from the mussel Mytilus edulis. *Cell Tissue Res* **289**, 537-45.

Price H.A. (1983) Structure and formation of the byssus complex in *Mytilus* (Mollusca, Bivalvia). *J Moll Stud.* **49**, 9-17.

Pritchard C.C. & Nelson P.S. (2008) Gene expression profiling in the developing prostate. *Differentiation* **76**, 624-40.

Radhakrishnan Y., Fares M.A., French F.S. & Hall S.H. (2007) Comparative genomic analysis of a mammalian beta-defensin gene cluster. *Physiol Genomics* **30**, 213-22.

Raikow D.F. (2004) Food web interactions between larval bluegill (*Lepomis macrochirus*) and exotic zebra mussels (*Dreissena polymorpha*). *Can. J. Fish. Aquat. Sci.* **61**, 497-504.

Rajagopal S., Van Der Velde G. & Jenner H.A. (2002) Does status of attachment influence survival time of zebra mussel, *Dreissena polymorpha*, exposed to chlorination? *Environ Toxicol Chem* **21**, 342-6.

Rajagopal S., van der Velde G., van der Gaag M. & Jenner H.A. (2005) Byssal detachment underestimates tolerance of mussels to toxic compounds. *Mar Pollut Bull* **50**, 20-9.

Renwrantz L. & Stahmer A. (1983) Opsonizing properties of an isolated hemolymph agglutinin and demonstration of lectin-like recognition molecules at the surface of hemocytes from *Mytilus edulis*. *J Comp Physiol* **149**, 535-46.

Richards O.W. (1929) The conduction of the nervous impulse through the pedal ganglion of *Mytilus Biol Bull* **56**, 32-40.

Roberts L. (1990) Zebra mussel invasion threatens U.S. waters: damage estimates soar into the billions for the zebra mussel, just one of many invaders entering U.S. waters via ballast water. *Science* **249**, 1370-2.

Roch P. (1999) Defense mechanisms and disease prevention in farmed marine invertebrates. *Aquaculture* **172**, 125-45.

Ross J. (1996) Control of messenger RNA stability in higher eukaryotes. *Trends Genet* **12**, 171-5.

Rzepecki L.M. & Waite J.H. (1993a) The byssus of the zebra mussel, *Dreissena polymorpha*. I: Morphology and in situ protein processing during maturation. *Mol Mar Biol Biotechnol* **2**, 255-66.

Rzepecki L.M. & Waite J.H. (1993b) The byssus of the zebra mussel, *Dreissena polymorpha*. II: Structure and polymorphism of byssal polyphenolic protein families. *Mol Mar Biol Biotechnol* **2**, 267-79.

Schloesser D.W., Nalepa T.F. & Mackie G.L. (1996) Zebra mussel infestation of unionid bivalves (Unionidae) in North America. *Am Zoologist* **36**, 300-10.

Seo J.K., Crawford J.M., Stone K.L. & Noga E.J. (2005) Purification of a novel arthropod defensin from the American oyster, *Crassostrea virginica*. *Biochem Biophys Res Commun* **338**, 1998-2004.

Shoemaker D.D., Schadt E.E., Armour C.D., He Y.D., Garrett-Engele P., McDonagh P.D., Loerch P.M., Leonardson A., Lum P.Y., Cavet G., Wu L.F., Altschuler S.J., Edwards S., King J., Tsang J.S., Schimmack G., Schelter J.M., Koch J., Ziman M., Marton M.J., Li B., Cundiff P., Ward T., Castle J., Krolewski M., Meyer M.R., Mao M., Burchard J., Kidd M.J., Dai H., Phillips J.W., Linsley P.S., Stoughton R., Scherer S. & Boguski M.S. (2001) Experimental annotation of the human genome using microarray technology. *Nature* **409**, 922-7.

Skubinna J.P., Coon T.G. & Batterson T.R. (1995) Increased abundance and depth of submersed macrophytes in response to decreased turbidity in Saginaw Bay, Michigan. *J Great Lake Res* **21**, 476-88.

Smith A.M. & Callow J.A. (2006) *Biological adhesives*. Springer, Berlin, Germany.

Smyth G.K. (2005) Limma: linear models for microarray data. In: *Bioinformatics and Computational Biology Solutions using R and Bioconductor* (eds. by Gentleman R, Carey V, Dudoit S, Irizarry R & Huber W), pp. 397 - 420. Springer, New York.

Snyder F.L., Hilgendorf M.B. & Garton D.W. (1997) Zebra Mussels in North America: The invasion and its implications. Ohio Sea Grant, Ohio State University, Columbus, OH.

Soares M.B., Bonaldo M.F., Jelene P., Su L., Lawton L. & Efstratiadis A. (1994) Construction and characterization of a normalized cDNA library. *Proc Natl Acad Sci U S A* **91**, 9228-32.

Song L., Xu W., Li C., Li H., Wu L., Xiang J. & Guo X. (2006) Development of expressed sequence tags from the bay scallop, *Argopecten irradians irradians*. *Mar Biotechnol (NY)* **8**, 161-9.

Stears R.L., Martinsky T. & Schena M. (2003) Trends in microarray analysis. *Nat Med* **9**, 140-5.

Stothard P. (2000) The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques* **28,** 1102, 4.

Sturn A., Quackenbush J. & Trajanoski Z. (2002) Genesis: cluster analysis of microarray data. *Bioinformatics* **18,** 207-8.

Suh Y.J., Cho S.A., Shim J.H., Yook Y.J., Yoo K.H., Kim J.H., Park E.Y., Noh J.Y., Lee S.H., Yang M.H., Jeong H.S. & Park J.H. (2008) Gene discovery analysis from mouse embryonic stem cells based on time course microarray data. *Mol Cells* **26,** 338-43.

Sun C. & Waite J.H. (2005) Mapping chemical gradients within and along a fibrous structural tissue, mussel byssal threads. *J Biol Chem* **280,** 39332-6.

Tafalla C., Gomez-Leon J., Novoa B. & Figueras A. (2003) Nitric oxide production by carpet shell clam (*Ruditapes decussatus*) hemocytes. *Dev Comp Immunol* **27,** 197-205.

Tamura K., Dudley J., Nei M. & Kumar S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24,** 1596-9.

Tempelman R.J. (2005) Assessing statistical precision, power, and robustness of alternative experimental designs for two color microarray platforms based on mixed effects models. *Vet Immunol Immunopathol* **105,** 175-86.

Thompson J.D., Higgins D.G. & Gibson T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22,** 4673-80.

Tiscar P.G. & Mosca F. (2004) Defense mechanisms in farmed marine molluscs. *Vet Res Commun* **28 Suppl 1,** 57-62.

Toomey M.B., Mccabe D. & Marsden J.E. (2002) Factors affecting the movement of adult zebra mussels (*Dreissena polymorpha*). *J. N. Am. Benthol. Soc.* **21,** 468-75.

Villalva C., Trempat P., Zenou R.C., Delsol G. & Brousset P. (2001) Gene expression profiling by suppression subtractive hybridization (SSH): an example for its application to the study of lymphomas. *Bull Cancer* **88,** 315-9.

Waite J.H. (1983a) Adhesion in byssally attached bivalves. *Biol Rev* **58**, 209-31.

Waite J.H. (1983b) Evidence for a repeating 3,4-dihydroxyphenylalanine- and hydroxyproline-containing decapeptide in the adhesive protein of the mussel, *Mytilus edulis* L. *J Biol Chem* **258**, 2911-5.

Waite J.H. (1987) Nature's underwater adhesive specialist. *Int J Adhes Adhes* **7**, 9-14.

Waite J.H. (1992) The formation of mussel byssus: anatomy of a natural manufacturing process. *Results Probl Cell Differ* **19**, 27-54.

Waite J.H. (2002) Adhesion à la Moule. *Integr Comp Biol* **42**, 1172-80.

Waite J.H., Andersen N.H., Jewhurst S. & Sun C. (2005) Mussel adhesion: finding the tricks worth mimicking. *J Adhes* **81**, 297-317.

Waite J.H. & Qin X. (2001) Polyphosphoprotein from the adhesive pads of *Mytilus edulis*. *Biochemistry* **40**, 2887-93.

Waite J.H., Qin X.X. & Coyne K.J. (1998) The peculiar collagens of mussel byssus. *Matrix Biol* **17**, 93-106.

Walker J.M. (2002) SDS polyacrylamide gel electrophoresis of proteins. In: *The protein protocols handbook* (ed. by Walker JM), pp. 61-7. Humana Press, Totowa, NJ.

Wang L., Song L., Zhao J., Qiu L., Zhang H., Xu W., Li H., Li C., Wu L. & Guo X. (2009) Expressed sequence tags from the zhikong scallop (*Chlamys farreri*): discovery and annotation of host-defense genes. *Fish Shellfish Immunol* **26**, 744-50.

Wang W., Cole A.M., Hong T., Waring A.J. & Lehrer R.I. (2003) Retrocyclin, an antiretroviral theta-defensin, is a lectin. *J Immunol* **170**, 4708-16.

Wang X. & Feuerstein G.Z. (2000) Suppression subtractive hybridisation: application in the discovery of novel pharmacological targets. *Pharmacogenomics* **1**, 101-8.

Warner S.C. & Waite J.H. (1999) Expression of multiple forms of an adhesive plaque protein in an individual mussel, *Mytilus edulis*. . *Mar Biol* **134**, 729-34.

Wettenhall J.M. & Smyth G.K. (2004) limmaGUI: a graphical user interface for linear modeling of microarray data. *Bioinformatics* **20**, 3705-6.

Wiens M., Koziol C., Hassanein H.M., Muller I.M. & Muller W.E. (1999) A homolog of the putative tumor suppressor QM in the sponge *Suberites domuncula*: downregulation during the transition from immortal to mortal (apoptotic) cells. *Tissue Cell* **31**, 163-9.

Wit E., Nobile A. & Khanin R. (2005) Near-optimal designs for dual channel microarray studies. *J Royal Stat Soc, C* **54**, 817-30.

Wu C.-F. & Hamada M. (2000) *Experiments: Planning, Analysis, and Parameter Design Optimization*. Wiley, New York.

Xu W. & Faisal M. (2007) Matrilin-like molecules produced by circulating hemocytes of the zebra mussel (Dreissena polymorpha) upon stimulation. *Dev Comp Immunol* **31**, 1205-10.

Xu W. & Faisal M. (2008) Putative identification of expressed genes associated with attachment of the zebra mussel (*Dreissena polymorpha*). *Biofouling* **24**, 157-61.

Xu W. & Faisal M. (2009a) Development of a cDNA microarray of zebra mussel (*Dreissena polymorpha*) foot and its use in understanding the early stage of underwater adhesion. *Gene* **436**, 71-80.

Xu W. & Faisal M. (2009b) Identification of the molecules involved in zebra mussel (*Dreissena polymorpha*) hemocytes host defense. *Comp Biochem Physiol B Biochem Mol Biol* **154**, 143-9.

Yang D., Chertov O., Bykovskaia S.N., Chen Q., Buffo M.J., Shogan J., Anderson M., Schroder J.M., Wang J.M., Howard O.M. & Oppenheim J.J. (1999) Beta-defensins: linking innate and adaptive immunity through dendritic and T cell CCR6. *Science* **286**, 525-8.

Yang Y.S., Mitta G., Chavanieu A., Calas B., Sanchez J.F., Roch P. & Aumelas A. (2000) Solution structure and activity of the synthetic four-disulfide bond Mediterranean mussel defensin (MGD-1). *Biochemistry* **39**, 14436-47.

Yano M., Nagai K., Morimoto K. & Miyamoto H. (2006) Shematrin: a family of glycine-rich structural proteins in the shell of the pearl oyster *Pinctada fucata*. *Comp Biochem Physiol B Biochem Mol Biol* **144**, 254-62.

Yu M. & Deming T.J. (1998) Synthetic Polypeptide Mimics of Marine Adhesives. *Macromolecules* **31**, 4739-45.

Zdobnov E.M. & Apweiler R. (2001) InterProScan--an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* **17**, 847-8.

Zhao H., Robertson N.B., Jewhurst S.A. & Waite J.H. (2006) Probing the adhesive footprints of Mytilus californianus byssus. *J Biol Chem* **281**, 11090-6.

Zhao H. & Waite J.H. (2006) Linking adhesive and structural proteins in the attachment plaque of Mytilus californianus. *J Biol Chem* **281**, 26150-8.

Zhao J., Song L., Li C., Ni D., Wu L., Zhu L., Wang H. & Xu W. (2007) Molecular cloning, expression of a big defensin gene from bay scallop Argopecten irradians and the antimicrobial activity of its recombinant protein. *Mol Immunol* **44**, 360-8.

Zhu G., Spellman P.T., Volpe T., Brown P.O., Botstein D., Davis T.N. & Futcher B. (2000) Two yeast forkhead genes regulate the cell cycle and pseudohyphal growth. *Nature* **406**, 90-4.

Zou J., Rodriguez-Zas S., Aldea M., Li M., Zhu J., Gonzalez D.O., Vodkin L.O., DeLucia E. & Clough S.J. (2005) Expression profiling soybean response to Pseudomonas syringae reveals new defense-related genes and rapid HR-specific downregulation of photosynthesis. *Mol Plant Microbe Interact* **18**, 1161-74.