AN ABSTRACT OF THE DISSERTATION OF

Chao Chen for the degree of Doctor of Philosophy in Animal Sciences presented on July 31, 2013.

Title: Quantitative and Molecular Genetics of Larval Susceptibility to *Vibrio tubiashii* in Pacific Oysters (*Crassostrea gigas*)

Abstract approved:

John Killefer, William D. Hohenboken

Vibriosis caused by the bacterial pathogen *Vibrio tubiashii* is one of several factors contributing to mass larval mortalities of the Pacific oyster (*Crassostrea gigas*) in Pacific Northwest shellfish hatcheries in recent years. Genetically improved strains of the Pacific oyster resistant to *V. tubiashii* would be highly desirable. This study investigates quantitative and molecular genetics of resistance to *V. tubiashii* in Pacific oyster larvae that could be used for marker-assisted selection.

In experiment 1, we measured the survival of larvae samples from 30 full-sib families under *V. tubiashii*-free conditions and when exposed to 3 different concentrations of *V. tubiashii*. We found that heritability of survival was surprisingly high but also variable, depending upon *V. tubiashii* concentration. We also found strong genotype by environment interactions, meaning families responded differently to changes in the environments. In experiment 2, we scanned the oyster genome for single nucleotide polymorphisms (SNP) associated with *V. tubiashii* resistance by comparing allele frequencies between two larvae pools: survivors of a laboratory *V. tubiashii* challenge and contemporary unchallenged

controls in each of four full-sib families. In total, we identified 262 SNPs with significantly different allele frequencies between the survivors and the contemporary controls. In experiment 3, we used a tag-base RNA-Seq procedure to quantify transcriptomes of 3 samples for each family: larvae taken at the beginning of the experiment (0-hour control), larvae raised for 12 hours without *V. tubiashii* (12-hour control), and larvae raised for 12 hours with *V. tubiashii* (12-hour treatment). We identified 12 differentially expressed genes between resistant and sensitive families. In addition, we confirmed the observation of Taris et al. (2009) that transcriptional differences between resistant and sensitive families exists prior to a bacterial challenge, supporting the hypothesis that difference in resistance to *V. tubiashii* is due to inherit differences in gene expression rather than differential responses to pathogen challenge.

These experiments provide both insight into the biology of larval resistant to *V. tubiashii* in Pacific oyster, and candidate genes and molecular markers that can be used for marker-assisted selection to develop genetically improved oyster strains resistant to *V. tubiashii*.

Quantitative and Molecular Genetics of Larval Susceptibility to *Vibrio tubiashii* in Pacific Oysters (*Crassostrea gigas*)

by

Chao Chen

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of

the requirements for the

degree of

Doctor of Philosophy

Presented July 31, 2013

Commencement June 2014

Doctor of Philosophy dissertation of Chao Chen presented on July 31, 2013.

APPROVED:

Co-Major Professor, representing Animal Sciences

Co-Major Professor, representing Animal Sciences

Head of the Department of Animal and Rangeland Sciences

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Chao Chen, Author

ACKNOWLEDGEMENTS

This dissertation would not be possible without the help and support from generous people around me. Although this dissertation bears my name, many people have contributed to it.

I would like to express my gratitude and appreciation to Mark Camara and Eli Meyer, for their guidance and support in every aspect of my research.

I would also like to thank my other committee members, John Killefer, William Hohenboken, George Waldbusser, and Shawn Townsend, who have given me advice and help in various ways. I would like to extend my appreciation to Chris Langdon, who has provided me advice and experimental materials, and to Gerd Bobe for his help with scientific writing.

Big thanks to the nice people in Department of Animal and Rangeland Sciences and Hatfield Marine Science Center, especially staff of the Molluscan Broodstock Program, as well as members of the Meyer lab, who made my research a great experience.

Last, but not least, I would like to thank my family for their unconditional support all the time, despite the long distance.

CONTRIBUTION OF AUTHORS

Dr. Mark Camara was involved in the experimental design of Chapter 2, 3 and 4, and the editing of Chapter 2 and 3. Dr. Eli Meyer assisted with data collection and analysis and editing of Chapter 3 and 4.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

Figure Page

LIST OF FIGURES (Continued)

Figure Page

LIST OF TABLES

Table Page

Quantitative and Molecular Genetics of Larval Susceptibility to *Vibrio tubiashii* in Pacific Oysters (*Crassostrea gigas*)

Chapter 1

1. General Introduction

1.1 The Oyster Industry

According to the Pacific Coast Shellfish Growers Association (PCSGA), current shellfish production and sales on the Pacific west coast are approximately 88 million pounds and 117 million US dollars, respectively, of which 73 million pounds (83%) and 73 million US dollars (62%) are attributable to oysters. Shellfish aquaculture provides thousands of family-wage jobs in coastal communities. Locally raised shellfish are also exported and in 2012 generated more than $16 million foreign trade in the coastal states of Washington, Oregon and California, according to the Foreign Agricultural Service, providing substantial economic benefits to coastal communities in the Pacific Northwest.

Pacific oysters were introduced from Japan to the Pacific Northwest in 1903 (Pauley et al. 1988). The commercial culture of Pacific oysters has grown rapidly since then. The Pacific oyster is preferred by the industry over the native Olympia oyster because it is larger and faster growing. However, Pacific oysters have only established natural populations in a few places in the Pacific Northwest because of high flushing rates of most West Coast estuaries and the generally low seawater temperature compared to optimum temperatures for spawning and settlement. There are three sources from which oyster growers can acquire *C. gigas* seed to plant on their farms: 1) imported seed from Japan, 2) naturally recruited seed, and 3) hatchery seed. Before the 1970's, growers relied heavily on imported seed from Japan except during World War II, when this was impossible. However, the price for imported seed

increased dramatically in the 1960's. Natural recruitment of *C. gigas* seed was successful for several years during the 1970's, taking the place of imported seed. However, natural recruitment is erratic and unpredictable. Hatchery production of oyster seed was developed in the 1970's and is more controlled and reliable than natural collection. Consequently, the industry quickly came to rely almost entirely on this technology. Moreover, because hatcheries enable control of spawning and fertilization, they make selective breeding possible. Breeding programs, in turn, provide hatcheries with genetically improved strains whose seed is more competitive than seed collected from natural populations (Clark & Langmo 1979).

1.2 Hatchery Problem and *Vibrio tubiashii*

1.2.1 Larval mortality in hatchery and *V. tubiashii*

In recent years, commercial shellfish hatcheries in the Pacific Northwest have reported an emerging disease that is impacting hatchery production of Pacific oyster larvae and seed. Whiskey Creek Hatchery in Oregon State, which produces almost 80% of all larvae for the West Coast oyster industry, experienced 70-80% larval mortalities in 2007 and 2008, while Taylor Hatchery in Washington State reported 60-70% mortality from 2007 to 2009. Moreover, there has been little natural recruitment for several years in areas where *C. gigas* had previously established naturalized populations (Dumbauld et al. 2011). Most farms rely on hatchery-purchased larvae for grow-out on their farms, and as a result, have been experiencing difficulties due to a shortage of seed for planting.

Larval mortality may be attributable to several factors and their interactions. Initially, losses of oyster larvae were linked to *V. tubiashii* outbreaks in the coastal seawater and

upwelling events (Elston et al. 2008). The upwelled seawater is cooler, and nutrient- and *Vibrio spp.*- enriched. It has been hypothesized that outbreaks of *V. tubiashii* may occur as a result of the warming of upwelled seawater by mixing with warmer-than-normal surface seawater, causing increased mortalities in several marine invertebrates, including abalone (Travers et al. 2009), corals (Vezzulli et al. 2010), and adult oysters (Garnier et al. 2007). Improved seawater treatment system in hatcheries helped in reducing larval mortality, however, pronounced mortalities still occurred in hatcheries in the absence of detectable *V. tubiashii*. Barton et al. (2012) reported a correlation between larval mortality and carbon dioxide levels (CO_2 partial pressure, P_{CO2}) in hatcheries. Mortality caused by high P_{CO2} may provide dead larval tissue for *V. tubiashii* to grow, which in turn may cause even greater larval mortality.

1.2.2 *Vibrio tubiashii*

V. tubiashii was first described by Tubiash et al. (1965; 1970) as a lethal pathogen of larval and juvenile bivalve mollusks, and as the cause of bacillary necrosis. It was designated as *Vibrio tubiashii* by Hada et al. (1984). *Vibrio* spp. are pathogenic to a variety of bivalves (e.g. Pacific oyster, Kumamoto oyster, eastern oyster, geoduck clam). Both larval and adult oysters can be affected, and susceptibility varies with size, smaller, younger animals being more susceptible (Gómez-León et al. 2008).

Toxic effects of *V. tubiashii* have been observed in all ages of larvae from D-stage to metamorphosis, and include loss of velar cilia, exfoliation of velar epithelial cells, failure to swim due to extensive velar damage, and eventual death. These effects are largely attributable to an extracellular metalloprotease produced by *V. tubiashii*, VtpA (Hasegawa et al. 2008),

although other factors are undoubtedly involved. However, recent findings showed that other secreted metalloprotease other than VtpA may cause larval mortality (C. Langdon, pers. comm.). Invasive infections of *V. tubiashii* occur in both laboratory challenges and hatchery epizootics. The pattern of invasive infection is identical to earlier and continuing reports of *Vibrio* spp. infections (Elston & Leibovitz 1980; Elston et al. 1981; Sindermann 1988; Elston et al. 1999). Infection begins with the attachment of bacterial cells to the shell surface followed by progressive growth along the internal shell surface and eventually abscessed infections of the mantle.

1.2.3 The need for genetically improved oyster strains

V. tubiashii seems to be cosmopolitan in coastal areas and highly opportunistic. Outbreaks of *V. tubiashii* in coastal waters are likely to increase in frequency as surface seawater temperature rises as a consequence of global warming, and what used to be "warmer-than-usual" surface seawater becomes the "new normal". Therefore *V. tubiashii* presents a continuing threat to shellfish producing facilities for the foreseeable future.

Although good husbandry practices such as better monitoring and improved water treatment systems can potentially reduce contamination levels and improve larval survival, a multifaceted approach is necessary to permanently resolve the problem and control the impacts of *V. tubiashii* on all susceptible phases of the culture process. Management of diseases often leads to economic losses, and toxins produced by *V. tubiashii* can bypass water treatment systems (Elston et al. 2008). Also, as mentioned above, death caused by other stresses (including high P_{CO2} in seawater) may provide *V. tubiashii* with substrates, which, combined with elevated seawater temperature in hatcheries, promotes growth of *V. tubiashii*

within hatcheries. High P_{CO2} also increases growth of *V. tubiashii* (Dorfmeier 2012), leading to further larval mortality. Therefore, the development of genetically improved oyster strains resistant to *V. tubiashii* would be highly desirable.

1.3 Defense Mechanisms in Oysters

Oysters do not have an antibody-based adaptive immune system like those of vertebrates and rely entirely on an innate immune system to combat pathogens. Shells and mucus serve as the first barrier against pathogens. If these defenses are breached, surface epithelial cells and hemocytes are the two major types of immune cells that defend oysters from pathogens (Schmitt et al. 2012). Although less well-studied, surface epithelial cells express receptors for both self/non-self-recognition and for immune effectors that kill pathogens. Hemocytes in oysters serve various functions including digestion and nutrient transport, wound healing, shell repair, excretion, and internal defense (Cheng 1996). Because oyster circulatory system is open, hemocytes are not confined to the interior of the heart and vessels because the oyster circulatory system is open, and they can be recruited to the site of infection upon injury. Phagocytosis by hemocytes plays a crucial role in immune response, which includes recognition, adhesion, ingestion, destruction and eventual elimination of foreign cells.

Three effectors are involved in the deactivation of foreign cells by hemocytes: reactive oxygen species (ROS, e.g. superoxides and peroxides), hydrolytic enzymes, and antimicrobial peptides and proteins (AMPs) (Schmitt et al. 2012). ROS production results from respiratory burst metabolism (Roch 1999), which is also associated with high levels of programmed cell death or apoptosis (Terahara & Takahashi 2008). For protection from ROS

damage, oysters rely on superoxide dismutase (SOD) and other antioxidant enzymes to detoxify ROS. Apoptosis in mollusks is thought to play an important role in the immune system in that it occurs without causing inflammation and therefore helps to maintain homeostasis (Sokolova 2009). AMPs are another kind of defense effector, including proline-rich AMPs, bactericidal/permeability-increasing (BPI) protein, and defensin-like peptides (defensins and big defensins). Different AMPs can work in synergy at sites of infection, increasing their effectiveness. It is suspected that the Pacific oyster lacks the systemic humoral antimicrobial response in that no increase of AMPs is observed after infection, and synergistic activities of AMPs compensate the low level of AMPs in the Pacific oysters (Schmitt et al. 2012).

Some examples demonstrate that bacteria can avoid being killed by hemocytes by damaging the phagocytes or inhibiting one of the steps in the process of phagocytosis, including changing physical or chemical properties of their surfaces, encapsulation, and preventing respiratory burst by the host (reviewed by: Canesi et al. 2002). Apoptosis can also be prevented by some pathogens that inhibit immune response (reviewed by: Sokolova 2009). A recent study on another oyster pathogen *Vibrio splendidus* showed that it can evade immune defense by preventing acidic vacuole formation and limiting ROS production (Duperthuy et al. 2011).

1.4 Genetic improvement

As mentioned above, genetically improved oyster strains resistant to *V. tubiashii* are needed. Genetic improvement through selective breeding requires the existence of genetic variation for the traits of interest, in this instance the relevant components of the oyster

immune system. Successful selective breeding of the closely-related eastern oyster (*C. virginica*) for resistance to two protozoan diseases that decimated their populations in the Chesapeake Bay (Ragone Calvo et al. 2003), coupled with the finding that this resistance was related to protease inhibitor activity that varied among families (Oliver et al. 2000), suggests that at least some components of innate immunity may be genetically variable in oysters. Furthermore, a more recent study (Gómez-León et al. 2008) found significant differences in the susceptibility of selectively bred strains of eastern oysters to several bacterial pathogens, including strains of *V. tubiashii* isolated from Pacific oysters (RE22 and RE101).

1.4.1 Conventional breeding

Conventional breeding requires knowledge of the genetic relationships among individuals and measurement of their phenotypes. Breeders make use of the resemblance among relatives to infer genetic parameters and estimate breeding values to identify genetically superior candidates, then use this information to increase the frequencies of favorable genes through selection without knowing the identities of the alleles involved.

Breeding progress is influenced by four factors in the so called "breeder's equation":

$$\Delta G = h^2 \sigma_P i / L$$

where ΔG is the genetic gain per unit time, h^2 is the narrow sense heritability of the trait, σ_p is the phenotypic standard deviation of the pertinent trait in the breeding population, i is selection intensity expressed in units of standard deviation from population mean, and L is the generation interval (i.e. length of time for a cycle of selection).

Heritability is the proportion of phenotypic variance attributable to genetic causes. Two types of heritabilities are often used: 1) narrow sense heritability, denoted by h^2, equals

the proportion of the phenotypic variance attributable to additive genetic effects, and 2) broad sense heritability, denoted by H^2, equals the proportion of the phenotypic variance attributable to all genetic effects, including additive effects and the non-additive effects.

To estimate heritability, one must to partition phenotypic variation into its genetic and environmental components. This can be done by analyzing the phenotypic resemblance among relatives with known levels of genetic relatedness. With appropriate mating designs that include both half- and full-sibs, genetic variation can be further partitioned into additive and non-additive components.

If estimates of the heritability and phenotypic variance are available, they can be combined with knowledge of generation interval and selection intensity to calculate the expected genetic gain over a certain period of time. By comparing the rates of estimated genetic gain from different breeding strategies, and with consideration of other factors such as costs and available resources, one can choose the most suitable strategy for genetic improvement.

However, conventional breeding is difficult for traits whose phenotypes are hard/expensive to measure or that cannot be measured non-destructively. Under such circumstances the use of molecular markers in selection, i.e. marker-assisted selection (MAS) is desirable.

1.4.2 Marker-assisted selection (MAS)

Marker assisted selection can greatly improve the efficiency and effectiveness of breeding. MAS relies on “tagging” of those genes affecting the trait of interest by molecular markers closely linked to those genes. While traditional breeding uses phenotypes as an

indicator of the state of the underlying genes, it falls short in situations where marker genotyping is preferable to phenotyping (reviewed in: Dekkers & Hospital 2002). This is especially true for traits whose phenotypes can only be measured under limited conditions. Some phenotypes are expressed in only one gender (e.g. milk production), and some can only be observed at certain developmental stages (e.g. grain yield). Additionally, individuals must be killed to measure some phenotypes. Therefore measurements could not be made on the individuals selected as breeders for the next generation. From the economic point of view, some phenotypes are expensive to measure, and using molecular markers may reduce overall costs. MAS also has the potential to be a better strategy than traditional breeding for traits whose phenotypes are poor indicators of the underlying genes due to low heritability or genotype-by-environment interaction.

For conventional breeding to improve *V. tubiashii* resistance, one would need to spawn adult oysters (which usually kills them) and challenge their offspring with *V. tubiashii* to be able to estimate their breeding values. Having to raise larvae just for progeny testing is expensive, wasteful, and time and labor consuming. The adult oysters being tested are sacrificed and therefore can't be used as selection candidates unless gametes are somehow preserved, for example by freezing (Lannan 1971; Yang et al. 2012). Also, the number of candidate oysters that can be tested is low due to limitations in resources and the fact that tank effects are difficult to control unless tests are well-replicated, requiring large numbers of containers. Moreover, candidate oysters can't be tested until they are sexually mature, and gamete quality is affected by the environment. However, if molecular markers strongly associated with resistance to *V. tubiashii* could be identified, these markers could be used to make selection and breeding decisions. DNA genotyping requires only small amounts of

DNA, therefore animals would not need to be sacrificed, and they could be tested before reaching maturity. Testing and selection at an early stage reduces the cost of maintaining large numbers of oysters and could potentially reduce the generation interval. DNA genotyping is fast and less labor intensive than phenotyping, greatly increasing the number of oysters that can be tested. Furthermore, DNA genotyping is cost effective. Therefore, marker-assisted selection is likely to be preferred for the *V. tubiashii* resistant trait.

Depending on the characteristics of species and traits, breeders can develop breeding strategies that combine molecular score and phenotypic data to maximum gain and/or reduce cost. MAS has been applied in various species, e.g. wheat, perennial ryegrass, cotton, etc. (reviewed in: FAO 2007). To perform MAS for resistance to *V. tubiashii* in oyster larvae, we must first identify marker(s) that are associated with *V. tubiashii* resistance. Two approaches are mainly used for identifying markers: the candidate gene based approach and the linkage based approach.

1.4.3 Marker development for marker-assisted selection

1.4.3.1 Candidate gene based approach

A candidate genes approach usually starts with generating a relatively short list of candidate genes based on all available information on the biochemical and genetic basis of the phenotype of interest, followed by the identification of markers for the selected genes. The markers are then tested for their association with the phenotype within a large, unrelated population. Markers found to be strongly associated with a trait can then be used in marker-assisted selection to alter that trait.

A great deal of prior information of the underlying mechanism of the trait is needed to make informed decisions about which genes are the best candidates. One common way is to compare the gene expression profiles of different groups of individuals under different conditions. Differentially expressed genes are hypothesized to be involved in the molecular process and can then be used as candidates for further marker development.

Several methods can be used to measure gene expression levels and thereby differentiate groups of animals or animals exposed to different conditions. Real-Time Polymerase Chain Reaction (RT-PCR), also called quantitative real-time polymerase chain reaction (qPCR), is widely used for testing single genes. RT-PCR is considered to be the standard for evaluating gene expression levels, and is used by itself as well as to validate results from other methods. The drawback of RT-PCR is that it can measure only one gene at a time, and primers need to be designed for each gene. It is not applicable when the purpose is to look at gene expression levels of a large number of genes or of the entire transcriptome. Other methods such as microarrays and RNA-seq are used to screen large numbers of genes. These techniques will be discussed in section 1.6.1.

1.4.3.2 Linkage based approach

Unlike the candidate gene based approach, the linkage based approach does not require prior understanding of the underlying mechanism of the trait of interest. Quantitative trait loci (QTL) are identified by a systematic scan of the whole genome for markers that are associated with a certain phenotype. The identified QTLs, in turn, can help to provide candidates for the candidate gene based approach.

1.4.4 Using markers for MAS

Markers can be classified into three categories based on the ways in which they can be used in MAS: 1) causative mutations (direct markers), 2) linked markers in population-wide linkage disequilibrium with the QTL (LD markers), and 3) linked markers in population-wide equilibrium with the QTL (LE markers) (Dekkers 2004). Direct markers are the easiest to apply in breeding programs and have the most potential for genetic improvement, but they are difficult to develop because causality is hard and expensive to demonstrate. Most studies for QTL of agriculturally important traits do not go this far. LE markers, on the other hand, are easy to develop. However, they can only be used on the families from which they were developed because linkage phases between markers and QTL differ among families. Also, linkage between LE markers and QTL are likely to break down by recombination after several generations because most LE markers are not close enough to the causal mutations. LD markers are close enough to the causal mutation that they are in population-wide linkage disequilibrium with the QTL, and thus can be used across the population. They are intermediate in terms of the difficulty of development and application.

1.4.5 Types of markers

Any kind of DNA sequence, in principle, can be used as a molecular marker if it meet two requirements: 1) variation(s) exist in the DNA, and 2) the differences can be detected. The best type of marker to use depends on the purpose of the project, resources, whether the markers are already developed, the number of markers and individuals to genotype, the organism of interest and the amount of available genomic information for the organism, etc. In general, three marker types are most widely used: Amplified Fragment

Length Polymorphisms (AFLP), Simple Sequence Repeats or microsatellites (SSR), and Single Nucleotide Polymorphisms (SNP).

AFLPs uses restriction endonucleases and Polymerase Chain Reaction (PCR) to generate fragments of different sizes, genotypes are called based on the presence of fragments of certain sizes (Vos et al. 1995). ALFP is flexible, and can generate a great deal of data without prior knowledge of DNA sequences. However, AFLP markers are dominant, and homozygotes with dominant alleles (AA) and heterozygotes (Aa) cannot be distinguished. Genotyping of AFLP is also labor intensive. SSRs are highly polymorphic and have a large number of alleles at each locus, which is very useful in applications like analysis of parentage (Kirst et al. 2005). However, SSR needs to be developed beforehand, and primers need to be designed for each locus. Null alleles, which are widespread in *C. gigas* (Li et al. 2003; Hedgecock et al. 2004), can also cause problems in calling genotypes. SNPs are a variation at a single nucleotide position and are the most abundant marker in the genome. This allows the construction of high density linkage maps. SNPs may or may not need to be developed beforehand depending on the method for genotyping. Numerous methods for SNP genotyping have been developed using difference in length of products, fluorescent signals, melting temperatures, sequences, etc. (Kim & Misra 2007). Among them, Restriction-site Associated DNA sequencing (RAD-seq) is gaining popularity for genome-wide genotyping. RAD-seq is especially useful in non-model organisms that don't have significant genomic sequences available, or organisms with large genome size. RAD-seq for genome-wide genotyping will be discussed in section 1.6.2.

1.5 Genome of *C. gigas*

The genome of *C. gigas* is highly polymorphic (Sauvage et al. 2007; Zhang et al. 2012). On one hand, this causes the problem of null alleles when using SSR as molecular markers in oysters (Li et al. 2003; Hedgecock et al. 2004). On the other hand, it also provides a good opportunity to develop SNPs. The high polymorphism may be the result of a large number of cell divisions during gamete formation, which generates more mutations (Li et al. 2002). As a consequence of the large number of mutations, oyster populations carry a high genetic load (Launey & Hedgecock 2001).

As an important mollusc species, the genome of *C. gigas* was recently sequenced (Zhang et al. 2012). The estimated genome size is 545 M to 637 M depending on the method used. Currently, the Pacific Oyster Genome Database (OysterDB) conctains 11,969 scaffolds and 28,027 genes. Also available is transcriptome information of different tissues, developmental stages, and under different stress conditions. This information will greatly facilitate research related to the biology and selective breeding of *C. gigas*.

Sequencing of the genome and transcriptome is accomplished mainly by next generation sequencing technologies, which will be discussed in the next section.

1.6. Next generation sequencing

Next generation sequencing (NGS) technology is gaining in popularity in recent years. NGS has fundamentally changed the way biological research is done. In general, samples are processed using protocols that generate pools of (short) DNA fragments (libraries) in various ways, and the next generation sequencing equipment generates hundreds of millions of short sequences. Depending on the way the libraries are prepared, NGS can be used for different applications, including chromatin immunoprecipitation (ChIP), gene

expression, genotyping, DNA methylation, diagnostics, etc. It is becoming the method of choice of many researchers due to the maturation of the technologies and the decreasing cost. I highlight two of the applications that I used in my research in more detail below: transcriptome profiling and genome-wide genotyping using RAD-seq.

1.6.1 Transcriptome profiling

Transcriptome profiling is the measurement of all the transcripts. Currently two methods are widely used for transcriptome profiling: microarray and RNA-seq.

A microarray is a collection of DNA probes placed on a chip. Each spot on the chip contains one kind of probe. For the purpose of transcriptome profiling, the probes represents thousands of genes. By hybridizing DNA made from samples to the chip, the amount of different transcripts in the samples, reflected by the intensity of fluorescent signals, can be determined (Miller & Tang 2009). Microarrays have been the method of choice for many because they are cheap, easy to prepare and have mature methods for data analysis (Reimers 2010). Microarrays yield comparable results to RNA-seq but with a smaller dynamic range (Wang et al. 2009). However, microarrays can only detect genes for which probes have been designed. Novel transcripts, which may be important genes, cannot be detected. Also, prior knowledge is needed to design the microarrays.

In contrast to microarrays, RNA-seq directly sequences transcripts in the sample and therefore can detect novel transcripts without prior knowledge of the transcriptomes. Expression levels of genes are reflected by the (normalized) counts of gene sequences (Dillies et al. 2012). Moreover, RNA-seq is able to detect alternative splicing (Toung et al. 2011), allele specific expression (ASE) (Zhang et al. 2009; Main et al. 2009; Skelly & Johansson

2011) and noncoding RNAs (Huang et al. 2011; Sun et al. 2012), etc. Many researchers are switching from microarrays to RNA-seq despite the higher cost of RNA-seq.

A way to reduce cost for RNA-seq is to use tag sequencing, which only sequences the end of transcripts and therefore greatly reduces the need for sequencing coverage. Although alternative splicing cannot be detected by tag sequencing, one can still identify novel transcripts. Tag sequencing, combined with the ability to multiplex by barcoding samples, is very useful in transcriptome profiling.

1.6.2 Genome-wide genotyping using Restrictive-site Associated DNA sequencing (RAD-seq)

Although NGS is getting cheaper, sequencing the whole genome for genome-wide genotyping is still not feasible for most organisms. It is feasible only in organisms with very small genomes like budding yeast (*Saccharomyces cerevisiae*) (Ehrenreich et al. 2010), whose genome size is only 12.5M. Pacific oyster genome size is at least 545M depending on the estimation method (Zhang et al. 2012), which is relatively small compared to other bivalve species. But it is still not feasible to cover the whole genome, at least not cost-effectively. Therefore, a target enrichment step that only covers a fraction of the genome is needed to prepare libraries for the next generation sequencing (Mamanova et al. 2010). Target enrichment is especially important in estimating allele frequencies where a high depth of coverage of the target genome region is required.

Restriction-site associated DNA (RAD) is an easy yet flexible method for genome-wide genotyping. This method uses restrictive endonucleases to target certain sites in the genome, followed by the amplification (and therefore the enrichment) of the regions adjacent to the cutting sites. RAD-seq has been used for polymorphism identification and genotyping

in various organisms (Miller et al. 2007; Hohenlohe et al. 2010; Bus et al. 2012). A variant of RAD-seq, 2bRAD uses type IIB restriction endonucleases to generate fragments uniform in size. Comparing to RAD-seq, 2bRAD is simpler and more flexible (Wang et al. 2012). One can obtain counts of appearance for each allele in each locus, and infer allele frequencies based on the counts of different alleles. To reduce the sequencing cost further, we can pool many individuals into one sample rather than sequence them individually. Pooling can also save time and labor, especially when the number of individuals is large (Sham et al. 2002). Also, for small organisms that don't have enough DNA for individual sequencing, pooling is the only option. Although genotypes of individuals in pooled samples cannot be inferred, allele frequencies can be estimated and compared between different pools. Samples for pools can be selected groups of individuals (case/control), or from different natural populations, depending on the purposes of the studies.

As far as we know, RAD, including 2bRAD, has not been used for estimating allele frequencies in pooled samples. Comparing allele frequencies between pools using HTS, however, has been used and proven to be effective and capable of locating QTLs if the genome if available (Kim et al. 2011; Parts et al. 2011; Boitard et al. 2012; Mullen et al. 2012; Liu et al. 2012). Because of the uniformity of fragment size in 2bRAD library preparation, 2bRAD is expected to work equally well, if not better than these methods in detecting differences in allele frequencies between pools of samples.

Chapter 2

2. Genetic variation for resistance to bacillary necrosis caused by *Vibrio tubiashii* in larval Pacific oysters (*Crassostrea gigas*)

Chao Chen[1], Mark D. Camara[2*]

[1]Deparment of Animal and Rangeland Sciences, Oregon State University, Corvallis, OR 97331, USA

[2]Cawthron Institute, 98 Halifax Street East, Private Bag 2, Nelson 7042, New Zealand

[*]Corresponding author

KEYWORDS: *Crassostrea gigas*, Pacific oyster, heritability, *Vibrio tubiashii*

2.1 Abstract

Since about 2006, shellfish hatcheries on the west coast of the U.S. have experienced severe mortalities of Pacific oyster (*Crassostrea gigas*) larvae, some of which have been attributed to the pathogenic bacterium *Vibrio tubiashii*. Genetically improved strains of the Pacific oyster resistant to *V. tubiashii* would, therefore, be highly beneficial for the Pacific oyster industry. Because selection-based genetic improvement is effective only if the breeding population harbors heritable genetic variation for relevant traits, the objective of this study is to estimate the broad sense heritability of survival of Pacific oyster larvae when exposed to this pathogen. We measured the survival of larvae samples from 30 full-sib families under *V. tubiashii*-free conditions and when exposed to 3 different concentrations of *V. tubiashii*. We found that heritability of survival was surprisingly high but also variable, depending upon *V. tubiashii* concentration. For examples, heritabilities estimated using a family model on proportion live scale ranged from 0.68 (control) to 0.80 (medium concentration of *V. tubiashii* treatment). In general, heritabilities were higher for low and

medium *V. tubiashii* treatment, and lower for control and high *V. tubiashii* treatments. We also found strong genotype by environment interactions, meaning families responded differently to changes in the environments. Therefore, performance must be evaluated under appropriate conditions for effective selective breeding.

2.2 Introduction

In recent years, commercial shellfish hatcheries in the Pacific Northwest have reported an emerging disease impacting hatchery production of Pacific oyster (*Crassostrea gigas*) larvae and seed. Elston et al. (2008) investigated the causes of this problem and found an association between blooms of *Vibrio tubiashii* during periods of unusually warm surface seawater temperatures and losses of oyster seed in commercial oyster hatcheries and nurseries. *V. tubiashii* has long been recognized as a significant bacterial pathogen of bivalve larvae (Tubiash 1965; 1970) and a causative agent of bacillary necrosis in commercial shellfish hatcheries and nurseries (Brown 1973; 1981; Brown & Losee 1978; Elston & Leibovitz 1980; Elston et al. 1981; Elston 1999). First described by Tubiash and coworkers (1965), *V. tubiashii* is highly pathogenic for larval and juvenile Pacific oysters. Bacterial outbreaks may originate from coastal ocean waters during the summer upwelling season. Elston et al. (2008) found that *V. tubiashii* is common in shellfish hatcheries, including in cultured algal food fed to larvae, ambient air, aeration systems, and even in stock sodium thiosulfate solutions used to dechlorinate hatchery equipment after chlorine sterilization. In addition, field sampling has resulted in consistent isolation of *V. tubiashii* from affected juvenile oysters up to 8mm shell height (R. Elston, pers. comm.).

Laboratory experiments using *V. tubiashii* isolates from Pacific Coast locations and Hawaii have described its toxic effects on larvae including variable degrees of loss of velar cilia, exfoliation of velar epithelial cells, failure to swim due to extensive velar damage and eventually death (Elston et al. 2008). Toxic effects have been observed in all ages of larvae from D-stage to metamorphosis, and invasive infections occur in both laboratory challenges and hatchery epizootics. The pattern of invasive infection is identical to earlier and continuing reports of *Vibrio* spp. infections (Elston and Leibovitz 1980; Elston et al. 1981; Sindermann 1988; Elston 1999). Infection begins with the attachment of bacterial cells to shell surface followed by progressive growth along the internal shell surface and abscessed infections of the mantle.

Currently, there are no simple remedies. Because *V. tubiashii* seems to be cosmopolitan in coastal areas and highly opportunistic, it cannot be eradicated from nature and thus presents a threat to shellfish producing facilities for the foreseeable future. While good husbandry practices such as better monitoring and improved water treatment systems can potentially reduce contamination levels and improve larval survival in the partially controlled hatchery environment, a multifaceted approach is necessary to permanently resolve the problem and control the effects of *V. tubiashii* on all susceptible phases of the culture process. Given the severity of the problems this bacterium can cause, the potential for airborne contamination that completely bypasses water treatment systems and directly infects larval culture tanks, and the infeasibility of eliminating the pathogen in nursery systems and grow-out areas where juvenile seed are in direct contact with untreated coastal waters, the development of genetically improved oyster strains with enhanced resistance to *V. tubiashii*

to augment the development of hatchery decontamination and control procedures would be highly desirable.

Genetic improvement through selective breeding relies upon the existence of heritable genetic variation for relevant traits, possibly including components of the oyster immune system. Successful selective breeding of the closely-related eastern oyster (*C. virginica*) for resistance to two protozoan diseases that have decimated oyster populations in the Chesapeake Bay (Ragone Calvo et al. 2003), coupled with the finding that resistance is related to protease inhibitor activity that varies among families (Oliver et al. 2000) suggests that at least some components of innate immunity may be genetically variable in oysters. Furthermore, a more recent study (Gómez-León et al. 2008) reported significant differences in the susceptibility of selectively bred strains of eastern oysters to several bacterial pathogens, including strains of *V. tubiashii* (RE22 and RE101) isolated from Pacific oysters.

Previously, Taris et al. (2009) used challenge experiments in which adult oysters from Pacific oyster families that had been previously characterized as resistant or sensitive to heat stress were exposed to cultured *V. tubiashii*, then characterized subsequent changes in gene expression patterns using a combination of qualitative transcription profiling of many genes' expression patterns (cDNA-AFLP) and quantitative (RT-qPCR) analyses of many specific genes. These experiments identified a number of genes whose expression patterns were altered in response to *V. tubiashii* exposure and/or showed constitutive differences in their mRNA levels in stress-tolerant vs. stress-sensitive families, thus providing indirect evidence that susceptibility to this pathogen is genetically variable, at least in adults. In this paper, we provided direct evidence that the susceptibility of Pacific oyster larvae to *V.*

tubiashii is genetically variable, and estimated the broad-sense heritability using 30 full-sib families.

2.3 Materials and methods

2.3.1 Biological Materials

We obtained 1-day-old larvae of 30 full-sib families from a routine spawn produced by the Molluscan Broodstock Program (MBP) on May 17, 2011 (Langdon et al. 2003). The selected families were unrelated based on their MBP pedigree information. Larvae from the 30 families were aliquoted in 50ml tubes with 10ml of 25℃ seawater at a density of 10 larvae per ml. Tubes were then transferred to random positions in a 25℃ room.

2.3.2 Bacterial preparation, challenge and sample protocol

V. tubiashii strain RE22 was grown on Luria-Bertani (LB) medium supplemented with 1% sodium chloride (LB-1% NaCl) and 100μg ml-1 of ampicillin (Ap) at 30℃ overnight on a shaker. Before use, log-phase bacterial culture was centrifuged gently and washed with seawater.

V. tubiashii were added to the larvae in appropriate volumes to obtain initial densities of 0, 3×10^4 ml^{-1}, 10^5 ml^{-1}, and 3×10^5 ml^{-1}. Each of these treatments was replicated three times for each of the 30 families. After 24 hours, we added a stock solution of neutral red to each tube to give a final concentration of 1:10,000 (Crippen & Perrier 1974). The larvae were killed by adding two drops of 10% formalin two hours after the addition of neutral red. Tubes were then transferred to 4℃.

All of the larvae in each tube were transferred to a watch glass using a transfer pipet, and the tubes were then checked under a dissection scope to make sure all the larvae were transferred. The watch glass was then photographed under a dissecting microscope. Pictures for all samples were taken within 2 days after staining.

2.3.3 Live-dead determinations

Larvae were categorized as live or dead based on the intensity of red color. Larvae with a color of dark red were scored as live, and pink and white larvae were scored as dead (Figure 2.1). Counts of live and dead larvae for each tube were obtained.

2.3.4 Statistical analyses

Statistical analyses were conducted using ASReml 3 (Gilmour et al. 2009) software for mixed models. To test for significant differences in mean survival among challenge treatments, significant variation among family-specific survival and family by treatment interactions, we analyzed models in which overall variation in larval survival was partitioned into components attributable to the four *V. tubiashii* treatments (a fixed effect) and to variation among families and family x treatment interaction (random effects). We did these analyses on two measurement scales: 1) the observed scale with the response variable being the arcsin-square root transformed proportion of live larvae in each tube and variation among replicate tubes within a family/treatment combination as the residual variance, and 2) the threshold-scale with individual survival scored as a binary variable linked to the implicit continuously distributed liability scale, which we modeled using the standardized probit function. Because threshold scale analyses use binary individual-level data, we included an

additional random effect for replicate tubes nested within family/treatment combinations to our threshold model, and the implicit unit variance of the probit function served as the residual variance. Statistical significance was determined for the fixed effect of *V. tubiashii* treatment using the F-ratios and P-values estimated by ASReml, but the tests for significant variation attributable to families and family-by-treatment interaction were tested using single degree of freedom likelihood ratio tests comparing the full model and restricted models from which we dropped each of these factors separately.

The linear model for the analysis on the transformed proportion-scale was:

$$y_{ijk} = \mu + t_i + f_j + (ft)_{ij} + e_{ijk}$$

where y_{ijk} is the dependent variable (transformed proportion live), μ is the overall mean, t_i is the fixed treatment effect , f_j is the random family effect , $(ft)_{ij}$ is the interaction between family and treatment, and e_{ijk} is the residual error.

The linear model for the analysis on the probit-scale was:

$$y_{ijkl} = \mu + t_i + f_j + r_{ijk} + (ft)_{ij} + e_{ijkl}$$

where y_{ijk} is the dependent variable, μ is the overall mean, t_i is the fixed treatment effect , f_j is the random family effect , r_{ijk} is the effect of replicate tubes nested with family/treatment combination, $(ft)_{ij}$ is the interaction between families and treatments, and e_{ijk} is the residual error.

Because both of these analyses found highly significant interaction (see RESULTS), and to explore these more thoroughly, we next performed separate analyses within each level of the *V. tubiashii* treatment and estimated the broad sense heritability for larval survival within each concentration, again at both the observed scale of proportion of surviving larvae and the probit scale. We did these analyses first using models that partitioned the variance in

the arcsin-square root transformed proportion of survivors among the full-sib families and treated this variance component as a direct estimate of the broad sense genetic variance rather than of half the genetic variance and the ratio of this variance component to the total phenotypic variance as the broad sense heritability. We took this approach rather than the normal approach of doubling the variance among full-sibs to estimate H^2 because in contrast to pairs of full-sib individuals that share half their genomes, replicate tubes from the same full-sib family are, in aggregate, genetically identical. We also analyzed the proportion data using a pedigree-based animal model which produces a direct estimate of the genetic component of variance and thereby obtained another estimate of the broad sense heritability as the ratio of the animal component of variance to the total phenotypic variance.

In addition, similar to the overall model above, we analyzed survival as a threshold trait using a probit link function and included replicates nested within families in the model to obtain heritability estimates on the threshold scale. These threshold-scale heritabilities and their standard errors, however, must be re-scaled to be comparable to those on the proportion scale, and we used the formula developed by Dempster and Lerner (1950) for this purpose:

$$H^2_{obs} = H^2_{thr}\frac{p(1-p)}{z^2}$$

where p is the mean proportion of surviving larvae in the reference population and z is the ordinate of the standardized normal distribution corresponding to a probability of p. Standard errors are re-scaled using the same factor.

The family model for the analysis on the transformed proportion-scale was:

$$y_{ij} = \mu + f_i + e_{ij}$$

where y_{ij} is the dependent variable, μ is the overall mean, f_i is the family effect as a random effect, and e_{ij} is the residual error.

The family model for the analysis on the probit-scale was:

$$y_{ijk} = \mu + f_i + r_{ij} + e_{ijk}$$

where y_{ijk} is the dependent variable, μ is the overall mean, f_j is the family effect as a random effect, r_{ij} is the replicate tubes nested with families, and e_{ijk} is the residual error.

The animal model for the analysis on the probit-scale was:

$$y_i = \mu + a_i + e_i$$

where y_i is the dependent variable, μ is the overall mean, a_i is individual animal value as a random effect, and e_i is the residual error.

Finally, we estimated the pairwise broad sense genetic correlation between all combinations of the *V. tubiashii* treatments as the correlation between the family-specific best linear unbiased estimators (BLUE) of the predicted values of the proportion of larvae surviving. For this analysis, we used the predicted values for each family produced when we analyzed each treatment separately for variation among families and estimated the correlations between them using PROC CORR in the SAS statistical package (SAS Institute Inc. 2009).

2.4 Results

Mean survival of each family under each condition is shown in Figure 2.2. In general, survival declined with the increasing *V. tubiashii* concentration. Some families are more resistant than others in that their survival remained fairly constant as *V. tubiashii* concentration increased up until the highest concentration, where they still had from 30% to 50% survival. Other families, in contrast, did well under the control environment but experienced high mortalities if *V. tubiashii* was present at all, and had nearly 100% mortality

under the high *V. tubiashii* concentration. ANOVA shows that there is a significant family effect ($p < 0.001$), treatment effect ($p < 0.0001$), and family × treatment interaction ($p < 0.0001$) on both the proportion live scale and the probit scale (Table 2.1).

Estimated broad sense heritabilities are shown in Table 2.2. Overall, heritabilities are surprisingly high, with some differences among treatments. Heritabilities are generally higher for the low and medium *V. tubiashii* treatment, and lowest in the control and high *V. tubiashii* treatments.

Figure 2.3 shows the correlation of ranks of families between pairs of environments. Correlations are higher for similar environments, and are significant except for control vs. high concentration of *V. tubiashii*. However, the five best families under different treatments are quite different (Table 2.3).

2.5 Discussion

In this study, we estimated the heritability of *V. tubiashii* resistance in Pacific oyster larvae using 30 full-sib families, and found the heritabilities are high in general, and that there is significant genotype by environment interactions.

2.5.1 Variability in survival

Larval survival is highly variable among families (Figure 2.2), especially under medium and high *V. tubiashii* concentrations. There is less family-level variation in the extreme conditions (control and high concentration of *V. tubiashii*) because family survival in general was pushed to the extremes. We can also categorize the families with low survival in the presence of *V. tubiashii* into two groups: families weak overall and families sensitive to *V.*

tubiashii. The "weak overall" families exhibit lower survival than other families even in the control group, indicating they may be weak in the first place and could be vulnerable to all kinds of stresses, whereas the "sensitive" families have medium to high survival under the control condition but very low survival in the presence of *V. tubiashii*, indicating they are sensitive specifically to this pathogen. The two kinds of families may represent two different mechanisms of sensitivity to *V. tubiashii*. A deeper study of the physiological and metabolic differences between these two phenotypes may be helpful in future research to look at the molecular mechanism of the trait by reducing potential confounding factors.

2.5.2 Estimation of heritabilities

In the family models, we estimated the genetic variance as 1 x the family variance rather than 2 x the family variance based on the idea that the genetic correlation between tubes equals 1 in that replicate tubes are more like "clones". Under the probit scale, the heritabilities in the family model are almost exactly double the heritabilities in the animal model in the three *V. tubiashii* treatments. This is not the case for the control environment, where heritabilities of the family and animal models are similar.

We attribute the differences between the family models and the animal models to the difference in model assumptions. In the family models, no assumption on gene action is made. The animal models, on the other hand, assume additive effects only. The results of the animal models, therefore, would be inaccurate in the presence of large non-additive genetic effects, resulting in the large differences in estimated heritabilities between the family models and animal models within the treatments. This implies the importance of non-additive gene action (dominance and/or epistasis) for *V. tubiashii* resistance. High levels of dominance

variance are common in traits related to fitness (Crnokrak & Roff 1995). In contrary to heritability estimates within treatments, the heritability estimated from the animal model is very close to the heritability estimated from the family model in the control, indicating that gene action in the absence of *V. tubiashii* challenge is mainly additive. This is different from the treatments because they are essentially two different traits (survival with/without the presence of *V. tubiashii*).

Overall, the heritabilities are very high, meaning there is substantial genetic variation in susceptibility to *V. tubiashii*. Two factors may cause the high heritabilities: 1) the trait has large genetic variance, and 2) the environmental variance is low. In our short term challenge (1 day) with well controlled environments, the environmental variance might be lower than that for most life-history traits, especially for the group treated with the high concentration of *V. tubiashii*, where most tubes had near-zero survival.

Because broad-sense heritability estimates the proportion of the phenotypic variance attributable to genetic effects, the high broad-sense heritability of *V. tubiashii* resistance provides a good opportunity to develop genetically improved lines that are resistant to *V. tubiashii* using line breeding, which utilize both additive genetic effect and non-additive genetic effect.

2.5.3 Genotype by environment interaction

Our data shows that there is significant genotype by environment interaction ($p < 0.0001$). In other words, families rank differently in different environmental conditions. Because of genotype by environment interactions, family performance should be tested under different conditions. Selection of individuals for breeding the next generation can be based on

an index that incorporates the performances under all conditions. Another possible option is to test families in the most common commercial environment. Careful selection of the most common environment would therefore be important. As we mentioned above, outbreaks of *V. tubiashii* may be more frequent and severe in the future because of global warming. Trends in the change of seawater conditions need to be considered when choosing the test environment.

2.6 Conclusion

In our study, we used 30 full-sib families to assess genetic variation of larval resistance to *V. tubiashii* in *C. gigas*. We found high heritabilities and suggestions of important dominance genetic variance, indicating rapid improvement is. Genotype by environment interactions suggest that family performance should be tested under multiple environments.

2.7 Acknowledgements

Chen thanks financial support from the China Scholarships Council (CSC). We thank Molluscan Broodstock Program (MBP) staff for providing oyster larvae, and Dacey Mercer and Yang Lu for their help in sample collection.

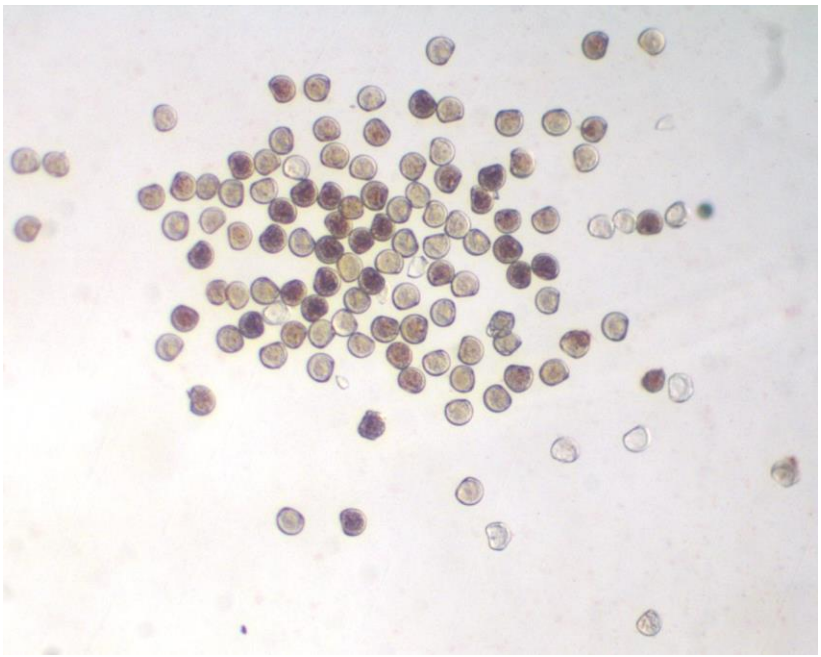

Figure 2.1. Larvae from one of the tubes. Larvae alive at the time of sample were deep red; larvae dead were white to pink.

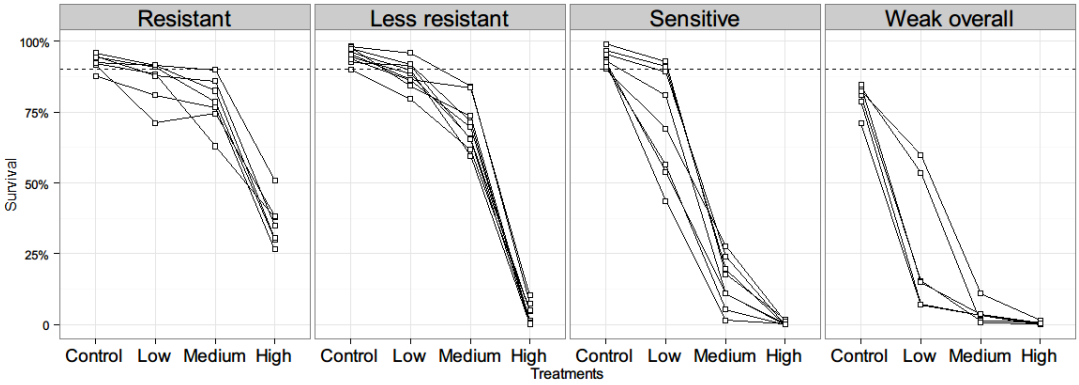

Fig 2.2. Mean family survival. Each line represents the survival of one family under the four treatments. Families were categorized into four groups and graphed separately for clarity: 1) resistant families whose survival is higher than 50% at medium concentration and 20% at high concentration, 2) less resistant families whose survival is higher than 50% at medium concentration but lower than 20% at high concentration, 3) sensitive families whose survival is higher than 90% under control condition but lower than 50% at medium concentration, and 4) overall weak families whose survival is lower than 80% under control condition. The dashed line represents 90% survival.

Table 2.1. Hypothesis tests for family models.

		Measurement Scale			
		Proportion live		Probit	
Model Effect	**df**	**Test statistic**	**p-value**	**Test statistic**	**p-value**
Fixed Effects (F-ratio)					
V. tubiashii Treatment	3	148.5	< 0.001	151.7	< 0.001
Random Effects (Likelihood ratio)					
Family	1	33.5	< 0.0001	26.7	< 0.0001
Interaction	1	126.2	< 0.0001	97.0	< 0.0001

Table 2.2. Estimated broad sense heritabilities.

Treatment	**Model**	**Scale**	**Heritability**	**Stand Error**
Control	Family	Proportion Live	0.68	0.08
		$\sin^{-1}$sqrt(Proportion)	0.65	0.09
		Re-scaled Probit	0.34	0.19
	Animal	Re-scaled Probit	0.35	0.08
Low	Family	Proportion Live	0.79	0.06
		$\sin^{-1}$sqrt(Proportion)	0.80	0.06
		Re-scaled Probit	0.69	0.12
	Animal	Re-scaled Probit	0.32	0.05
Medium	Family	Proportion Live	0.80	0.05
		$\sin^{-1}$sqrt(Proportion)	0.82	0.05
		Re-scaled Probit	0.82	0.11
	Animal	Re-scaled Probit	0.34	0.04
High	Family	Proportion Live	0.64	0.09
		$\sin^{-1}$sqrt(Proportion)	0.76	0.06
		Re-scaled Probit	1.22	0.22
	Animal	Re-scaled Probit	0.58	0.09

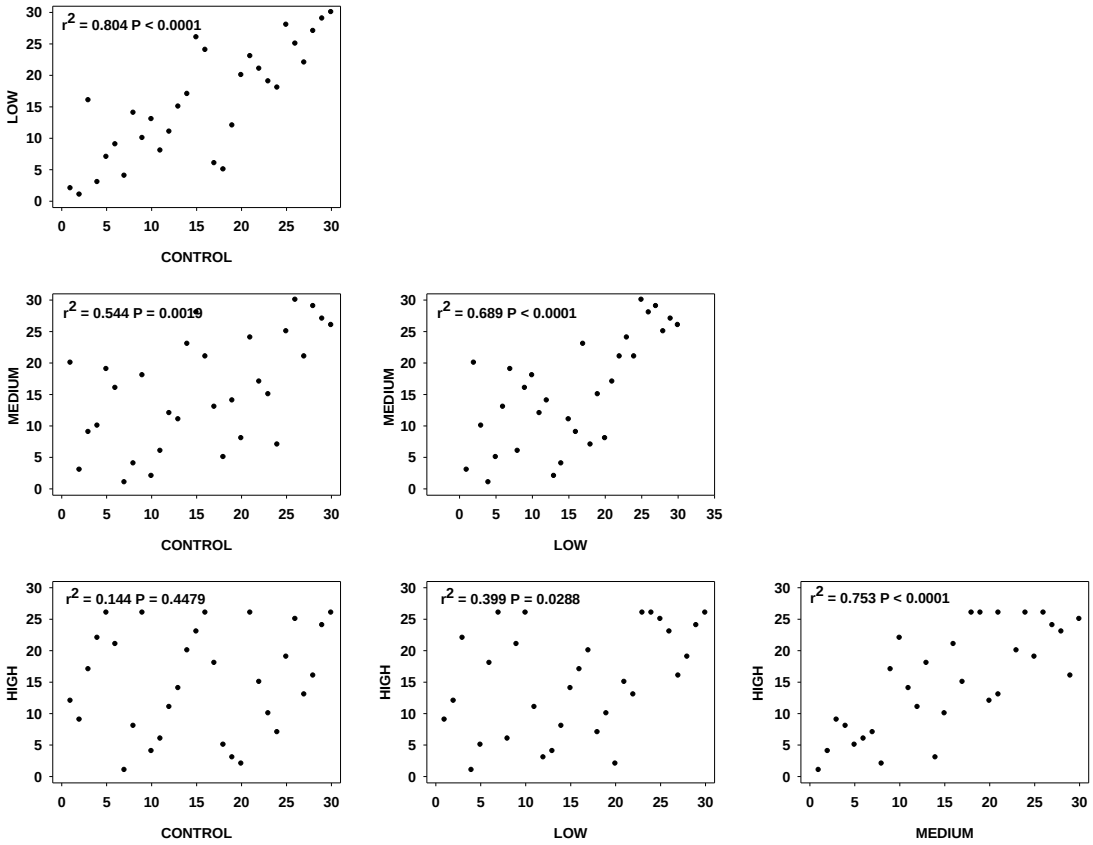

Fig 2.3. Pairwise correlation of family ranks. Vertical and horizontal axis are ranks of families under different treatments.

Table 2.3. Five best families in each treatment.

	Treatments			
Rank	**Control**	**Low**	**Medium**	**High**
1	Family 3	Family 1	Family 28	Family 28
2	Family 1	Family 3	Family 11	Family 26
3	Family 15	Family 18	Family 1	Family 8
4	Family 18	Family 28	Family 17	Family 11
5	Family 21	Family 30	Family 30	Family 30

Chapter 3

3. Genome-wide bulk segregant analysis of allele frequencies in the Pacific oyster (*Crassostrea gigas*) using 2b-RAD reveals markers linked to larval resistance to *Vibrio tubiashii*

Chao Chen[1], Eli Meyer[2], Mark Camara[3]

1 Department of Animal and Rangeland Sciences, Oregon State University, Corvallis, OR 97331, USA

2 Department of Zoology, Oregon State University, Corvallis, OR 97331, USA

3 Cawthron Institute, 98 Halifax Street East, Private Bag 2, Nelson 7042, New Zealand

KEYWORDS: *Crassostra gigas*, Pacific oyster, *Vibrio tubiashii*, RAD, allele frequency

3.1 Abstract

Vibriosis caused by the bacterial pathogen *Vibrio tubiashii* is one of several factors contributing to mass larval mortalities of the Pacific oyster (*Crassostrea gigas*) at Pacific Northwest shellfish hatcheries in recent years. Genetically improved strains of the Pacific oyster resistant to *V. tubiashii* could help to reduce these mortalities, and if molecular markers with strong associations with resistance were available, it would be possible to use marker-assisted selection to improve this trait. In this study, we scanned the oyster genome for single nucleotide polymorphisms (SNP) associated with *V. tubiashii* resistance by comparing allele frequencies between two larvae pools: survivors of a laboratory *V. tubiashii* challenge and contemporary unchallenged controls in each of four full-sib families. In total, we identified 262 SNPs with significantly different allele frequencies between the survivors and the contemporary controls. With further development, such markers could be used for marker-assisted selection to improve oyster strains.

3.2 Introduction

The Pacific oyster, *Crassostrea gigas*, is a major shellfish aquacultural species. In the Pacific Northwest, current shellfish production and sales are approximately 88 million pounds and 117 million US dollars, of which 73 million pounds (83%) and 73 million US dollars (62%) derive from oysters, according to the Pacific Coast Shellfish Growers Association (PCSGA). Oyster farms in the Pacific Northwest rely heavily on hatcheries for reliable seed supply, but commercial hatcheries have experienced severe larval mortalities in recent years resulting in a shortage of hatchery-produced seed oysters. Elston et al. (2008) observed a correlation between mass mortalities in hatcheries and outbreaks of a re-emerging pathogen, *Vibrio tubiashii*, during periods of unusually warm surface seawater temperature.

Mortalities of oyster larvae are expected to be more frequent and severe in the future due to rising surface seawater temperature associated with global warming, favoring outbreaks of *V. tubiashii*. *V. tubiashii* cannot be eradicated from nature because of its cosmopolitan and opportunistic nature. *V. tubiashii* is found in humid air and can therefore bypass water treatment systems in shellfish hatcheries (Elston et al. 2008). Moreover, juvenile seed are planted on marine farms where they are in direct contact with untreated seawater. As a consequence, genetically improved oyster strains resistant to *V. tubiashii* are highly desirable.

Because the direct measurement of *V. tubiashii* resistance is difficult and complex, marker-assisted selection (MAS) using molecular markers to predict this phenotype has advantages over traditional breeding. However, it requires the development of molecular markers either within, or tightly linked to, the underlying causative genes. A cost-effective

way to detect associations between molecular markers and traits is bulk segregant analysis (BSA), which quantifies differences in allele frequencies between pools of samples (Michelmore et al. 1991).

The objective of this experiment is to identify molecular markers associated with resistance to *V. tubiashii*. We conducted a bulk segregant analysis by comparing allele frequencies between pools of survivors of a *V. tubiashii* challenge and unchallenged control pools in four Pacific oyster families. We used Restriction site-Associated DNA (RAD) using type IIB restriction endonucleases (2b-RAD) (Wang et al. 2012) as a cost-effective way to estimate allele frequencies in > 5,700 loci for each family and identified 262 SNPs associated with *V. tubiashii* resistance. Gene enrichment analysis highlighted some biological processes that were possibly involved.

3.3 Materials and methods

3.3.1 Biological materials

Each of the four full-sib families used in this study were created by first crossing two individuals from outcrossed families from the Molluscan Broodstock Program (MBP) (Langdon et al. 2003) that differed in survival and/or growth rate to produce F_1 progeny in 2005. Family 1 (ARS family # 10) was created by mating a large male from a high-surviving family to a small female from a low-surviving family. Family 2 (ARS family # 2) was created by mating a large male from a high-growth family to a small female from a low-growth family. Family 3 (ARS family # 3) was created by mating a large male from a high-growth family to a small female from a low-surviving family. Family 4 (ARS family # 4) was created by mating a small male from a low-growth family to a large female from a high-growth

family. In 2011, F_1 individuals were transferred to land-based conditioning systems at the Hatfield Marine Science Center in Newport, Oregon. After two weeks of acclimation, seawater temperature was increased 1 ℃ per day to 18 ℃. These F_1 individuals were kept in 18 ℃ seawater until gamete maturation (six weeks or longer). Four F_2 families (families 1, 2, 3 and 4), were produced by mating two individuals from the same F_1 family (Figure 3.1). Samples of mantle tissue from the eight F_1 parents were preserved in 70% ethanol. Crossing and larval culturing were conducted following standard MBP hatchery and nursery protocols (Langdon et al. 2003).

3.3.2 Bacterial preparation and challenge

V. tubiashii strain RE22 was grown on Luria-Bertani (LB) medium supplemented with 1% sodium chloride (LB-1% NaCl) and 100μg ml^{-1} of ampicillin (Ap) at 30 ℃ overnight on a shaker. Before use, the log-phase bacterial culture was centrifuged gently and washed with seawater.

V. tubiashii were added to 1-day-old larvae in appropriate volumes to obtain final bacterial densities of 3×10^5 ml^{-1} in the treatment tanks, whereas no *V. tubiashii* was added to the control tanks. Both the controls and the treatments were done in triplicate for each of the 4 families. Survival in the treatment tanks was monitored (Figure 3.2), and larval samples were collected when mortalities exceeded 50% (24 hours for families 2 and 3, 48 hours for families 1 and 4). Larvae were separately sieved from each tank, washed with distilled water and preserved in 70% ethanol.

3.3.3 Library preparation and read processing

DNA of F_1 parents and their pooled larval progeny were extracted using Qiagen DNeasy Blood & Tissue kit. Sequencing libraries were prepared for F_1 parents and F_2 larval pools using AlfI as the restrictive endonuclease. Libraries were prepared and reads were processed following the protocol of Wang et al. (2012). All sequencing was done on Applied Biosystems SOLiD™ system at University of Texas at Austin. We extracted AlfI cutting sites from the published oyster genome (Zhang et al. 2012), and used the extracted sequences for read mapping. For each of the four sets of parents, we retained a locus only if it met the following criteria: 1) the locus had at least 20X coverage, 2) the locus was heterozygous in at least one parent, and 3) the count for the minor allele of a locus was no less than 5. The remaining polymorphic loci and the SNPs they contained were assigned back to the scaffolds in the draft oyster genome.

3.3.4 Statistical analysis

Replicate larval pools with extremely low coverage due to low number of reads were excluded in subsequent analyses. In the 10 remaining larval pools, allele counts of replicate pools for each family were summed up and used for statistical tests. False discovery rate corrected p-values were calculated for each family using Fisher's Exact Test. All statistical analysis was done in the R statistical software.

3.3.5 Gene enrichment analysis

Because of the large extent of linkage equilibrium and relative small size of the scaffolds in the published oyster genome, we also did analyses at the scaffold level. Information of genes in the selected scaffolds (see Table 3.1 for the list of scaffolds) was

extracted from the oyster draft genome (Zhang et al. 2012). Gene ontology enrichment analysis of the extracted genes was done using the Gene Ontology Enrichment Analysis Software Toolkit (GOEAST) (Zheng & Wang 2008) using all annotated genes as background. Adrian Alexa's improved weighted scoring algorithm (Alexa et al. 2006) and Yekutieli's procedure to control false discovery rate under dependency (Benjamini & Yekutieli 2001) were used for analysis.

3.4 Results and discussion

3.4.1 Sequencing and SNP genotyping

Our sequencing produced approximately 121.4 million 50-base single-end SOLiD reads, generating 4.4 Gb of sequence data. A total of 69,901 AlfI recognition sites were extracted from the published oyster genome (Zhang et al. 2012) and used as a reference. 95.1 million reads passed the quality filters, of which 24.9 million were mapped to the reference. Average coverage per AlfI sites was 266, 187, 234, and 132 for the four families. The sequencing yields and mapping efficiencies of each individual family are summarized in Table 3.2.

Genotyping of parental samples identified 17,429 within-family polymorphic loci (SNPs), 10,846 (62.2%) of which were polymorphic in only one family. Only 568 (3.3%) SNPs occurred in all four families. The distribution of sequencing coverage showed a tailed normal distribution (Figure 3.3). Sequencing coverage is high, with mean coverage of 200.

3.4.2 Comparing allele frequencies

Distribution of differences in allele frequency between the controls and the treatments is shown in Figure 3.4. 36% of SNPs showed allele frequency differences > 10%, and only 0.2% (59) of SNPs showed allele frequency differences > 50%. Of all the SNPs, 32, 7, 2 and 221 showed significant differences in allele frequencies between the controls and the treatments at a false discovery rate of 0.05 in families 1-4, respectively. Nine SNPs showed significant differences in more than one family (Table 3.3).

The size of *C. gigas* genome is 637 M or 545 M depending on the estimation method used (Zhang et al. 2012). The current oyster genome assembly contains 11,969 scaffolds and 28,027 genes. The largest scaffold is ~1.96 M in length; 98% of the scaffolds are shorter than 0.53 M. Because of the small scaffold size, the sliding window approach for mapping quantitative trait loci (QTL) for BSA, proposed by Magwene et al. (2011), is not feasible in our analysis. We compared the families at the scaffold level for two reasons: First, only a small proportion (3.3%) of SNPs are shared by all families, whereas all scaffolds are shared by families, making direct family to family comparison possible. Second, the extent of linkage disequilibrium is large and the size of scaffold is small, therefore there is little recombination within a scaffold after one generation. Genome length of *C. gigas* is 776 cM for males and 1020 cM for females (Hubert & Hedgecock 2004), and the length of DNA corresponding to 1 cM after one generation is between 0.53 M and 0.82 M. Since 98% of the scaffolds are shorter than 0.53 M, the scaffolds can be treated as a single locus for the sake of family to family comparison.

To be able to compare the scaffolds, we assigned the smallest P-values of SNPs in a scaffold as the P-value for the scaffold. There are 32, 6, 3 and 205 scaffolds with P-value < 0.05 in the four families, respectively. Among them, 17 scaffolds showed up in more than one

family (Table 3.1). Because of the large differences in the number of significant scaffolds in different families, we ranked the scaffolds based on their P-values (smaller is better) and looked at the top 50 scaffolds in each family. All together 175 scaffolds were among the top 50 scaffolds in at least one family. One of the scaffolds, scaffold100, is among the top 50 in all four families, 22 scaffolds are common in 2 families (Table 3.4), and the remaining 152 scaffolds showed up in only one family.

3.4.3 Reliability of allele frequency estimation

Several methods that use high throughput sequencing technologies to estimate allele frequencies in pooled samples have successfully identified previously known QTLs (Kim et al. 2011; Parts et al. 2011; Boitard et al. 2012; Mullen et al. 2012; Liu et al. 2012), but to our knowledge, 2b-RAD has not been previously used for estimating allele frequencies in pooled samples. To evaluate the precision of these estimates, we first calculated logarithm of odds ratio of two genotypes for each locus in each replicate. Correlations of logarithm of odds ratio between replicates were calculated for loci with different coverage. We found that correlations were high with medium to high coverage (Figure 3.5). The correlation is > 0.9 for 50X coverage and more than 0.95 for 100X or more coverage.

Some SNPs/scaffolds with P-value < 0.05 were identified in multiple families, highlighting genomic regions associated with *Vibrio* resistance in more than one families. Comparing family 1 and family 4, 9 of the 32 SNPs with significant changes in allele frequencies in family 1 also showed significant differences in family 4. 12 SNPs were polymorphic only in family 1. The remaining 12 SNPs were polymorphic in family 4 but did not show significant differences. Possible explanations are: (1) More than one mechanism is

responsible for *V. tubiashii* resistance and different sets of genes are involved in different families. (2) There is a large number of false positives even though FDR correction was used during statistical analysis.

3.4.4 Genes in the scaffolds with P-value < 0.05 in more than one family

Significant SNPs identified through association between genotypes and phenotypes were rarely the causal SNPs; the causal SNPs were likely to be located in the genes in surrounding genomic regions that were in linkage disequilibrium with the significant SNPs. Therefore, it is useful to explore genes near the significant SNPs. Because the genomic scaffolds were relatively small, we looked at all the genes in the same scaffold as the significant SNPs.

A total of 849 genes in those scaffolds with P-value < 0.05 in more than one family were extracted from the oyster genome and used for gene ontology enrichment analysis. Several Gene Ontology (GO) terms were enriched in these genes (Table 3.5). Notably, several highly represented GO terms, discussed below, can potentially influence the outcome of immune responses in oysters.

Initiation of immune response requires the recognition of an invading pathogen. Peptidoglycan is a major component of many bacterial cell walls and can be recognized by the host peptidoglycan recognition proteins (PGRP). The PGRP family is found in diverse metazoans including insects, mammals, and mollusks, and plays an important role in innate immunity. PGRPs can recognize peptidoglycan and initiate signal transduction, resulting in the activation of other immune genes including NF-κB like genes (Dziarski & Gupta 2006). Some PGRPs are directly involved in antimicrobial processes via N-acetylmuramoyl L-

alanine amidase activity (Mellroth & Steiner 2006). Such PGRPs are enriched in our gene set (GO term: "peptidoglycan catabolic process", "N-acetylmuramoyl-L-alanine amidase activity"). The PGRP family has previously been characterized in *C. gigas* (Itoh & Takahashi 2008; Itoh & Takahashi 2009), and one PGRP showed an increased expression upon *V. tubiashii* challenge (Itoh & Takahashi 2009).

Enriched "Rab GTPase binding" and "sequence-specific DNA binding transcription factor activity" categories may indicate signaling cascades in response to infection. Rab GTPase family is a member of Ras GTPase superfamily, which is activated during stress responses (Zhang et al. 2012). Several Ras GTPase genes and one NF-κB gene are located in genomic scaffold1009, where we found 2 SNPs with significant differences in allele frequency. The NF-κB pathway is highly conserved (Hoffmann 1999), and plays an important role in regulating immune response (Lemaitre & Hoffmann 2007). Several components in the NF-κB pathway have been characterized in *C. gigas* (Montagnani et al. 2004; Montagnani et al. 2008), although their exact role in immune response is unclear. Increases in expression of NF-κB gene (Roberts et al. 2009) or in activity of NF-κB protein (Zhu & Wu 2008) have been observed in oyster hemocytes upon challenge with bacterial or bacteria related product. A Toll-like receptor, known to activate NF-κB, is up-regulated in oysters surviving after virulent *Vibrio* infection compared to avirulent *Vibrio* infection (de Lorgeril et al. 2011).

NF-κB is also involved in the regulation of apoptosis (Farhana et al. 2005), which is thought to play an important role in mollusk immune systems (Sokolova 2009). Apoptosis can be triggered by Ca^{2+}-dependent proteases, i.e. calpain (Momeni 2011), which is also enriched in our gene set (GO term: "calcium-dependent cysteine-type endopeptidase

activity"). Calpain is activated by increased Ca^{2+} concentration (Momeni 2011), and interestingly, two GO terms related to the regulation of Ca^{2+} homeostasis (GO term: "calcium ion transport", "inositol 1, 4, 5-trisphosphate-sensitive calcium-release channel activity"), are also over-represented. Inositol 1, 4, 5-trisphosphate (IP_3) is a secondary messenger in signal transduction. Binding of IP_3 to its receptors regulates a number of Ca^{2+} dependent processes, including apoptosis (Boehning et al. 2004). Taken together, these observations suggest the involvement of the NF-κB regulatory pathway and regulation of apoptosis in successful immune response to *V. tubiashii*.

Note that the list of genes generated in our study is only from scaffolds associated with *V. tubiashii* resistance, and genes on the list are not necessarily involved in immune response. Some of these genes may simply be located in the same scaffolds as genes involved in the immune response. Also, since *C. gigas* is not a model organism, the annotation may not describe the function of the genes in oysters. Functions of the annotated genes were tested in model organisms but not oysters, and the actual function of a gene may differ in different organisms (Deng et al. 2010). Therefore, extra caution should be taken in drawing conclusions from the GO analysis.

3.4.5 2b-RAD sequencing of pooled samples for testing marker-trait association

In this study, we use 2b-RAD sequencing to compare allele frequencies in pooled samples and identified SNPs associated with *V. tubiashii* resistance and genes potentially involved in this process. Pooling and selective genotyping of individuals with extreme phenotypes has been used with various techniques as a cost-effective way to search for marker-trait association (see a discussion in Xu & Crouch 2008). Pooling may be the only

option when the amount of DNA from a single individual is not enough for genotyping as in the case of this study, or when pools of samples are easier to obtain compared to large number of individual samples. To obtain a high marker density for genome scan, a SNP microarray can be used (Butcher et al. 2004; Meaburn et al. 2005), which requires a microarray developed beforehand, requiring additional time and money. Another possible approach is whole genome resequencing, although this is cost-effective and feasible only in organisms with small genomes (Ehrenreich et al. 2010). Instead of a whole genome scan, resequencing of selected genes or regions in the genome can greatly reduce the overall cost (Mullen et al. 2012). This, however, requires ample prior knowledge of the trait to make informed decisions about which genes or genomic regions to study. 2b-RAD presents the following advantages: 1) it can generate large number of SNPs to achieve high marker density without the need for prior knowledge of DNA sequences. 2) It is cost-effective because it only resequences a portion of the genome. 3) The number of loci can be adjusted using adaptors with selective nucleotides to amplify a subset of the loci, further reducing the cost. 4) Also, the protocol of 2b-RAD is simpler than other RAD protocols, making preparation of large number of samples easier (Wang et al. 2012).

In our study, we have neither individual genotypes nor the complete genome to be able to map QTLs. Instead we have a list of putative SNPs. Further testing of the association between SNPs identified in this study and *V. tubiashii* resistance in oyster larvae in the breeding population is therefore a logical next step. Once marker-trait associations have been validated, these SNPs can be used in marker-assisted selection.

3.5 Conclusion

In this study, we used a cost-effective method, 2b-RAD sequencing, to identify loci differing in allele frequencies at the genome scale between pools of samples of two kinds: survivors of a laboratory *V. tubiashii* challenge and contemporary unchallenged controls, to identify SNPs associated with larval resistance to *V. tubiashii* in Pacific oysters. With validation, these SNPs can be used for marker-assisted selection to increase the resistance of Pacific oyster larvae to *V. tubiashii*.

3.6 Acknowledgements

Chen thanks financial support from the China Scholarships Council (CSC). We thank Molluscan Broodstock Program (MBP) staff for providing help in raising oyster larvae, and Dr. Gerd Bobe and Mr. Gu Mi for the advice in data analysis, and Dr. William Hohenboken for the advice on the manuscript.

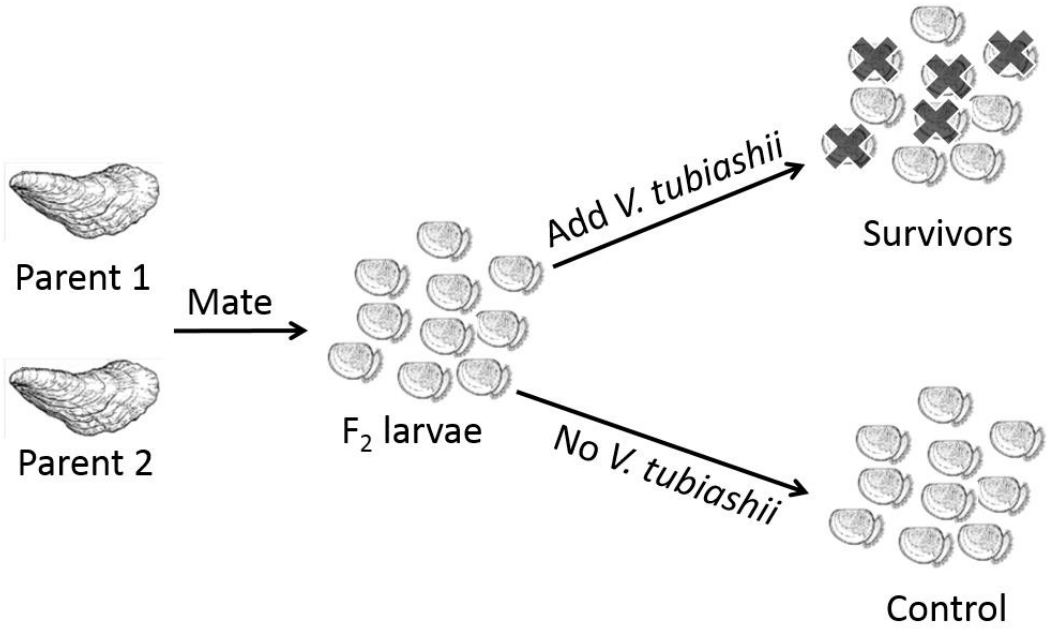

Figure 3.1. Crossing design and sampling. Parent 1 and parent 2 are brother/sister, F_2 larvae were divided into 6 groups. Three of them were treated with *V. tubiashii* until mortalities were greater than 50%; the other three were raised without exposure to *V. tubiashii*.

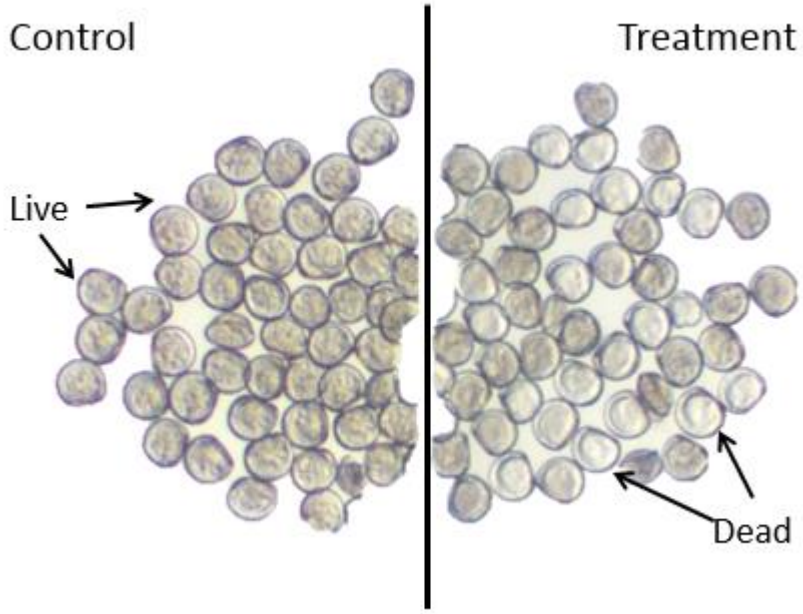

Figure 3.2. Larvae samples from one of the families. Larvae from the control pool are on the left and larvae from the treatment pool are on the right. Live larvae have tissues inside the shells whereas dead larvae have empty shells.

Table 3.1. Scaffolds containing SNP with P-value < 0.05 in more than one family

Scaffold[1]	SNP position	Alleles of SNP	P-value[2]	Family
scaffold100	397499	G/T	9.4×10^{-4}	4
	397499	G/T	1.4×10^{-2}	2
scaffold1070	98905	C/T	1.6×10^{-3}	4
	98912	A/C	6.6×10^{-3}	1
scaffold1328	307440	A/G	3.1×10^{-4}	4
	307440	A/G	1.2×10^{-2}	1
scaffold1512	85	C/T	1.2×10^{-6}	4
	85	C/T	1.4×10^{-3}	1
scaffold37010	10635	G/T	1.8×10^{-5}	4
	10635	G/T	1.2×10^{-2}	1
scaffold397	44272	A/G	5.8×10^{-6}	1
	44272	A/G	4.8×10^{-2}	4
scaffold43956	26238	A/C	1.5×10^{-2}	4
	26252	A/T	3.8×10^{-2}	1
scaffold733	1039525	C/T	1.6×10^{-2}	4
	1039525	C/T	3.2×10^{-2}	1
scaffold756	337938	C/T	4.5×10^{-3}	4
	337938	C/T	2.0×10^{-2}	1
scaffold987	211620	A/C	1.4×10^{-4}	4
	211620	A/C	3.4×10^{-2}	1
scaffold1599	52815	A/G	1.9×10^{-2}	1
	545370	C/G	3.4×10^{-4}	4
scaffold169	108979	A/G	2.7×10^{-2}	3

	191066	C/T	9.4×10^{-4}	4
scaffold203	167810	A/G	5.0×10^{-2}	4
	649467	A/T	2.0×10^{-2}	1
scaffold210	169111	A/G	2.1×10^{-3}	2
	410030	A/C	1.1×10^{-2}	4
scaffold43742	130520	A/G	1.0×10^{-2}	4
	299946	A/C	4.5×10^{-2}	1
scaffold675	297397	A/T	1.3×10^{-3}	2
	66088	A/C	1.5×10^{-2}	4
scaffold71	215400	G/T	2.1×10^{-2}	1
	98507	A/C	3.1×10^{-2}	4

[1]: Assembly by Zhang et al. (2012)
[2]: False discovery rate corrected P-value < 0.05

Table 3.2. Sequencing and SNP statistics

	Family 1	**Family 2**	**Family 3**	**Family 4**
Total sequencing reads generated (M)	22.2	15.5	15.9	10.8
Number of high quality reads (M)	17.5 (79%)	11.9 (77%)	12.1 (76%)	8.5 (78%)
Reads mapped to the reference[1] (M)	5.3 (30%)	3.5 (29%)	4.1 (34%)	3.0 (35%)
Number of RAD-tag	35063	37156	36262	32865
Average coverage per RAD-tag	266	187	234	132
Number of bases covered[2] (G)	1.1	1.1	1.1	1.0
Number of SNPs	6,585	6,825	5,784	7,859
Number of significant SNPs[3]	32	7	2	221

[1]: The reference is AlfI sites extracted from the genome;
[2]: Removing enzyme recognition sites;
[3]: False discovery rate corrected P-value < 0.05; Percentage in parentheses represent proportions related to the category above each field; counts for each family are the sum of the control and the treatments.

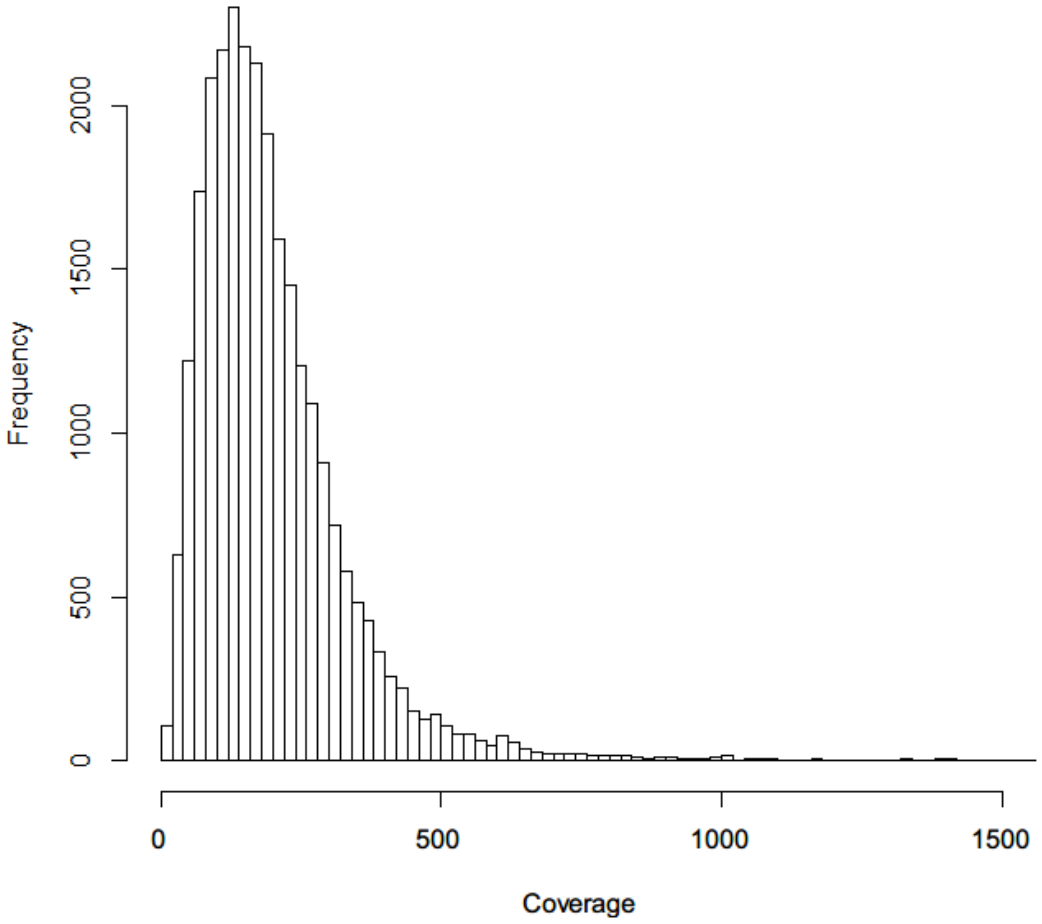

Figure 3.3. Histogram of the distribution of coverage of the polymorphic loci in all families.

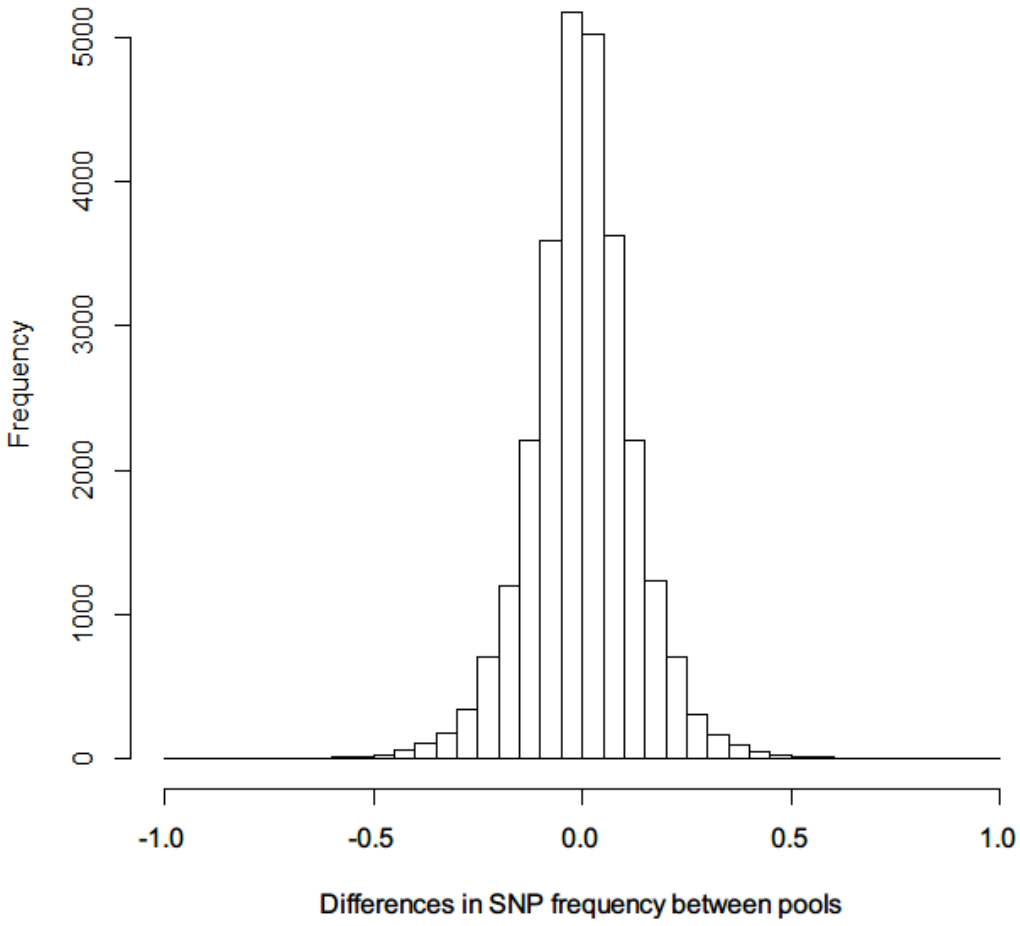

Figure 3.4. Histogram of the distribution of the allele frequency differentials observed between control and survivors pools.

Table 3.3. SNPs with P-values < 0.05 in more than one family

Scaffold	Position	Alleles	P-value[1]	family
scaffold100	397499	G/T	9.4×10^{-4}	4
			1.4×10^{-2}	2
scaffold1328	307440	A/G	3.1×10^{-4}	4
			1.2×10^{-2}	1
scaffold1512	85	C/T	1.2×10^{-6}	4
			1.4×10^{-3}	1
scaffold37010	10635	G/T	1.8×10^{-5}	4
			1.2×10^{-2}	1
scaffold397	44272	A/G	5.8×10^{-6}	1
			4.8×10^{-2}	4
scaffold43956	26238	A/C	1.5×10^{-2}	4
			3.8×10^{-2}	1
scaffold733	1039525	C/T	1.6×10^{-2}	4
			3.2×10^{-2}	1
scaffold756	337938	C/T	4.5×10^{-3}	4
			2.0×10^{-2}	1
scaffold987	211620	A/C	1.4×10^{-4}	4
			3.4×10^{-2}	1

[1]: False discovery rate corrected P-value < 0.05

Table 3.4. Scaffolds that are among top 50 in more than one family

Scaffold[1]	Family	Rank in family
scaffold100	1	33
	2	6
	3	3
	4	29
scaffold1070	1	7
	4	35
scaffold1324	1	45
	2	49
scaffold1328	1	8
	4	22
scaffold1341	2	37
	4	2
scaffold1360	3	44
	4	49
scaffold1512	1	6
	4	3
scaffold157	3	49
	4	6
scaffold1599	1	14
	4	25
scaffold1665	1	46
	3	25
scaffold1672	2	7
	3	7
scaffold169	3	2
	4	31

scaffold198	1	34
	2	17
scaffold203	1	16
	2	24
scaffold247	1	48
	2	18
scaffold323	1	39
	3	28
scaffold332	1	20
	2	25
scaffold37010	1	11
	4	8
scaffold4	2	30
	4	15
scaffold43956	1	28
	2	28
scaffold492	2	45
	3	32
scaffold962	2	9
	3	41
scaffold987	4	16
	10	24

[1]: Assembly by Zhang et al. (2012)

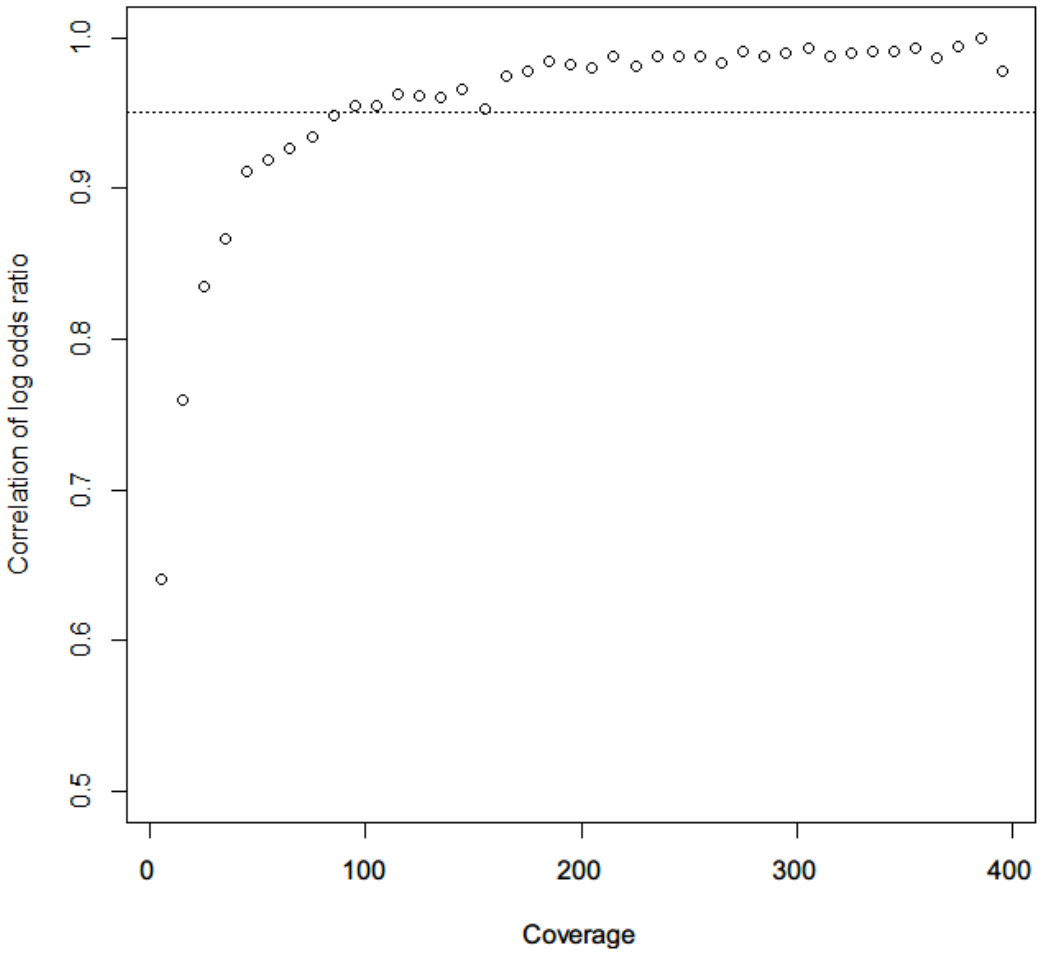

Figure 3.5. Correlation of logarithm of odds ratios between replicates at different coverage. Coverage was the smaller coverage of the two replicates for each locus. Correlations were calculated in each increasing windows of 10X starting from 0. Within each window, odds ratio of the two genotypes for each SNP in each replicate is calculated and a correlation of the logarithm of odds ratio of all loci between two replicates were calculated. The dashed line correspond to 0.95.

Table 3.5. Enriched gene ontology terms

GOID	GO Term	Gene counts[2]	Background gene counts[2]	P-value[1]
Biological process				
GO:0006816	calcium ion transport	3 (849)	27 (28,027)	4.9×10^{-2}
GO:0009311	oligosaccharide metabolic process	4 (849)	6 (28,027)	1.2×10^{-5}
GO:0007600	sensory perception	2 (849)	9 (28,027)	2.9×10^{-2}
GO:0006614	SRP-dependent cotranslational protein targeting to membrane	2 (849)	7 (28,027)	1.8×10^{-2}
GO:0009253	peptidoglycan catabolic process	3 (849)	9 (28,027)	2.1×10^{-3}
Molecular function				
GO:0004499	N,N-dimethylaniline monooxygenase activity	5 (849)	15 (28,027)	6.6×10^{-5}
GO:0050660	flavin adenine dinucleotide binding	6 (849)	72 (28,027)	2.4×10^{-2}
GO:0050661	NADP binding	5 (849)	21 (28,027)	3.9×10^{-4}
GO:0004842	ubiquitin-protein ligase activity	6 (849)	57 (28,027)	8.3×10^{-3}
GO:0003700	sequence-specific DNA binding transcription factor activity	17 (849)	339 (28,027)	3.7×10^{-2}
GO:0005507	copper ion binding	7 (849)	73 (28,027)	7.5×10^{-3}
GO:0004198	calcium-dependent cysteine-type endopeptidase activity	3 (849)	20 (28,027)	2.3×10^{-2}
GO:0005272	sodium channel activity	3 (849)	25 (28,027)	4.1×10^{-2}
GO:0005220	inositol 1,4,5-trisphosphate-sensitive calcium-release channel activity	2 (849)	6 (28,027)	1.3×10^{-2}

GO:0017137	Rab GTPase binding	2 (849)	6 (28,027)	1.3×10^{-2}
GO:0004806	triglyceride lipase activity	2 (849)	6 (28,027)	1.3×10^{-2}
GO:0008745	N-acetylmuramoyl-L-alanine amidase activity	3 (849)	9 (28,027)	2.2×10^{-3}
Cellular component				
GO:0005634	nucleus	31 (849)	679 (28027)	1.7×10^{-2}
GO:0005795	Golgi stack	4 (849)	6 (28027)	2.9×10^{-5}

[1]: False discovery rate corrected P-value
[2]: Numbers in parentheses are the total number of genes in the group (849 for our gene list and 28,027 for the background genes)

Chapter 4

4. Whole transcriptome profiling by tag-based RNA-seq of Pacific oyster (*Crassostrea gigas*) families differing in resistance to *Vibrio tubiashii*

Chao Chen[1], Eli Meyer[2], Mark Camara[3*]

[1]Department of Animal and Rangeland Sciences, Oregon State University, Corvallis, OR 97331, USA

[2]Department of Zoology, Oregon State University, Corvallis, OR 97331, USA

[3]Cawthron Institute, 98 Halifax Street East, Private Bag 2, Nelson 7042, New Zealand

[*] Corresponding author

KEYWORDS: *Crassostrea gigas*, Pacific oyster, larval mortality, *Vibrio tubiashii*, gene expression, RNA-Seq, immunity

4.1 Abstract

The bacterial pathogen *Vibrio tubiashii* is one of several factors causing severe larval mortality of the Pacific oyster (*Crassostrea gigas*) on the west coast of the United States in recent years. Strains of oysters genetically resistant to *V. tubiashii* would be highly desirable. To understand the biology of resistance to *V. tubiashii* and to identify candidate genes related to resistance to *V. tubiashii* for marker development, we compared differences in gene expression between previously identified oyster families resistant versus sensitive to *V. tubiashii* infection. We used a tag-base RNA-Seq procedure to quantify transcriptomes of 3 samples for each family: larvae taken at the beginning of the experiment (0-hour control), larvae raised for 12 hours without *V. tubiashii* (12-hour control), and larvae raised for 12 hours with *V. tubiashii* (12-hour treatment). We identified 12 differentially expressed genes between resistant and sensitive families, including some genes involved in regulation of

immune response, apoptosis, and possibly development. In addition, we confirmed the observation of Taris et al. (2009) that transcriptional differences between resistant and sensitive families exists prior to a bacterial challenge, supporting the hypothesis that difference in resistance to *V. tubiashii* is due to inherit differences in gene expression rather than differential responses to pathogen challenge. The differentially expressed genes could be used as candidates for development of molecular markers that can be used for marker-assisted selection.

4.2 Introduction

The Pacific oyster is an important aquaculture species around the world. In the Pacific Northwest, oysters account for 83% of shellfish production in production weight, according to the Pacific Coast Shellfish Growers Association (PCSGA). Farmers rely heavily on seed produced by hatcheries but have been experiencing supply limitations in recent years due to high larval mortalities in commercial shellfish hatcheries. A pathogen, *Vibrio tubiashii*, was one of the possible causes of the mortalities. A correlation has been observed between larval mortality and outbreak *V. tubiashii*, during upwelling events in years when surface seawater temperature is higher than normal (Elston et al. 2008). Larval mortalities in hatcheries caused by *V. tubiashii* is expected to be more frequent and serious due to rising surface seawater temperature as a result of global warming.

Larval resistance to *V. tubiashii* is highly heritable (Chen, unpublished), and selective breeding is a promising approach to develop resistant strains. Marker-assisted selection can be used to facilitate evaluation of performance and selection by breeders. However, genes

and/or molecular markers need to be developed for marker-assisted selection, requiring the understanding of the molecular basis of *V. tubiashii* resistance.

In Chapter 2, we have identified families sensitive and resistant to *V. tubiashii*. In this study, we looked at the transcriptome profiles of those sensitive and resistant families using the tag-based RNA-Seq procedure developed by Meyer et al. (2011) on the Applied Biosystems SOLiD (SOLiD) high-throughput sequencing platform. Transcriptome profiling is a useful tool for understanding underlying mechanisms for complex traits. A number of studies have used this approach to describe responses to pathogens and stressful conditions in oysters (Rosa et al. 2012; de Lorgeril et al. 2011; Fleury & Huvet 2012; Zhao et al. 2012; Lang et al. 2009; Chapman et al. 2011). We compared gene expression between resistant and sensitive families, aiming to understand the relevant aspects of oyster innate immunity and to identify important candidate genes that can be used for marker-assisted selection. The objective of this study is to better understand what makes oyster larvae resistant to *V. tubiashii* at the gene expression level.

4.3 Materials and methods

4.3.1 Biological materials

1-day-old larvae from 6 families that were identified to be resistant or sensitive to *V. tubiashii* (Chapter 2) were obtained from a routine spawn produced by the Molluscan Broodstock Program (MBP) in May 17, 2011 (Langdon et al. 2003). Larvae from each family were aliquoted in 1 liter beakers with 900ml of 25 ℃ seawater at a density of 20 larvae ml^{-1}. Beakers were then transferred to pre-assigned random positions in a 25 ℃ room. The random

positions were assigned to the beakers using an online randomizer (http://www.randomizer.org/).

4.3.2 Bacteria preparation, challenge and sampling

V. tubiashii strain RE22 was grown on Luria-Bertani (LB) medium supplemented with 1% sodium chloride (LB-1% NaCl) and 100μg ml^{-1} of ampicillin (Ap) at 30 ℃ overnight on a shaker table. Log-phase bacterial culture were centrifuged gently and washed with seawater before use.

For each family, a single sample with more than 20,000 larvae was taken at the beginning of the experiment (0-hour control). Three replicate beakers of larvae were raised for 12 hours without *V. tubiashii* (12-hour control). Another three replicate beakers of larvae were raised for 12 hours with *V. tubiashii* at a density of 3×10^4 cells ml^{-1} (12-hour treatment). All larvae were raised at a density of 20 ml^{-1} at 25 ℃ without food. After 12 hours, larvae were sieved, combined and preserved in 1.5 ml microcentrifuge tubes with 1 ml of RNAlater solution. Samples were held at 4 ℃ overnight and transferred to -20 ℃ the next day.

4.3.3 RNA extraction and RNA-Seq library preparation

RNA was extracted using Qiagen RNeasy Mini Kit. Sequencing libraries were prepared following the protocol of Meyer et al. (2011). RNA was first fragmented by heat, and first-strand cDNA was synthesized using a primer containing oligo dT. cDNA was then amplified and sample specific barcodes were incorporated into the PCR products as part of the primer. All sequencing was done on SOLiD™ system at University of Texas at Austin.

4.3.4 Read processing and identification of differentially expressed genes

To reduce computational time, sequencing reads were filtered before mapping. Reads with long homopolymers (≥ 10 bp) and low quality reads (more than 23 bases with score < 20) were excluded. The high quality reads were then mapped to our reference using SHRiMP package (version 2.1.1b) (Rumble et al. 2009). The reference we used was an assembly made using Roche *de novo* assembler (version 2.6, with "-cdna" option) from publicly available data (accession No.: SRA026751). We did not use the established collection of CDS from the oyster genome (Zhang et al. 2012) because fewer reads were mapped to the CDS than to our assembly (10% for the CDS versus 32% for our assembly). A possible explanation for this is that CDS does not include the 3' untranslated region (3'-UTR) of mRNA, and a large portion of our reads were from the 3'-UTR region. Due to high variation in sequencing coverage among samples, 7 out of 18 samples had very low coverage and were excluded from subsequent analysis. The remaining 11 samples included four 0-hour-controls (two from resistant families and two from sensitive families), three 12-hour-controls (two from resistant families and one from sensitive families, and four 12-hour-treatments (two from resistant families and two from sensitive families). Among the 11 samples, differentially expressed genes (DEG) were detected using the R package DESeq (Anders & Huber 2010). To test the effects of time, *V. tubiashii* treatment, family type, and their interactions, on gene expressions, we used the following linear model:

$$Y_{ijk} = \mu + T_i + M_j + F_k + (TF)_{ik} + (MF)_{jk} + e_{ijk}$$

where Y_{ijk} is the dependent variable (read counts of genes), μ is the overall mean, T_i is the treatment effect, M_j is the time effect, F_k is the effect of family type (resistant or sensitive),

$(TF)_{ik}$ is the interaction between treatment and family type, $(MF)_{jk}$ is the interaction between time and family type, and e_{ijk} is the residual error. Calculated P-values were adjusted using false discovery rate (Benjamini & Yekutieli 2001) method in R (R Core Team 2013).

4.3.5 Validation of differential expression using quantitative reverse transcription polymerase chain reaction (qRT-PCR)

Fourteen differentially expressed genes and 4 candidate reference genes were selected for qRT-PCR. Candidate reference genes were selected based on their constant expression across all samples in RNA-Seq (coefficient of variation < 5%).

Primers were designed using Primer 3 (http://bioinfo.ut.ee/primer3-0.4.0/). T_M of all primers ranged from 57 ℃ to 61 ℃, and amplicon lengths ranged from 100 to 250 bp. Thirteen genes (9 DEG and 4 reference genes) successfully amplified product with target size (checked on 2% agarose gels). Efficiencies of the 13 primer pairs was tested using serial dilutions of cDNA, and primer pairs with efficiency lower than 1.8 or higher than 2.2 were excluded from subsequent analysis (1 of 13 primer pairs was excluded). We also did melting curve analysis on the remaining 12 primers, all of those primers amplified without non-specific products, andfor every gene, melting curves of different samples clustered together, with very little variation.

Before qRT-PCR, RNA's were treated with DNase I and precipitated with Lithium Chloride (LiCl) to remove DNA contaminations. cDNA was synthesized using SuperScript® II reverse transcriptase (Invitrogen™). To test for remaining DNA contamination, we included a control reaction without reverse transcriptase (-RT) for each sample. We then conducted qRT-PCR using –RT controls for each sample and for each gene, and excluded any sample/gene that showed amplification (2 of 12 genes excluded).

In total, 6 differentially expressed genes and 4 reference genes were tested on 8 samples. All qRT-PCR were done in triplicate for each gene/sample combination. Relative expressions were calculated following Livak & Schmittgen (2001), with the amount of RNA added to the reaction normalized using the average C_T of the four reference genes.

4.3.6 Gene annotation

12 genes that were differentially expressed between the resistant and sensitive families. We used BLAST (Altschul et al., 1990) to ascertain similarity of the 12 genes to coding sequences (CDS) of the published oyster genome (Zhang et al. 2012). Annotations of the best matches with expected value less than 10^{-10} were used as annotation for the genes.

4.4 Results and discussion

4.4.1 Sequencing statistics

Our sequencing produced approximately 230 million 50 bp single end reads, of which ~164 million (71%) passed all quality filters. For the 11 samples used in differential gene expression analysis, a total of 99 million reads were produced (9 million per sample), 67% of which passed the quality filters. Of all the high quality reads, 21 million (1.9 million per sample) were mapped to the reference.

Gene expression levels for most genes were low (Figure 1), with the 25%, 50%, 75% quantiles being 1.8, 5.2 and 15.6 reads per million per sample, respectively. The highest expressed gene was an 18S rRNA gene, with the expression level of 212,748 reads per million per sample.

4.4.2 Differential gene expression

We found no statistically significant effect for treatment, treatment × family interaction, or time × family type interaction for any of the genes examined. 55 genes were differentially expressed (FDR corrected P-value < 0.1) between 0-hour and 12-hour samples, and 55 genes differentially expressed between the resistant and sensitive families.

The presence of time effects and absence of treatment effects indicates the DEG between 0-hour and 12-hour samples were caused by time rather than impact of the *V. tubiashii* challenge. To further examine this result, we did two additional comparisons: 1) between 0-hour-control and 12-hour-control samples, and 2) between 0-hour-control and 12-hour-treatment samples. We found 43 DEG for the 0-hour-control versus 12-hour-control comparison, and 45 DEG for the 0-hour-control versus 12-hour-treatment comparison (FDR adjusted P-value < 0.1). We then compared the fold changes of all the DEG in the two comparisons and found the change in gene expression appeared well correlated between the two comparisons (Figure 4.2). In other words, the treatment of *V. tubiashii* did not have detectable effects on gene expression.

The existence of family type effects and the absence of treatment effects is in agreement with previous findings of Taris et al. (2009) looking at differences in gene expression of adult oysters between resistant and sensitive families following challenge of *V. tubiashii*. They found that the differential gene expression between resistant and sensitive families existed before the bacterial challenge was applied, supporting the hypothesis that difference in resistance to *V. tubiashii* is due to inherit differences in gene expression rather than differential responses to pathogen challenge.

4.4.3 qRT-PCR validation of gene expression

To evaluate gene expression levels measured by RNA-Seq, we performed qRT-PCR on 8 samples using the same RNA as used for RNA-Seq. Overall, there is good agreement between qRT-PCR and RNA-seq on relative gene expression levels among samples (Figure 4.3), with regression P-value = 3.4×10^{-12} and $R^2 = 0.6629$, indicating high accuracy of our measurements using RNA-Seq. Among individual genes, all except for Gelsolin showed agreement between RNA-Seq and qRT-PCR. Disagreement between qRT-PCR and other methods has been reported previously in Pacific oysters (Taris et al. 2008), possibly due to differences in primer binding during qRT-PCR in different individuals. Regression P-value equaled 3.4×10^{-14} and R^2 equaled 0.78 after excluding the Gelsolin gene.

4.4.4. Potential roles of DEG between the resistant and the sensitive families for oyster immunity

Table 4.1 shows the list of DEG between resistant and sensitive families. Of the 12 genes, we found no hit when BLAST against the oyster CDS database. Among the remaining 10 genes, 5 genes were not annotated, 5 genes were annotated and are thought to be involved in immune response as discussed below.

Complement C1q protein plays a role in the modulation of innate and adaptive immune response (Sontheimer et al. 2005; Yuste et al. 2006). Taris et al. (2009) found that C1q gene was differentially expressed between high- and low-survival families in the Pacific oysters infected with *V. tubiashii*. Up-regulation of C1q gene was also found in the Pacific oysters after expoure to another *Vibrio* species, *Vibrio splendidus* (de Lorgeril et al. 2011). Identified as an immune gene in the Mediterranean mussel, *Mytilus galloprovincialis*, C1q gene was highly diverse among individuals at both the genomic sequence and gene

expression levels (Gestal et al. 2010). C1q domain protein and inhibitor of apoptosis (IAP) genes, as will be mentioned below, are highly over-represented in the oyster genome (Zhang et al., 2012), indicating their important role in oyster immune defense.

Baculoviral inhibitor of apoptosis (IAP) repeat-containing protein 2 is a member of IAP protein family, which suppresses programmed cell death (apoptosis) (Deveraux & Reed 1999). IAPs were found to be up-regulated in the Pacific oysters infected with *V. splendidus* (de Lorgeril et al. 2011) and upon challenge by various other stressors (Zhang et al. 2012). In our study, IAP exhibited higher expression in the resistant families than in the sensitive families, indicating promotion of apoptosis may play an important role in successful defense against *V. tubiashii*. Apoptosis protects the host by eliminating a pathogen without causing inflammation (Sokolova 2009). Some pathogens can inhibit apoptosis of the host cells to protect themselves (Böttger et al. 2008).

The conservative NF-κB pathway plays an important role in the regulation of apoptosis (Lemaitre & Hoffmann 2007). NF-κB family members share a DNA binding domain, Rel homology domain (RHD) (Karin & Ben-neriah 2000). In our study, a gene with RHD was differentially expressed between the resistant and the sensitive families, suggesting that the NF-κB pathway may be involved in protecting oysters from *V. tubiashii*.

Another differentially expressed gene in our study, E3 ubiquitin-protein ligase, was found to participate in stress responses in Pacific oysters (Zhang et al. 2012). Zhang et al. (2013) found that E3 ubiquitin ligase played a negative role in regulating immune response to double-strand DNA (dsDNA). In our study, E3 ubiquitin ligase also expressed at a lower level in the resistant families.

β-lactamase is an enzyme that provides resistance to β-lactam antibotics (penicillins for example) in many bacteria (Neu 1969). Interestingly, a gene with a β-lactamase-related domain was expressed at a lower level in the resistant families than in the sensitive families in our study. However, the role of β-lactamase in the oyster immune system is not clear. A possible explanation is β-lactamase may be involved in regulating the dynamics of the commensal bacteria, some of which may secrete antibiotic to inhibit other bacteria, including pathogens like *V. tubiashii*. The higher expression of β-lactamase in the sensitive families may indirectly facilitate infection by *V. tubiashii*. Currently, little is known about commensal bacteria in oysters. However, some characteristics of microflora in oysters have led to the hypothesis that oyster microflora may be essential for immunity and maintenance of homeostasis (Schmitt et al. 2012). Alternatively, it is possible the β-lactamase plays other roles in the immune system other than degrading β-lactam antibotics.

Of the 5 DEG with unknown function, 4 showed high expression levels in larval stages and low or no expression in adult oysters (Zhang et al. 2012). Presumably, these genes were involved in regulation of larval development rather than immune response. However, different expression levels of these genes between the resistant and sensitive families may represent differences in energy allocation and/or energy efficiency, as was proposed as an explanation for summer mortality syndrome in the Pacific oysters (Li et al. 2007).

4.5 Conclusion

In our study, we used RNA-Seq to examine differences in gene expression between Pacific oyster families resistant versus sensitive to *V. tubiashii*. Our results confirm the findings by Taris et al. (2009) that differences in gene expression between sensitive and

resistant families exist prior to infection by *V. tubiashii*. We also identified genes related to resistance to *V. tubiashii*, providing insights into the oyster immune system, as well as candidate genes that can be used for development of molecular markers for marker-assisted selection (MAS).

4.6 Acknowledgements

Chen thanks financial support from the China Scholarships Council (CSC). We thank Molluscan Broodstock Program (MBP) stuff for providing oyster larvae, Dacey Mercer and Yang Lu for their help in sample collection, and Jacob Snelling for his help in library preparation.

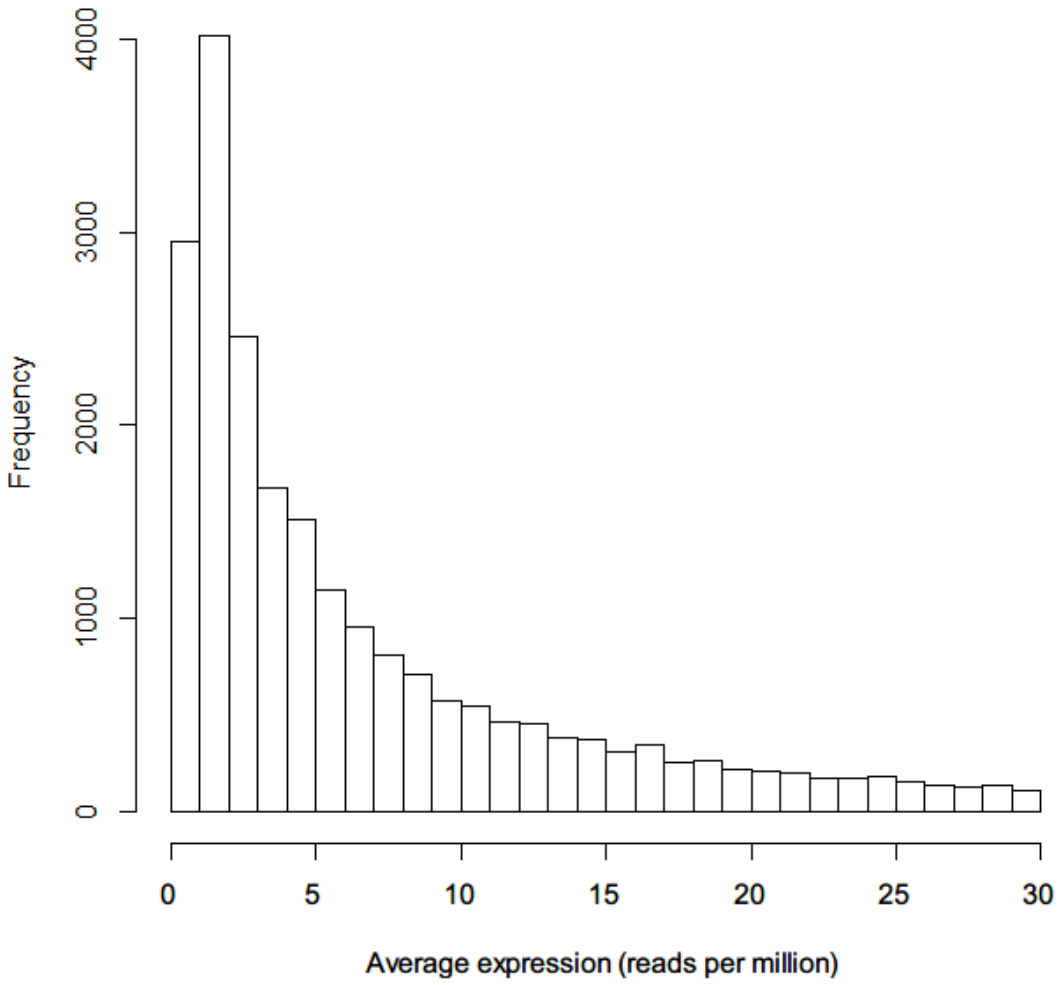

Figure 4.1. Distribution of expression levels for genes with ≥ 1 read per sample. 3,634 genes with expression higher than 30 reads per million per sample were not shown.

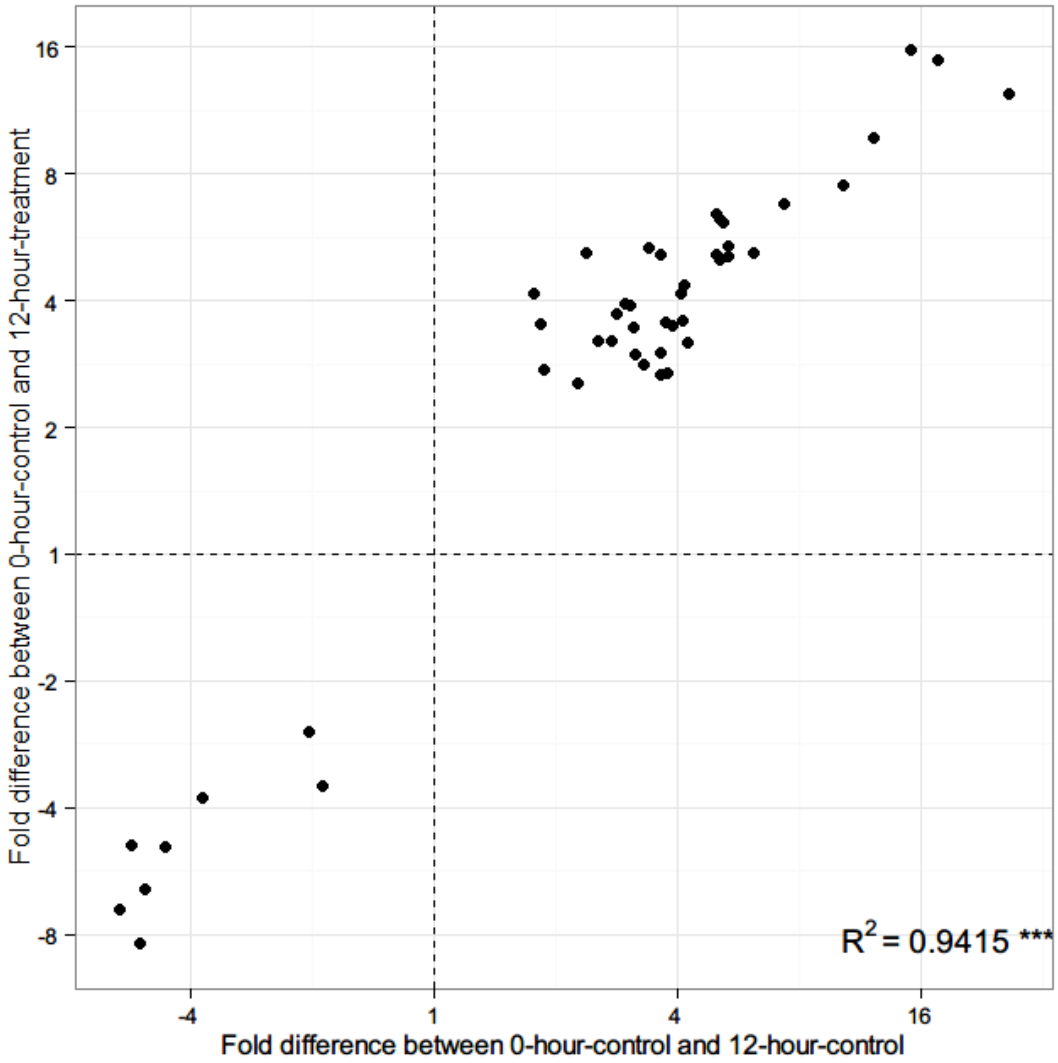

Figure 4.2. Comparison of fold changes of DEG between 0-hour-control and 12-hour-control (x axis), and 0-hour-control and 12-hour-treatment (y axis). Each point represents the fold change in expression level in the two comparisons. Gene expression fold changes were similar for all DEG in the two comparisons.

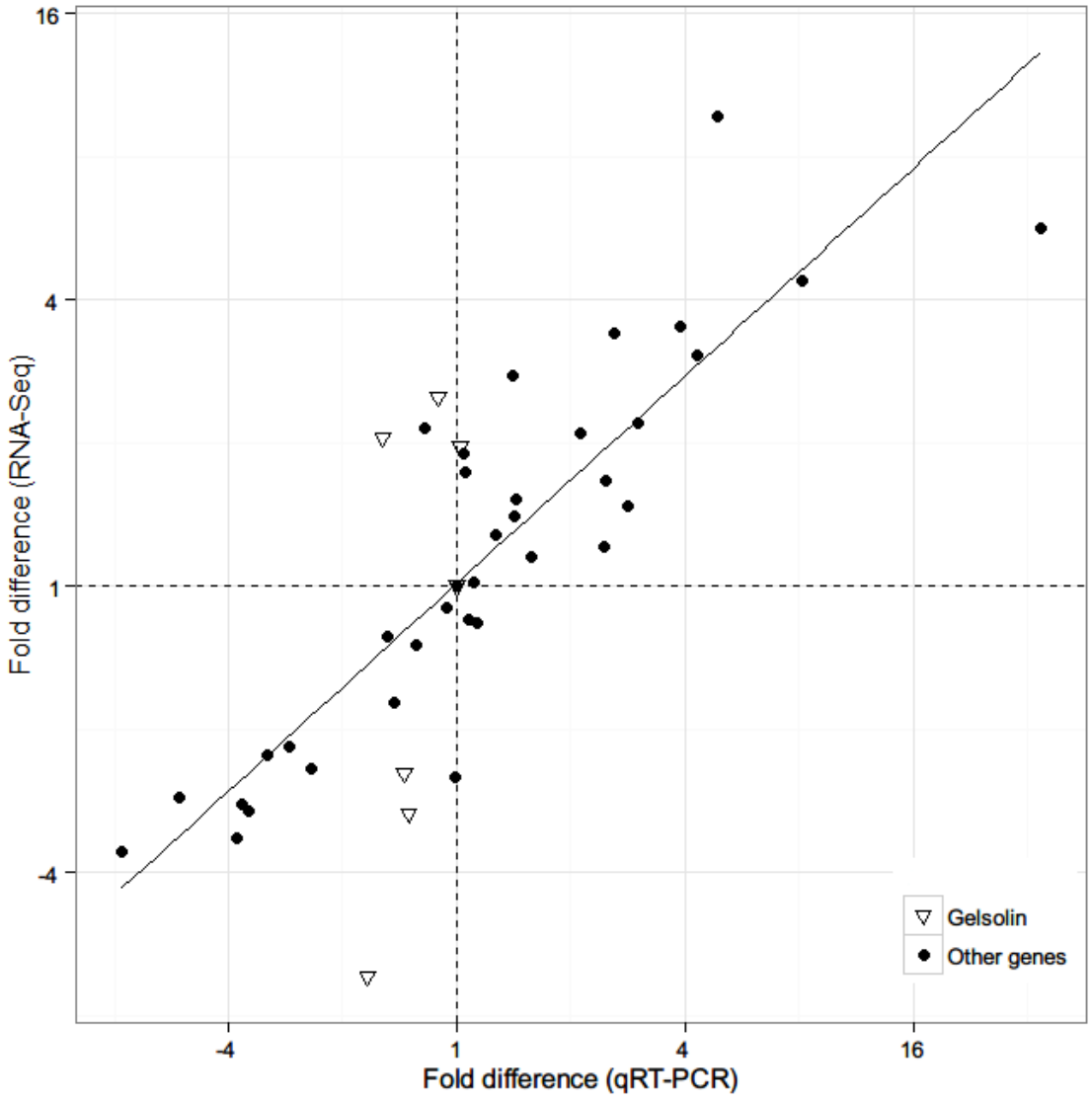

Figure 4.3. Comparison of gene expression levels assessedusing RNA-seq and qRT-PCR. Each point represents the expression of a gene in one sample relative to the references in RNA-Seq and qRT-PCR. All genes except Gelsolin (open triangle) showed agreement between qRT-PCR and RNA-Seq in relative expression levels (regression P-value = 3.4×10^{-12} and $R^2 = 0.6629$ for all genes, and regression P-value = 3.4×10^{-14} and $R^2 = 0.78$ after excluding the Gelsolin gene).

Table 4.1. List of differentially expressed genes between resistant and sensitive families.

IsogroupID	P-value[1]	Best hit in CDS	Expected value	Function
Isogroup03216	5.6×10^{-2}	CGI_10014672	0	C1q domain protein
Isogroup04882	1.6×10^{-2}	CGI_10026316	0	Baculoviral inhibition of apoptosis protein
Isogroup19161	4.1×10^{-2}	CGI_10016487	4×10^{-46}	Nuclear factor of activated T-cells 5
Isogroup19279	2.0×10^{-2}	CGI_10024137	0	E3 ubiquitin-protein ligase
Isogroup08194	1.4×10^{-6}	CGI_10004573	0	Uncharacterized protein ZK945.1 (with Beta-lactamasetype transpeptidase fold domain)
Isogroup02880	2.6×10^{-4}	CGI_10013253	2×10^{-86}	Unknown
Isogroup29028	1.6×10^{-2}	CGI_10002555	5×10^{-11}	Unknown
Isogroup32546	4.1×10^{-2}	CGI_10010207	10^{-134}	Unknown
Isogroup33091	5.6×10^{-2}	CGI_10010208	10^{-140}	Unknown
Isogroup00503	5.6×10^{-2}	CGI_10003275	3×10^{-76}	Unknown
Isogroup26632	9.5×10^{-2}	No hit		
Isogroup17845	4.1×10^{-2}	No hit		

[1]: FDR adjusted P-value

Bibliography

Alexa, A., Rahnenführer, J., & Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, *22*(13), 1600–7. doi:10.1093/bioinformatics/btl140

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410.

Anders, S., & Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biology*, *11*(10), R106. doi:10.1186/gb-2010-11-10-r106

Barton, A., Hales, B., Waldbusser, G. G., Langdon, C., & Feely, R. A. (2012). The Pacific oyster, *Crassostrea gigas*, shows negative correlation to naturally elevated carbon dioxide levels: Implications for near-term ocean acidification effects. *Limnology and Oceanography*, *57*(3), 698–710. doi:10.4319/lo.2012.57.3.0698

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, *29*(4), 1165–1188.

Boehning, D., Patterson, R. L., & Snyder, S. H. (2004). Apoptosis and calcium: new roles for cytochrome c and inositol 1,4,5-trisphosphate. *Cell Cycle*, *3*(3), 252–4.

Boitard, S., Schlötterer, C., & Nolte, V. (2012). Detecting selective sweeps from pooled next-generation sequencing samples. *Molecular Biology and Evolution*, *29*(9), 2177–86. doi:10.1093/molbev/mss090

Böttger, S., Jerszyk, E., Low, B., & Walker, C. (2008). Genotoxic stress-induced expression of p53 and apoptosis in leukemic clam hemocytes with cytoplasmically sequestered p53. *Cancer research*, *68*(3), 777–82. doi:10.1158/0008-5472.CAN-06-0968

Brown, C. (1973). The effects of some selected bacteria on embryos and larvae of the American oyster, *Crassostrea virginica*. *Journal of Invertebrate Pathology*, *21*(3), 215–223. doi:http://dx.doi.org/10.1016/0022-2011(73)90206-1

Brown, C. (1981). A study of two shellfish-pathogenic *Vibrio* strains isolated from a Long Island hatchery during a recent outbreak of disease. *Journal of Shellfish Research*, *1*(1), 83–87.

Brown, C., & Losee, E. (1978). Observations on natural and induced epizootics of vibriosis in *Crassostrea virginica* larvae. *Journal of Invertebrate Pathology*, *31*(1), 41–47. doi:http://dx.doi.org/10.1016/0022-2011(78)90107-6

Bus, A., Hecht, J., Huettel, B., Reinhardt, R., & Stich, B. (2012). High-throughput polymorphism detection and genotyping in *Brassica napus* using next-generation RAD sequencing. *BMC Genomics*, *13*(1), 281. doi:10.1186/1471-2164-13-281

Butcher, L. M., Meaburn, E., Liu, L., Fernandes, C., Hill, L., Al-Chalabi, A., … Craig, I. W. (2004). Genotyping pooled DNA on microarrays: a systematic genome screen of thousands of SNPs in large samples to detect QTLs for complex traits. *Behavior Genetics*, *34*(5), 549–55. doi:10.1023/B:BEGE.0000038493.26202.d3

Canesi, L., Gallo, G., Gavioli, M., & Pruzzo, C. (2002). Bacteria-hemocyte interactions and phagocytosis in marine bivalves. *Microscopy Research and Technique*, *57*(6), 469–76. doi:10.1002/jemt.10100

Chapman, R. W., Mancia, A., Beal, M., Veloso, A., Rathburn, C., Blair, A., … Sanger, D. (2011). The transcriptomic responses of the eastern oyster, *Crassostrea virginica*, to environmental conditions. *Molecular Ecology*, *20*(7), 1431–49. doi:10.1111/j.1365-294X.2011.05018.x

Cheng, T. (1996). Hemocytes: forms and functions. (V. S. Kennedy, R. I. E. Newell, A. F. Eble Eds.) *The Eastern Oyster: Crassostrea Virginica* (2nd ed.). College Park, MD, USA: University of Maryland Sea Grant Publications.

Clark, J., & Langmo, R. (1979). Oyster seed hatcheries on the US West Coast: an overview. *Marine Fisheries Review, 41(12), 10-16.*

Crippen, R. W., & Perrier, J. L. (1974). Use of neutral red and Evans blue for live-dead determinations of marine plankton. *Stain Technology*, *49*(2), 97–104.

Crnokrak, P., & Roff, D. A. (1995). Dominance variance: associations with selection and fitness. *Heredity*, *75*(5), 530–540. doi:10.1038/hdy.1995.169

De Lorgeril, J., Zenagui, R., Rosa, R. D., Piquemal, D., & Bachère, E. (2011). Whole transcriptome profiling of successful immune response to *Vibrio* infections in the oyster *Crassostrea gigas* by digital gene expression analysis. *PLoS ONE*, *6*(8), e23142. doi:10.1371/journal.pone.0023142

Dekkers, J. (2004). Commercial application of marker-and gene-assisted selection in livestock: strategies and lessons. *Journal of Animal Science*, *82*, E313–28.

Dekkers, J. C. M., & Hospital, F. (2002). The use of molecular genetics in the improvement of agricultural populations. *Nature Reviews Genetics*, *3*(1), 22–32. doi:10.1038/nrg701

Dempster, E. R., & Lerner, I. M. (1950). Heritability of threshold characters. *Genetics*, *35*(2), 212–36.

Deng, C., Cheng, C.-H. C., Ye, H., He, X., & Chen, L. (2010). Evolution of an antifreeze protein by neofunctionalization under escape from adaptive conflict. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(50), 21593–8. doi:10.1073/pnas.1007883107

Deveraux, Q. L., & Reed, J. C. (1999). IAP family proteins − suppressors of apoptosis. *Genes and development*, *13*(3), 239–252.

Dillies, M.-A., Rau, A., Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., … Jaffrézic, F. (2012). A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Briefings in Bioinformatics*. doi:10.1093/bib/bbs046

Dorfmeier, E. (2012). *Ocean acidification and disease: how will a changing climate impact Vibrio tubiashii growth and pathogenicity to Pacific oyster larvae?* M.S. Thesis, University of Washington, USA.

Dumbauld, B. R., Kauffman, B. E., Trimble, A. C., & Ruesink, J. L. (2011). The Willapa bay oyster reserves in Washington State: fishery collapse, creating a sustainable replacement, and the potential for habitat conservation and restoration. *Journal of Shellfish Research*, *30*(1), 71–83. doi:10.2983/035.030.0111

Duperthuy, M., Schmitt, P., Garzón, E., Caro, A., Rosa, R. D., Le Roux, F., … Destoumieux-Garzón, D. (2011). Use of OmpU porins for attachment and invasion of *Crassostrea gigas* immune cells by the oyster pathogen *Vibrio splendidus*. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 2993–8. doi:10.1073/pnas.1015326108

Dziarski, R., & Gupta, D. (2006). The peptidoglycan recognition proteins (PGRPs). *Genome Biology*, *7*(8), 232. doi:10.1186/gb-2006-7-8-232

Ehrenreich, I. M., Torabi, N., Jia, Y., Kent, J., Martis, S., Shapiro, J. A., … Kruglyak, L. (2010). Dissection of genetically complex traits with extremely large pools of yeast segregants. *Nature*, *464*(7291), 1039–42. doi:10.1038/nature08923

Elston, R., Frelier, P., & Cheney, D. (1999). Extrapallial abscesses associated with chronic bacterial infections in the intensively cultured juvenile Pacific oyster *Crassostrea gigas*. *Diseases of Aquatic Organisms*, *37*(2), 115–120.

Elston, R., & Leibovitz, L. (1980). Pathogenesis of experimental vibriosis in larval American oysters, *Crassostrea virginica*. *Canadian Journal of Fisheries and Aquatic Sciences*, *37*(6), 964–978. doi:10.1139/f80-126

Elston, R., Leibovitz, L., Relyea, D., & Zatila, J. (1981). Diagnosis of vibriosis in a commercial oyster hatchery epizootic: diagnostic tools and management features. *Aquaculture*, *24*, 53–62.

Elston, R. A., Hasegawa, H., Humphrey, K. L., Polyak, I. K., & Häse, C. C. (2008). Re-emergence of *Vibrio tubiashii* in bivalve shellfish aquaculture: severity, environmental drivers, geographic extent and management. *Diseases of Aquatic Organisms*, *82*(2), 119–34. doi:10.3354/dao01982

FAO. (2007). *Marker-assisted selection: current status and future perspectives in crops, livestock, forestry and fish*. (E. Guimarães, J. Ruane, B. D. Scherf, A. Sonnino, & J. D. Dargie, Eds.) Rome, Italy: Food and Agriculture Organization of the United Nations.

Farhana, L., Dawson, M., & Fontana, J. (2005). Apoptosis induction by a novel retinoid-related molecule requires nuclear factor-κB activation. *Cancer Research*, *65*(11), 4909–4917.

Fleury, E., & Huvet, A. (2012). Microarray analysis highlights immune response of Pacific oysters as a determinant of resistance to summer mortality. *Marine Biotechnology*, *14*(2), 203–17. doi:10.1007/s10126-011-9403-6

Garnier, M., Labreuche, Y., Garcia, C., Robert, M., & Nicolas, J. L. (2007). Evidence for the involvement of pathogenic bacteria in summer mortalities of the Pacific oyster *Crassostrea gigas*. *Microbial Ecology*, *53*(2), 187–96. doi:10.1007/s00248-006-9061-9

Gestal, C., Pallavicini, A., Venier, P., Novoa, B., & Figueras, A. (2010). MgC1q, a novel C1q-domain-containing protein involved in the immune response of Mytilus galloprovincialis. *Developmental and comparative immunology*, *34*(9), 926–34. doi:10.1016/j.dci.2010.02.012

Gilmour, A. R., Gogel, B. J., Cullis, B. R., and Thompson, R. (2009). ASReml User Guide Release 3.0. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK www.vsni.co.uk

Gómez-León, J., Villamill, L., Salger, S. A., Sallum, R. H., Remacha-Triviño, A., Leavitt, D. F., & Gómez-Chiarri, M. (2008). Survival of eastern oysters *Crassostrea virginica* from three lines following experimental challenge with bacterial pathogens. *Diseases of Aquatic Organisms*, *79*(2), 95–105. doi:10.3354/dao01902

Hada, H. S., West, P. A., Lee, J. V., Stemmler, J., & Colwell, R. R. (1984). *Vibrio tubiashii* sp. nov., a pathogen of bivalve mollusks. *International Journal of Systematic Bacteriology*, *34*(1), 1–4. doi:10.1099/00207713-34-1-1

Hasegawa, H., Lind, E., Boin, M., & Häse, C. (2008). The extracellular metalloprotease of *Vibrio tubiashii* is a major virulence factor for Pacific oyster (*Crassostrea gigas*) larvae. *Applied and Environmental Microbiology*, *74*(13), 4101–10. doi:10.1128/AEM.00061-08

Hedgecock, D., Li, G., Hubert, S., Bucklin, K., & Ribes, V. (2004). Widespread null alleles and poor cross-species amplification of microsatellite DNA loci cloned from the Pacific oyster, *Crassostrea gigas*. *Journal of Shellfish Research*, *23*(2), 379–385.

Hoffmann, J. A. (1999). Phylogenetic perspectives in innate immunity. *Science*, *284*(5418), 1313–1318. doi:10.1126/science.284.5418.1313

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics*, *6*(2), e1000862. doi:10.1371/journal.pgen.1000862

Huang, R., Jaritz, M., Guenzl, P., Vlatkovic, I., Sommer, A., Tamir, I. M., … Pauler, F. M. (2011). An RNA-Seq strategy to detect the complete coding and non-coding transcriptome including full-length imprinted macro ncRNAs. *PLoS ONE*, *6*(11), e27288. doi:10.1371/journal.pone.0027288

Hubert, S., & Hedgecock, D. (2004). Linkage maps of microsatellite DNA markers for the Pacific oyster *Crassostrea gigas*. *Genetics*, *168*(1), 351–62. doi:10.1534/genetics.104.027342

Itoh, N., & Takahashi, K. G. (2008). Distribution of multiple peptidoglycan recognition proteins in the tissues of Pacific oyster, *Crassostrea gigas*. *Comparative Biochemistry and Physiology, Part B*, *150*(4), 409–17. doi:10.1016/j.cbpb.2008.04.011

Itoh, N., & Takahashi, K. G. (2009). A novel peptidoglycan recognition protein containing a goose-type lysozyme domain from the Pacific oyster, *Crassostrea gigas*. *Molecular Immunology*, *46*(8-9), 1768–74. doi:10.1016/j.molimm.2009.01.022

Karin, M., & Ben-neriah, Y. (2000). Phosphorylation meets ubiqutination: the control of NF-κB activity. *Annual review of immunology*, *18*, 621–663.

Kim, S., & Misra, A. (2007). SNP genotyping: technologies and biomedical applications. *Annual Review of Biomedical Engineering*, *9*, 289–320. doi:10.1146/annurev.bioeng.9.060906.152037

Kim, S. Y., Lohmueller, K. E., Albrechtsen, A., Li, Y., Korneliussen, T., Tian, G., … Nielsen, R. (2011). Estimation of allele frequency and association mapping using next-generation sequencing data. *BMC Bioinformatics*, *12*(1), 231. doi:10.1186/1471-2105-12-231

Kirst, M., Cordeiro, C. M., Rezende, G. D. S. P., & Grattapaglia, D. (2005). Power of microsatellite markers for fingerprinting and parentage analysis in *Eucalyptus grandis* breeding populations. *Journal of Heredity*, *96*(2), 161–6. doi:10.1093/jhered/esi023

Lang, R., Bayne, C., & Camara, M. (2009). Transcriptome profiling of selectively bred Pacific oyster *Crassostrea gigas* families that differ in tolerance of heat shock. *Marine Biotechnology*, *11*(5), 650–68. doi:10.1007/s10126-009-9181-6

Langdon, C., Evans, F., Jacobson, D., & Blouin, M. (2003). Yields of cultured Pacific oysters *Crassostrea gigas* Thunberg improved after one generation of selection. *Aquaculture*, *220*(1-4), 227–244. doi:10.1016/S0044-8486(02)00621-X

Lannan, J. (1971). Experimental self-fertilization of the Pacific oyster, *Crassostrea gigas*, utilizing cryopreserved sperm. *Genetics*, *68*(4), 599–601.

Launey, S., & Hedgecock, D. (2001). High genetic load in the Pacific oyster *Crassostrea gigas*. *Genetics*, *159*(1), 255–65.

Lemaitre, B., & Hoffmann, J. (2007). The host defense of *Drosophila melanogaster*. *Annual Review of Immunology*, *25*, 697–743. doi:10.1146/annurev.immunol.25.022106.141615

Li, G., Hubert, S., Bucklin, K., Ribes, V., & Hedgecock, D. (2003). Characterization of 79 microsatellite DNA markers in the Pacific oyster *Crassostrea gigas*. *Molecular Ecology Notes*, *3*(2), 228–232. doi:10.1046/j.1471-8286.2003.00406.x

Li, W., Yi, S., & Makova, K. (2002). Male-driven evolution. *Current Opinion in Genetics & Development*, *12*(6), 650–6.

Li, Y., Qin, J. G., Abbott, C. A., Li, X., & Benkendorff, K. (2007). Synergistic impacts of heat shock and spawning on the physiology and immune health of *Crassostrea gigas*: an explanation for summer mortality in Pacific oysters. *American journal of physiology. Regulatory, integrative and comparative physiology*, *293*(6), R2353–62. doi:10.1152/ajpregu.00463.2007

Liu, S., Yeh, C.-T., Tang, H. M., Nettleton, D., & Schnable, P. S. (2012). Gene mapping via Bulked Segregant RNA-Seq (BSR-Seq). *PloS ONE*, *7*(5), e36406. doi:10.1371/journal.pone.0036406

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods*, *25*(4), 402–8. doi:10.1006/meth.2001.1262

Magwene, P. M., Willis, J. H., & Kelly, J. K. (2011). The statistics of bulk segregant analysis using next generation sequencing. *PLoS Computational Biology*, *7*(11), e1002255. doi:10.1371/journal.pcbi.1002255

Main, B. J., Bickel, R. D., McIntyre, L. M., Graze, R. M., Calabrese, P. P., & Nuzhdin, S. V. (2009). Allele-specific expression assays using Solexa. *BMC Genomics*, *10*, 422. doi:10.1186/1471-2164-10-422

Mamanova, L., Coffey, A., Scott, C., Kozarewa, I., Turner, E. H., Kumar, A., … Turner, D. J. (2010). Target-enrichment strategies for next-generation sequencing. *Nature Methods*, *7*(2), 111–8. doi:10.1038/NMETH.1419

Meaburn, E., Butcher, L. M., Liu, L., Fernandes, C., Hansen, V., Al-Chalabi, A., … Schalkwyk, L. C. (2005). Genotyping DNA pools on microarrays: tackling the QTL problem of large samples and large numbers of SNPs. *BMC Genomics*, *6*, 52. doi:10.1186/1471-2164-6-52

Mellroth, P., & Steiner, H. (2006). PGRP-SB1: an *N*-acetylmuramoyl *L*-alanine amidase with antibacterial activity. *Biochemical and Biophysical Research Communications*, *350*(4), 994–9. doi:10.1016/j.bbrc.2006.09.139

Meyer, E., Aglyamova, G. V, & Matz, M. V. (2011). Profiling gene expression responses of coral larvae (*Acropora millepora*) to elevated temperature and settlement inducers using a novel RNA-Seq procedure. *Molecular Ecology*, *20*(17), 3599–616. doi:10.1111/j.1365-294X.2011.05205.x

Michelmore, R. W., Paran, I., & Kesseli, R. V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(21), 9828–32.

Miller, M. B., & Tang, Y.-W. (2009). Basic concepts of microarrays and potential applications in clinical microbiology. *Clinical Microbiology Reviews*, *22*(4), 611–33. doi:10.1128/CMR.00019-09

Miller, M., Dunham, J., & Amores, A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research*, *17*(2), 240–248. doi:10.1101/gr.5681207.high-throughput

Momeni, H. (2011). Role of calpain in apoptosis. *Cell Journal*, *13*(2), 65–72.

Montagnani, C., Kappler, C., Reichhart, J. M., & Escoubas, J. M. (2004). *Cg*-Rel, the first Rel/NF-κB homolog characterized in a mollusk, the Pacific oyster *Crassostrea gigas*. *FEBS Letters*, *561*(1-3), 75–82. doi:10.1016/S0014-5793(04)00124-3

Montagnani, C., Labreuche, Y., & Escoubas, J. M. (2008). Cg-IκB , a new member of the IκB protein family characterized in the pacific oyster *Crassostrea gigas*. *Developmental and Comparative Immunology*, *32*(3), 182–190. doi:10.1016/j.dci.2007.06.001

Mullen, M. P., Creevey, C. J., Berry, D. P., McCabe, M. S., Magee, D. A., Howard, D. J., … Waters, S. M. (2012). Polymorphism discovery and allele frequency estimation using high-throughput DNA sequencing of target-enriched pooled DNA samples. *BMC Genomics*, *13*(1), 16. doi:10.1186/1471-2164-13-16

Neu, H. C. (1969). Effect of beta-lactamase location in *Escherichia coli* on penicillin synergy. *Applied microbiology*, *17*(6), 783–6.

Oliver, J. L., Gaffney, P. M., Allen, S. K., Faisal, M., & Kaattari, S. L. (2000). Protease inhibitory activity in selectively bred families of Eastern oysters. *Journal of Aquatic Animal Health*, *12*(2), 136–145.

Parts, L., Cubillos, F. A., Warringer, J., Jain, K., Salinas, F., Bumpstead, S. J., … Liti, G. (2011). Revealing the genetic structure of a trait by sequencing a population under selection. *Genome Research*, *21*(7), 1131–8. doi:10.1101/gr.116731.110

Pauley, G. B., Van Der Raay, B., & Troutt, D. (1988). Species profiles: life histories and environmental requirements of coastal fishes and Invertebrates (Pacific Northwest), Pacific oyster. *U.S. Fish and Wildlife Service Biological Report*, *82* (11.85). U. S. Army Corps of engineers, TR EL-82.4. 28pp

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Ragone Calvo, L. M., Calvo, G. W., & Burreson, E. M. (2003). Dual disease resistance in a selectively bred eastern oyster, *Crassostrea virginica*, strain tested in Chesapeake Bay. *Aquaculture*, *220*(1-4), 69–87. doi:10.1016/S0044-8486(02)00399-XRagone Calvo, L. M., Calvo, G. W., & Burreson, E. M. (2003). Dual disease resistance in a selectively bred eastern oyster, *Crassostrea virginica*, strain tested in Chesapeake Bay. *Aquaculture*, *220*(1-4), 69–87. doi:10.1016/S0044-8486(02)00399-X

Reimers, M. (2010). Making informed choices about microarray data analysis. *PLoS Computational Biology*, *6*(5), e1000786. doi:10.1371/journal.pcbi.1000786

Roberts, S., Goetz, G., White, S., & Goetz, F. (2009). Analysis of genes isolated from plated hemocytes of the Pacific oyster, *Crassostreas gigas*. *Marine Biotechnology*, *11*(1), 24–44. doi:10.1007/s10126-008-9117-6

Roch, P. (1999). Defense mechanisms and disease prevention in farmed marine invertebrates. *Aquaculture*, *172*(1), 125–145. doi:10.1016/S0044-8486(98)00439-6

Rosa, R. D., de Lorgeril, J., Tailliez, P., Bruno, R., Piquemal, D., & Bachère, E. (2012). A hemocyte gene expression signature correlated with predictive capacity of oysters to survive *Vibrio* infections. *BMC Genomics*, *13*(1), 252. doi:10.1186/1471-2164-13-252

Rumble, S. M., Lacroute, P., Dalca, A. V, Fiume, M., Sidow, A., & Brudno, M. (2009). SHRiMP: accurate mapping of short color-space reads. *PLoS computational biology*, *5*(5), e1000386. doi:10.1371/journal.pcbi.1000386

SAS Institute Inc. (2009). SAS/STAT 9.2 User's Guide (2nd ed.). Cary, NC: SAS Institute Inc.

Sauvage, C., Bierne, N., Lapègue, S., & Boudry, P. (2007). Single nucleotide polymorphisms and their relationship to codon usage bias in the Pacific oyster *Crassostrea gigas*. *Gene*, *406*(1-2), 13–22. doi:10.1016/j.gene.2007.05.011

Schmitt, P., Rosa, R. D., Duperthuy, M., de Lorgeril, J., Bachère, E., & Destoumieux-Garzón, D. (2012). The antimicrobial defense of the Pacific oyster, *Crassostrea gigas*. How diversity may compensate for scarcity in the regulation of resident/pathogenic microflora. *Frontiers in Microbiology*, *3*(May), 160. doi:10.3389/fmicb.2012.00160

Sham, P., Bader, J., & Craig, I. (2002). DNA pooling: a tool for large-scale association studies. *Nature Reviews Genetics*, *3*(11), 862–71. doi:10.1038/nrg930

Sindermann, C. (1988). Vibriosis of larval oysters. (C. J. Sindermann & D. V. Lightner, Eds.) *Developments in Aquaculture and Fisheries Science*, 271–274.

Skelly, D., & Johansson, M. (2011). A powerful and flexible statistical framework for testing hypotheses of allele-specific gene expression from RNA-seq data. *Genome Research*, *21*(10), 1728–37. doi:10.1101/gr.119784.110

Sokolova, I. (2009). Apoptosis in molluscan immune defense. *Invertebrate Survival Journal*, *6*, 49–58.

Sontheimer, R. D., Racila, E., & Racila, D. M. (2005). C1q: its functions within the innate and adaptive immune responses and its role in lupus autoimmunity. *The Journal of investigative dermatology*, *125*(1), 14–23. doi:10.1111/j.0022-202X.2005.23673.x

Sun, L., Zhang, Z., Bailey, T. L., Perkins, A. C., Tallack, M. R., Xu, Z., & Liu, H. (2012). Prediction of novel long non-coding RNAs based on RNA-Seq data of mouse Klf1 knockout study. *BMC Bioinformatics*, *13*, 331. doi:10.1186/1471-2105-13-331

Taris, N., Lang, R. P., & Camara, M. D. (2008). Sequence polymorphism can produce serious artefacts in real-time PCR assays: hard lessons from Pacific oysters. *BMC genomics*, *9*, 234. doi:10.1186/1471-2164-9-234

Taris, N., Lang, R. P., Reno, P. W., & Camara, M. D. (2009). Transcriptome response of the Pacific oyster (*Crassostrea gigas*) to infection with *Vibrio tubiashii* using cDNA AFLP differential display. *Animal Genetics*, *40*(5), 663–77. doi:10.1111/j.1365-2052.2009.01894.x

Terahara, K., & Takahashi, K. (2008). Mechanisms and immunological roles of apoptosis in molluscs. *Current Pharmaceutical Design*, *14*(2), 131–137.

Toung, J. M., Morley, M., Li, M., & Cheung, V. G. (2011). RNA-sequence analysis of human B-cells. *Genome Research*, *21*(6), 991–8. doi:10.1101/gr.116335.110

Travers, M.-A., Basuyaux, O., Le Go ̈c, N., Huchette, S., Nicolas, J.-L., Koken, M., & Paillard, C. (2009). Influence of temperature and spawning effort on *Haliotis tuberculata* mortalities caused by *Vibrio harveyi*: an example of emerging vibriosis linked to global warming. *Global Change Biology*, *15*(6), 1365–1376. doi:10.1111/j.1365-2486.2008.01764.x

Tubiash, H. S., Chanley, P. E., & Leifson, E. (1965). Bacillary necrosis, a disease of larval and juvenile bivalve mollusks. *Journal of Bacteriology*, *90*(4), 1036–1044.

Tubiash, H. S., Colwell, R. R., & Sakazaki, R. (1970). Marine vibrios associated with bacillary necrosis, a disease of larval and juvenile bivalve mollusks. *Journal of Bacteriology*, *103*(1), 271–2.

Vezzulli, L., Previati, M., Pruzzo, C., Marchese, A., Bourne, D. G., & Cerrano, C. (2010). *Vibrio* infections triggering mass mortality events in a warming Mediterranean Sea. *Environmental Microbiology*, *12*(7), 2007–19. doi:10.1111/j.1462-2920.2010.02209.x

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., … Kuiper, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, *23*(21), 4407–14.

Wang, S., Meyer, E., McKay, J. K., & Matz, M. V. (2012). 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nature Methods*, *9*(8), 808–810. doi:10.1038/nmeth.2023

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, *10*(1), 57–63. doi:10.1038/nrg2484

Xu, Y., & Crouch, J. H. (2008). Marker-assisted selection in plant breeding: from publications to practice. *Crop Science*, *48*(2), 391–407.

Yang, H., Hu, E., Cuevas-Uribe, R., Supan, J., Guo, X., & Tiersch, T. R. (2012). High-throughput sperm cryopreservation of eastern oyster *Crassostrea virginica*. *Aquaculture*, *344-349*, 223–230. doi:10.1016/j.aquaculture.2012.03.018

Yuste, J., Ali, S., Sriskandan, S., Botto, M., & Brown, J. S. (2006). Roles of the alternative complement pathway and C1q during innate immunity to streptococcus pyogenes. *Journal of immunology*, *176*(10), 6112–20.

Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., & Xu, F., Yang, P., Zhang, L., Wang, X., … Wang, J. (2012). The oyster genome reveals stress adaptation and complexity of shell formation. *Nature*, *490*(7418), 49–54. doi:10.1038/nature11413

Zhang, K., Li, J. B., Gao, Y., Egli, D., Xie, B., Deng, J., … Church, G. M. (2009). Digital RNA allelotyping reveals tissue-specific and allele-specific gene expression in human. *Nature Methods*, *6*(8), 613–8. doi:10.1038/nmeth.1357

Zhang, Z., Bao, M., Lu, N., Weng, L., Yuan, B., & Liu, Y.-J. (2013). The E3 ubiquitin ligase TRIM21 negatively regulates the innate immune response to intracellular double-stranded DNA. *Nature immunology*, *14*(2), 172–8. doi:10.1038/ni.2492

Zhao, X., Yu, H., Kong, L., & Li, Q. (2012). Transcriptomic responses to salinity stress in the Pacific oyster *Crassostrea gigas*. *PLoS ONE*, *7*(9), e46244. doi:10.1371/journal.pone.0046244

Zheng, Q., & Wang, X.-J. (2008). GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis. *Nucleic Acids Research*, *36*(Web Server issue), W358–63. doi:10.1093/nar/gkn276

Zhu, B., & Wu, X. (2008). Identification of outer membrane protein ompR from rickettsia-like organism and induction of immune response in *Crassostrea ariakensis*. *Molecular Immunology*, *45*(11), 3198–204. doi:10.1016/j.molimm.2008.02.026