

Spring 2016

## Differences in Sockeye Salmon Antibody Composition: Testing The Immunological Imprinting Hypothesis

Maxwell Elliott Chappell

College of William and Mary, [mechappell@email.wm.edu](mailto:mechappell@email.wm.edu)

Follow this and additional works at: <https://scholarworks.wm.edu/etd>



Part of the [Biology Commons](#)

---

### Recommended Citation

Chappell, Maxwell Elliott, "Differences in Sockeye Salmon Antibody Composition: Testing The Immunological Imprinting Hypothesis" (2016). *Dissertations, Theses, and Masters Projects*. Paper 1477068424.

<http://doi.org/10.21220/S21S3F>

This Thesis is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Dissertations, Theses, and Masters Projects by an authorized administrator of W&M ScholarWorks. For more information, please contact [scholarworks@wm.edu](mailto:scholarworks@wm.edu).

Differences in Sockeye Salmon Antibody Composition: Testing the  
Immunological Imprinting Hypothesis

Maxwell Elliott Chappell

Williamsburg, Virginia

Bachelor of Science, College of William and Mary

A Thesis presented to the Graduate Faculty  
of the College of William and Mary in Candidacy for the Degree of  
Master of Science

Department of Biology

The College of William and Mary  
August, 2016



## APPROVAL PAGE

This Thesis is submitted in partial fulfillment of  
the requirements for the degree of

Master of Science



---

Maxwell Elliott Chappell

Approved by the Committee, June, 2016



---

Committee Chair

Professor Patty Zwollo, Biology  
The College of William and Mary



---

Chancellor Professor Lizabeth Allison, Biology  
The College of William and Mary



---

Associate Professor Kurt Williamson, Biology  
The College of William and Mary

## COMPLIANCE PAGE

Research approved by

Institutional Animal Care and Use Committee

Protocol number(s): IACUC-2013-12-17-9139 pxzwol

Date(s) of approval: 12/17/2015-12/17/2016

## ABSTRACT PAGE

Anadromous fish such as sockeye salmon return to their natal streams to spawn, during which they undergo significant physiological changes including the release of cortisol, a known immunosuppressive hormone. Our lab has proposed the Immunological Imprinting Hypothesis, which suggests that juvenile anadromous fish respond to pathogens specific to their natal site by producing protective long lived plasma cells (LLPCs) that constitutively produce antibodies against those pathogens. These LLPCs are believed to be highly cortisol resistant. Thus, fish returning to their natal streams have immunological protection from pathogens found at that specific location. I investigated the Immunological Imprinting Hypothesis through analysis of antibody composition and usage. Since 2009 samples of Sockeye Salmon spleen and anterior kidney have been harvested from two separate salmon runs in Alaska. Using quantitative PCR (qPCR) I examined the relative usage levels of specific  $V_H$  gene families between fish at different locations. To further investigate the “pathogen fingerprint” of given spawning sites, I also performed qPCR analysis in order to compare the pathogen loads of multiple pathogens from different sites, including Bacterial Kidney Disease (*Renibacterium salmoninarum*), Bacterial Coldwater Disease (*Flavobacterium psychrophilum*), and Infectious Hematopoietic Necrosis Virus (IHNV). Analysis of  $V_H$  family usage suggests that differences exist between certain spawning locations not only for selected individual  $V_H$  families, but also for multiple  $V_H$  families analyzed simultaneously. Likewise, pathogen loads and infection rates are found to differ frequently between many spawning sites, while probability of infection is shown to be dependent on location for each pathogen analyzed. Analysis of  $V_H$  usage and pathogen loads suggests several correlations that exist between specific usage patterns and lower pathogenic loads. Greater understanding of spawning fish immune functioning can potentially suggest a method of natural immunization against common fish pathogens and thus protect both farmed and wild populations. These differences in  $V_H$  usage patterns and pathogen infection rates between spawning sites provide strong evidence in support of the Immunological Imprinting Hypothesis.

## TABLE OF CONTENTS

|                  |     |
|------------------|-----|
| Acknowledgements | ii  |
| Dedications      | iii |
| List of Tables   | iv  |
| List of Figures  | v   |
| Introduction     | 1   |
| Methods          | 13  |
| Results          | 34  |
| Discussion       | 63  |
| Appendix         | 80  |
| Bibliography     | 85  |

## ACKNOWLEDGEMENTS

I would like to thank everyone who helped to make this possible. I would like to thank Dr. Patty Zwollo, who welcomed me into her lab and allowed me to work on such an incredibly interesting project. I truly can't thank her enough for all of her kindness, instruction, and patience (especially patience!). Beyond these she taught me a great appreciation for science and has helped point me toward a future involving scientific research. For all of these things and more I would like to thank her.

I would like to thank Lidia Epp for all of her help and support. I lost track of the number of times I went to her with questions and she always tried her best to point me in the right direction. Throughout the study she has held the data we gathered to a high standard and I know that the research is that much stronger because of it.

I would like to thank the members of my committee, Dr. Lizabeth Allison and Dr. Kurt Williamson, for helping point my research in the right direction. I feel that their feedback has helped to make this thesis a much more coherent and focused study.

I would like to thank the current and former members of the Zwollo lab for all of their help in making this possible.

I would like to thank Chris Pallister, as well as all of the others over the years who have helped to collect samples and data, allowing for a fantastic database to work from.

I would like to thank my fellow graduate students and friends for their support and understanding. In many ways they have helped to make the past two years a lot of fun!

Finally, I would like to thank my family for all that they have done; especially my parents. I don't think that I'll ever know everything that they've done for me over the years, but I am truly thankful for their love and support.



I would like to dedicate this thesis to my family

## LIST OF TABLES

|   |    |
|---|----|
| 1. Pathogen Primers, Standard Sequence, V <sub>H</sub> Family Sequences   | 20 |
| 2. MemHCmu, SecHCmu, and Tub Primers  | 28 |
| 3. Summary of V <sub>H</sub> usage comparisons between genders per site   | 43 |
| 4. Results of Hotelling T <sup>2</sup> tests analyzing differences in V <sub>H</sub> usage patterns between sites | 46 |
| 5. Results of statistical tests (ANOVA, Permutation, MANOVA) analyzing pathogen infection rates and loads         | 55 |
| 6. Results of site specific independence tests for pathogen infection rates                                       | 56 |

## LIST OF FIGURES

|   |    |
|---|----|
| 1. Visualization of the process of VDJ recombination of the heavy chain in <i>Oncorhynchus mykiss</i>   | 6  |
| 2. Visualization of the grouping of V <sub>H</sub> gene segment families  | 9  |
| 3. Map of Alaska with sampling locations highlighted  | 14 |
| 4. Analysis of the nucleocapsid gene of multiple IHNV strains from around the world   | 17 |
| 5. Sequence alignments of accepted <i>Oncorhynchus mykiss</i> V <sub>H</sub> families with other species  | 21 |
| 6. Sequence alignments of specific <i>Oncorhynchus mykiss</i> V <sub>H</sub> gene segments with other <i>Oncorhynchus mykiss</i> V <sub>H</sub> gene segments | 24 |
| 7. Formulas used to calculate fold change from collected C <sub>T</sub> values according to Livak and Schmittgen (2001)                                       | 32 |
| 8. Strip charts showing relative expression of MemHCmu and SecHCmu relative to MoK reference fish   | 35 |
| 9. Strip charts showing the distribution of V <sub>H</sub> fold change data according to location   | 38 |
| 10. Bar graph showing the average fold change relative to a control fish for each V <sub>H</sub> family divided by gender                                     | 42 |
| 11. Pie charts showing relative usage patterns of each V <sub>H</sub> family at each geographic location  | 45 |
| 12. Graphs showing frequency and load of IHNV infection   | 47 |
| 13. Graphs showing frequency and load of Rs infection   | 49 |
| 14. Graphs showing frequency and load of Fp infection   | 51 |
| 15. Graph showing frequency of worm infestation and picture illustrating worms  | 54 |
| 16. Correlation table showing relationships between variables measured and table with p-values  | 58 |
| 17. Non-metric dimensional scaling (NMDS) analysis between variables measured   | 60 |

# **Differences in Sockeye Salmon Antibody Composition: Testing the Immunological Imprinting Hypothesis**

## **Introduction**

Anadromous fish such as sockeye salmon return to their natal streams to spawn before subsequently dying. While our mechanistic understanding of how these fish return to their natal streams has increased over the years, the question remains of why they consistently return to the same body of water to spawn when it might involve traveling very long distances. During the spawning journey the immune systems of these fish gradually decline, a factor that could potentially help explain such an unusual life history.

Our lab recently proposed the Immunological Imprinting Hypothesis as a possible explanation for why spawning fish return to their natal bodies of water (Zwollo 2012). This hypothesis suggests that exposure to pathogens early in life can produce immunological memory specific to the unique pathogen “fingerprint” at that spawning site. This means that fish have increased protection from the combined set of pathogens in their natal streams relative to those from foreign environments. Such an immunological bias may produce a fitness advantage that justifies a long spawning journey. The basis of this hypothesis is rooted in the idea of immunological memory. A subset of immune cells known as long-lived plasma cells (LLPC's) are believed to persist for long periods of time and constitute a significant component of fish immunological memory (E. S. Bromage et al. 2004; Kaattari, Bromage, and Kaattari 2005). Interestingly, it has previously been shown that (LL)PCs are maintained in successfully spawned fish despite the widespread suppression of the immune system (Schouten et

al. 2013). This in turn could support a connection between the immune system and the biological motivation behind the spawning journey.

### The *Oncorhynchus* Genus and Anadromy

Anadromous fish, including Sockeye salmon (*Oncorhynchus nerka*), demonstrate a unique and well known life cycle. Upon hatching in freshwater streams or lakes most Sockeye salmon live for one to two years in their natal body of water, growing in size before traveling downstream to the ocean (Quinn and Myers 2004) (French et al. 1976) (Rounsefell 1958). After living and growing in the ocean for one to four years, the fish return to the exact same body of water where they hatched in order to spawn before dying shortly thereafter (Quinn and Myers 2004) (Rounsefell 1958). The ability of Pacific salmon, such as Sockeye salmon, to return to the exact same location with a high degree of accuracy is in part attributed to the phenomenon of olfactory imprinting. Juvenile salmon imprint on the specific chemical composition of their natal bodies of water, thus, when returning to spawn they are able to determine where they should go based on olfactory stimuli (Dittman and Quinn 1996; Ueda 2011).

The journeys to and from the ocean present immense challenges to the health of the animal, and specifically to the immune system. The transition from freshwater to saltwater as juveniles, and saltwater to freshwater as adults, requires significant physiological changes. When this is combined with sexual maturation in adults, it results in high levels of stress in returning fish. Migration from saltwater to freshwater requires changes to the fish physiology that allow for greater regulation of salt ions in the body. In Sockeye salmon this is accomplished through the release of glucocorticoid hormones like cortisol, which is thought to result in upregulation of the  $\alpha 1a$  isoform of the NKA sodium potassium ATPase, which is associated with freshwater migration (Flores et al. 2012). A gradual increase in cortisol levels as fish approach spawning sites has been

previously documented (Baker and Vynne 2014). However, in certain circumstances these hormones are known to adversely affect the immune system in Atlantic salmon (*Salmo salar*) (Gadan et al. 2012), and in Winter flounder (*Pleuronectes americanus* L.) (Carlson, Anderson, and Bodammer 1993) among others. Generally, short term exposure of cortisol has been associated with increased innate immune activity, including lysozyme activity and phagocytic activity in macrophages from the anterior kidney, as well as mobilization of lymphocytes from immune organs such as the spleen (Wang et al. 2005). Chronic stress, however, is associated with suppression of phagocytic and lysozyme activity (Wang et al. 2005), as well as suppression of the T-cell mediated immune response, suppression of lymphocyte mobilization (Dhabhar and McEwen 1997), and suppression of antibody production (Li et al. 2007), among other factors (Dhabhar 2014) (Kusnecov and Rossi-George 2002). This is especially critical at the start of and throughout the spawning journey as the fish are traveling between different environments that may be home to pathogens drastically different from what are normally encountered. It is not surprising then that spawning fish often display signs of infection with a variety of different pathological agents, be they viral, bacterial, fungal, or parasitic. The journey from the ocean to the natal stream can at times be quite long, covering hundreds of miles and pitting the fish against powerful currents and obstacles. Even when the fish reach the spawning ground often times spawning does not occur immediately as they need to fully mature sexually before spawning. For example, at the Lake Dalnee spawning ground in Kamchatka the average life span for sockeye salmon is 15 days (Hartman, Merrell, and Painter 1964). This means that spawning fish experience adverse conditions for prolonged periods of time, thus inhibiting the immune response and potentially increasing the period of time during which infection can occur. When combined these factors lead to a very small percentage of fertilized salmon eggs hatching, surviving to adulthood, and eventually returning to spawn. Recent tagging

studies indicate that smolt-to-adult survival in certain populations of wild Sockeye salmon are less than 0.2%, which does not include pre-smolt mortality (Wood et al. 2012).

### Plasma Cells and Immunological Memory

One of the most interesting capabilities of the immune system is that of immunological memory. Immunological memory refers to the ability of the immune system to recognize pathogens that have been encountered previously and produce an immune response much more quickly and efficiently than at original exposure. There are multiple components of immunological memory, including memory T cells, memory B cells, and long lived plasma cells (LLPCs). In teleost fish, which have limited affinity maturation (somatic hypermutation) and memory B cells (Ma, Ye, and Kaattari 2013), immunological memory is largely comprised of LLPCs, constitutively producing antibodies specific to pathogens encountered previously. Studies have shown that in humans humoral immunity can persist for decades, potentially due to the presence of LLPCs (Amanna, Carlson, and Slifka 2007). During the spawning journey, novel immune responses, including the formation and maturation of new B cells, are thought to be inhibited by chronically high levels of cortisol. Long lived plasma cells, however, appear to be retained in the fish through the spawning journey (Schouten et al. 2013). LLPCs are thought to be resistant to the negative effects of chronic cortisol exposure, even going as far as to represent a pathological condition in humans suffering from autoimmune diseases (Hoyer et al. 2008). Whereas normal autoreactive plasma cells would be inhibited by the presence of immunosuppressive or anti-inflammatory drugs, the resistance of LLPCs means that they continue to produce autoreactive antibodies long after normal plasma cells would have stopped. As long as LLPCs remain functional, there is a component of immunological memory at work and the fish maintains some

protection. However, in the event of exposure to novel pathogens the spawning fish are less able to mount an effective immune response.

It is important to note that teleost fish do not possess bone marrow, the primary site of hematopoiesis in mammals. Instead, the anterior portion of the teleost kidney is the hematopoietic site (Hansen and Zapata 1998). As a result the anterior kidney (a primary immune tissue) and the spleen (a secondary immune tissue) are among the sites of greatest interest for our analysis.

### Antibody Variation

One of the primary avenues of immune function comes in the form of antibodies. Antibodies, also known as immunoglobulins, consist of two heavy polypeptide chains and two light polypeptide chains that, when combined, form the quaternary structure of the protein (Williams and Barclay 1988). Antibodies are broadly divided into categories called classes or isotypes, which differ in their heavy chain constant regions (Figure 1), and are generally found to perform specialized biological functions. For example, humans have five different classes of immunoglobulins, some important for complement activation, others for binding to specific Fc receptors on immune cells. Fish rely heavily on a single isotype; immunoglobulin mu (IgM), but also produce immunoglobulin tau (IgT) (Hansen, Landis, and Phillips 2005), and immunoglobulin delta (IgD) (Wilson et al. 1997). IgT in fish is functionally equivalent to immunoglobulin alpha (IgA) in mammals and both function as part of the mucosal defense against pathogens (Zhang et al. 2010).

The ability of antibodies to bind to an almost infinite array of antigens is attributed in part to the portion of the antibody known as the variable domain (Figure 1). The variable domain is comprised of different gene segments: variations of the same gene that offer different structural variations. These differences in gene segment usage in the



variable domain in turn allow for changes in the structure of the antibody that result in varied binding affinities to select pathogens. When looking at the heavy chain of the IgM isotype, the variable domain consists of three separate parts that contribute to the observed diversity: the variable segment ( $V_H$ ), the diversity segment ( $D_H$ ), and the joining segment ( $J_H$ ) (Figure 1). IgT is unique in having the heavy chain constant region coding for it between the  $D_H$  and  $J_H$  segments (not shown in Figure 1) (Hansen, Landis, and Phillips 2005). As IgM is the primary systemic isotype used by fish it has been the best studied class of Igs (Salinas, Zhang, and Sunyer 2011; E. Bromage 2004).

(A.)

Germline Configuration of Gene



D to J Recombination



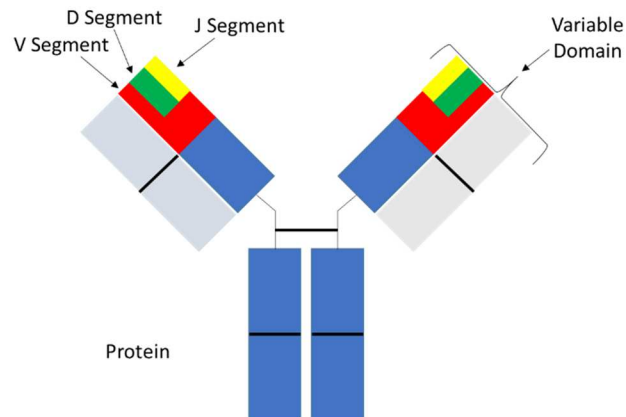
V to DJ Recombination



Transcription, Splicing



(B.)



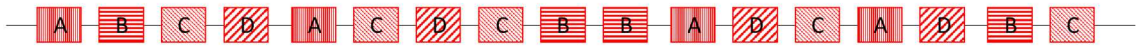
*Figure 1 (A.) Visualization of the process of VDJ recombination of the heavy chain gene in *Oncorhynchus mykiss*. Variable (V) segments are shown in red, diversity (D) segments in green, joining (J) segments in yellow and constant region segments in blue. The germline configuration is shown in the top line, indicating that multiple V, D, and J segments exist. Following D to J recombination a single D and a single J segment are randomly chosen (boxed in teal). Following V to DJ recombination a single V segment is randomly chosen (boxed in teal). Only the selected VDJ segment can be transcribed as the promoter (P) associated with the chosen V segment is in closer proximity to the constant region enhancer (E) following removal of J segments. (B). Schematic representation of an immunoglobulin molecule showing the heavy and light chains (blue and grey respectively). The variable domain is highlighted and the V segment is colored red, the D segment is colored green, and the J segment is colored yellow. Adapted from (“VDJ Recombination | Laika’s MedLibLog” 2014)*

For each  $V_H$  gene segment, different combinations of  $D_H$  and  $J_H$  segments can be used, and the possible combinations using these three types of segments lead to even more variation in antibody structure (Castro et al. 2013). Previous work by others has

suggested that in response to viral challenge fish not only experience a dramatic clonal expansion of B cells, but also a shift in  $V_H$  family usage. For example, previous studies have shown that in response to challenge with Viral Hemorrhagic Septicemia Virus (VHSV), expression of  $V_H3$ ,  $V_H4$ ,  $V_H5.1$ , and  $V_H9$  gene segments were all strongly affected relative to non-infected controls (Castro et al. 2013). This indicates that certain  $V_H$  gene segments were more effective against viral infection – resulting in increased expression – while others were less effective – resulting in a relative decrease in expression. In Rainbow trout (*Oncorhynchus mykiss*) 13  $V_H$  gene families have been defined that represent different versions of the  $V_H$  segment (Roman et al. 1996) (Brown, Kaattari, and Kaattari 2006). Currently, 57  $V_H$ , 9  $D_H$ , and 7  $J_H$  gene segments have been identified in Rainbow Trout according to the International Immunogenetics Information System (Lefranc et al. 2009). The 13  $V_H$  families are groupings of  $V_H$  gene segments based on sequence homology, with a >80% homology being indicative of members of the same family, while members of different families typically have <70% homology (Brodeur and Riblet 1984) (Figure 2).

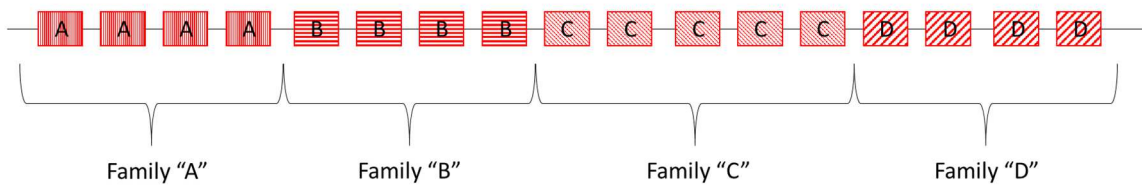
(A.)

Location of  $V_H$  Segments on the Genomic DNA



(B.)

$V_H$  Family Groupings by Sequence Homology



*Figure 2: Visualization of the grouping of  $V_H$  gene segment families. (A.) In the germline configuration  $V_H$  gene segments are randomly interspersed; however, according to sequence homology these  $V_H$  gene segments can be grouped together (in this figure by the pattern and letter seen on each segment) into families (shown in B).*

The genus *Oncorhynchus* is made up of the Pacific salmonids including Chum (*Oncorhynchus keta*), Chinook (*Oncorhynchus tshawytscha*), and Sockeye salmon (*Oncorhynchus nerka*), as well as Rainbow trout (*Oncorhynchus mykiss*). Many species in this genus exhibit the anadromous life history. It is known that members of the genus *Oncorhynchus* are highly related based on mitochondrial sequence analysis (Domanico and Phillips 1995). When comparing Sockeye salmon mitochondrial DNA to Rainbow trout mitochondrial DNA there is only about a 9% sequence divergence in the ATPase 6 gene (Domanico and Phillips 1995). Clearly there is a high degree of relatedness between seemingly different species of the genus *Oncorhynchus*, suggesting that analysis of *Oncorhynchus nerka* DNA sequences could logically begin through comparison to other members of the genus. Likewise, it has been demonstrated that

even in fish of different genera the heavy chain variable region (encoded by  $V_H$ ,  $D_H$ , and  $J_H$  segments) is highly conserved, so it is reasonable to assume that the variable domain in Rainbow trout would be similar to the variable domain in Sockeye salmon (Andersson and Matsunaga 1998). While it does not represent all of the variation that occurs in antibody structure, analysis of  $V_H$  family expression is an important first step toward determining whether there are differences in pathogenic challenges between spawning locations. As antibodies are produced in response to pathogens, fish responding to similar pathogens would be expected to react similarly.

### Pathogen Analysis

While observed variation in  $V_H$  segment usage could reasonably be attributed to differences in the pathogens that the fish are being exposed to, a direct analysis of the pathogens fish might encounter could provide additional useful information. Sockeye salmon must potentially deal with a wide range of pathogens before having the opportunity to reproduce. For Sockeye salmon in Alaska the most important pathogens include, *Renibacterium salmoninarum* (Rs) causing Bacterial Kidney Disease (BKD), *Aeromonas salmonicida* (AS) causing Furunculosis, *Ichthyobodo necator* (IN) causing Ichthyobodiasis, *Flavobacterium psychrophilum* (Fp) causing Bacterial Coldwater Disease (BCWD), and Infectious Hematopoietic Necrosis Virus (IHNV), causing Infectious Hematopoietic Necrosis (Meyers et al. 2008). Rs is an intracellular Gram-positive bacterium that produces lesions on fish and results in a distended abdomen. AS is a Gram-negative non-motile rod that produces visible lesions on fish infected with it Fp is a widespread bacterial pathogen affecting all salmonids; it is a Gram-negative rod that occurs frequently in very cold water with optimal growth between 15°C and 20°C (Bernardet and Kerouault 1989; Sugahara et al. 2010). IN is a kinetoplastid ecto-parasite that if left untreated can result in severe mortality in a variety of fish. IHNV is a

rhabdovirus, meaning it is a single-stranded negative sense RNA virus (Meyers et al. 2008; Eiras et al. 2008). IHNV in particular is an especially significant pathogen as it leads to severe morbidity in many cases and mortality rates approaching 100% in juveniles, while causing asymptomatic infection in adults (Kibenge et al. 2012). IHNV has been a problem of increasing importance in aquaculture and much effort has gone into developing more effective ways of dealing with IHNV outbreaks, including the development of vaccines (Purcell et al. 2004). Increasing prevalence of the disease in aquaculture settings can lead to increased infection rates in wild fish, thus reducing the occurrence of natural fish populations. As salmon progress toward their spawning grounds and their immune systems become more compromised they are likely faced with these as well as other pathogens.

#### Experimental Approach

Two specific aims were addressed in this thesis research:

- 1. Do anadromous fish, Sockeye salmon, demonstrate different antibody specificities at different locations?**
  - a. Does the usage of  $V_H$  gene families vary at different spawning sites?**
  - b. Does the usage of  $V_H$  gene families vary between two sites from the same spawning run?**

To test the validity of the Immunological Imprinting Hypothesis we proceeded via several different avenues. If true, we expected to find different antibody expression ( $V_H$ ) patterns in anterior kidney from fish at different spawning sites, while fish at the same site would have similar  $V_H$  patterns. According to the hypothesis this would, at least partially, occur as a result of the different “pathogen fingerprints” at each spawning site. We predicted that different  $V_H$  families would have a different abundance between sites.

For example,  $V_H$  5.1 usage could be more abundant at site 1 compared to site 2, while  $V_H$  1.1 may be more abundant at site 2 compared to site 1. In conjunction with previous work on this subject (Schouten et al. 2013), the relative expression of membrane bound and secreted heavy chain mu transcripts was evaluated in order to establish differences in B cell and plasma cell abundance and to provide a reference to compare various  $V_H$  family expression rates. Analysis of  $V_H$  family usage in fish at locations from the same spawning run could indicate that despite having the same origin, the spawning location can still be significantly different.

## **2. Do Sockeye salmon demonstrate a different prevalence of infection with common Sockeye salmon pathogens at different locations?**

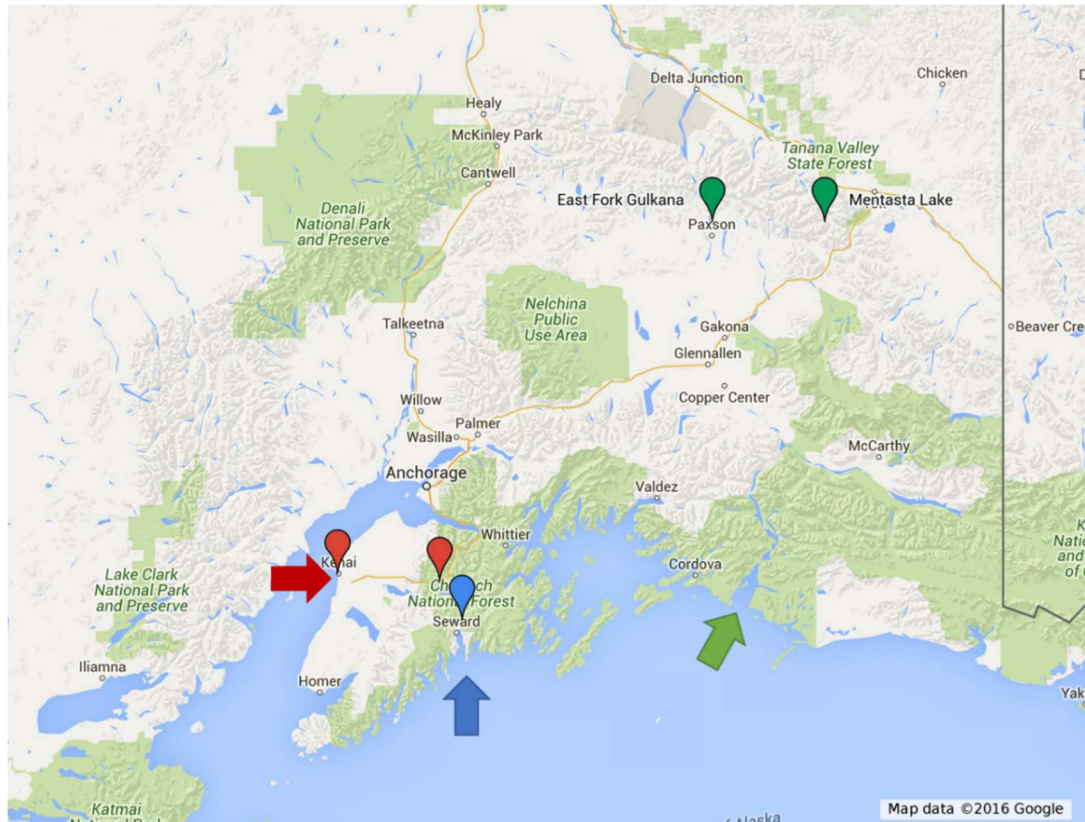
When approaching the “pathogen fingerprints” themselves, we expected to find differential representation of pathogens in the tissues of fish from different sites. These observed differences would theoretically not be present if all of the pathogens at different spawning sites were the same, or if it didn’t matter how the fish immunologically responded to them. Site-specific differences in average antibody ( $V_H$ ) expression and pathogen loads could therefore be seen as evidence in support of the Immunological Imprinting Hypothesis.

## Methods

### Collection of Samples

Since 2009, adult Sockeye salmon have been collected between the months of June and August at various geographic locations along the Kenai Peninsula in Alaska, including the Mouth of the Kenai River (MoK) (Latitude/Longitude: 60.55168/-151.27418), Quartz Creek (QC) (Latitude/Longitude: 60.49562/-149.69714), Mentasta Lake (ML) (Latitude/Longitude: 63.12460/-143.75423), East Fork Gulkana (EFG) (Latitude/Longitude: 63.12460/-145.49274), and Bear Lake (BL) (Latitude/Longitude: 60.1985/-149.35525) (Figure 3). Fish were captured using the “snagging” method at QC, ML, EFG, and BL, or were donated by local fishermen (at MoK). Fish were immediately euthanized via cerebral concussion. Approximately 100mg of desired tissues, including anterior kidney and spleen, were immediately removed and placed in 1.5mL microcentrifuge tubes containing 700µl of RNAlater® Stabilization Solution (Ambion, # AM7020). Following removal of relevant tissues, fish remains were returned to the water to minimize impact on the environment. RNALater stored samples were frozen at -20°C after approximately six hours at 4°C, and sent back to the College of William and Mary in Williamsburg, Virginia on dry ice for long term storage in a -20°C or -80°C freezer.





*Figure 3: Map of Alaska with sampling locations highlighted. Those sites labeled in red are part of the Kenai River run, in green are part of the Copper River run. The single site labeled in blue is associated with the Resurrection peninsula run. Arrows indicate the likely location where adult fish enter the river associated with the like-colored sampling site(s).*

### RNA Extraction

Total RNA was isolated using RNeasy RT (Molecular Research Centers, Inc). Frozen tissue samples stored in RNeasy were thawed and approximately 50mg of tissue was placed in 2mL lysing tubes containing 1mL of RNeasy RT. The type of lysing tube varied based on tissue type; lysing tubes with no beads were acceptable for the softer anterior kidney, while bead-containing Lysing Matrix F (MP Biomedicals) tubes were used for the harder spleen tissue. Tissues were subsequently homogenized using an Omni Beadruptor 24 (Omni International), with one 20 second cycle at speed 5.

Samples were immediately placed on ice for approximately 1 minute to minimize any heat damage during homogenization. Samples were then placed at room temperature. One by one 400µl of UltraPure™ DNase/RNase-Free Distilled Water (Invitrogen, Inc.) was added to the tubes, and one by one the tubes were shaken vigorously. The samples were allowed to incubate for 15 minutes at room temperature before being spun at 12000rcf at room temperature for 15 minutes using a 5430R centrifuge (Eppendorf AG). Following centrifugation lysing tubes were carefully removed from the centrifuge and approximately 1mL of each individual sample was moved to a clean microcentrifuge tube containing 600µl of a 75% ethanol mixture using Absolute 200 Proof Ethyl Alcohol (Pharmco-AAPER, #111000200) and shaken vigorously. Next, samples were allowed to incubate for 10 minutes at room temperature before being centrifuged at 12000rcf for 10 minutes at room temperature. The supernatant was removed and 600µl of 75% ethanol added to each tube to detach and wash the RNA pellet. Samples were then centrifuged at 10000rcf for 2 minutes. This wash step was repeated once more. After the second wash the supernatant was again poured off and the microcentrifuge tubes were inverted and left to dry for 30-45 minutes before being resuspended in 30-60µl of molecular grade water. RNA concentration was then measured using a Nanodrop ND-1000 Spectrophotometer (Thermo Fisher Scientific Inc.). The RNA was then stored at -80°C for future use.

### DNA Isolation

DNA isolation was performed using DNAzol (Molecular Research Center, Inc., #DN 127). Tissue samples were thawed and approximately 50mg of tissue was added to lysing tubes containing 1mL of DNAzol. Anterior kidney samples were disrupted in lysing tubes without beads, while spleen samples used the Lysing Matrix F tubes with beads.

Tissue samples in lysing tubes were then homogenized using the beadruptor with one 20 second cycle at speed 5. Samples were immediately placed on ice.

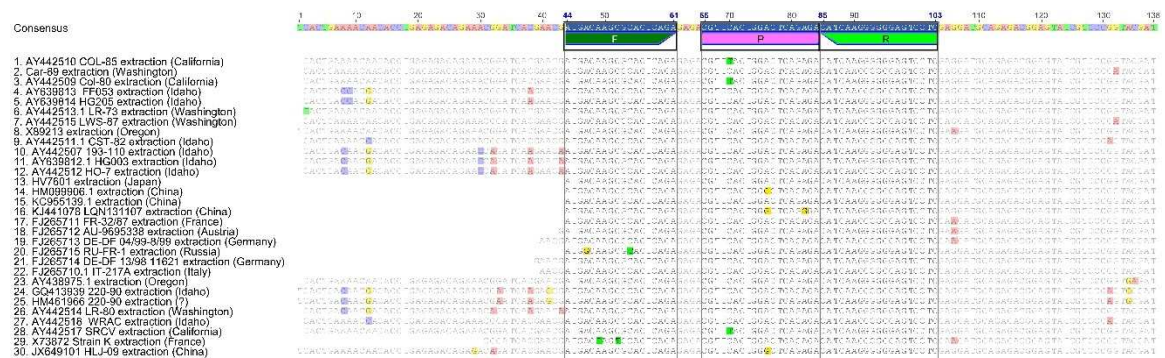
Next the homogenates were incubated at room temperature for 10 minutes. Samples were then centrifuged at 10000rcf at room temperature for 10 minutes. Upon completion as much of the supernatant as possible was transferred to a clean 1.5mL microcentrifuge tube containing 500 $\mu$ L of 100% ethanol. After addition of the supernatant each sample was immediately shaken vigorously. The samples were allowed to incubate for 3 minutes at room temperature before being centrifuged at room temperature for 5 minutes at a speed of 5000rcf. Following centrifugation the supernatant was removed. 800 $\mu$ L of 75% ethanol was added and the tubes were gently shaken, before being centrifuged again at 1000rcf for 2 minutes at room temperature. This wash step was repeated once more before the tubes were inverted and allowed to dry for 15-30 minutes. After this drying period the pellets were resuspended in 30-60 $\mu$ L of molecular grade water and the DNA concentration was measured using the Nanodrop. The DNA was then stored at -80°C for future use.

#### cDNA Synthesis

All work was done on ice to maintain stability of the RNA and reverse transcriptase enzyme. 1 $\mu$ g of RNA was added to 4 $\mu$ L of iScript™ Reverse Transcriptase Supermix for RT-qPCR (Bio-Rad Laboratories, Inc., #1708841) and RNase free water to a maximum volume of 20 $\mu$ L in a labeled PCR tube per sample. The PCR tubes were incubated in a 2720 Thermal Cycler (Applied Biosystems) under the following conditions: 25°C for 5 minutes, 42°C for 60 minutes, 85°C for 5 minutes, and 4°C for storage. 4-5 aliquots of 4-5 $\mu$ L cDNA each were stored at -80°C for future use.

#### TaqMan Assays

TaqMan assays for pathogen detection were taken from relevant literature for specific pathogens. Pathogens investigated were: Fp (Marancik and Wiens 2013), Rs (Sandell and Jacobson 2011), AS (Keeling et al. 2013), and IN (Isaksen et al. 2012). The assay for IHNV was created based on nucleocapsid (N-gene) sequence information submitted to Invitrogen. This sequence information was acquired from the National Center for Biotechnology Information and the Infectious Hematopoietic Necrosis Virus database (“IHN Virus Database” 2016) and chosen as being the most likely and most conserved sequence of the IHNV N-gene that would be present in Alaskan isolates of the virus. IHNV sequence comparison and assay components are shown in Figure 4.



**Figure 4. Analysis of the nucleocapsid gene of multiple Infectious Hematopoietic Necrosis Virus strains from around the world. Forward and reverse primers are highlighted in green with the forward labelled “F”, the reverse labeled “R”, and the probe in pink labeled “P”. The sequence submitted to Invitrogen was targeted to maximize homology between all strains, but primarily those that are geographically close to Alaska, represented by the U genotype as described in (Kurath 2003). Sequences above list first the GenBank Accession number, followed by the strain name of the given isolate, and lastly the location given in parentheses.**

All pathogen assays were ordered as Single Tube Custom TaqMan® Gene Expression Assays (Applied Biosystems), and all assays used a FAM reporter, NFQ

quencher, and a ROX reference. Organization and visualization of TaqMan assays were aided by the use of the program Geneious "Geneious Restricted 8.1.7 (<http://www.geneious.com>, (Kearse et al. 2012))". Primer set sequences and target amplicons are listed in Table I.

### Standardization of Assay

To successfully perform TaqMan qPCR and allow for the calculation of copy numbers a standard containing a desired number of amplicons for each assay was designed. Using the previously published TaqMan assays the amplicons were taken and arranged in a single continuous sequence separated by four thymine bases. Again, the software Geneious was used for the creation and visualization of the sequence. Each amplicon target was tested for specificity using BLAST (Basic Local Alignment Search Tool) (Altschul et al. 1990). After specificity was confirmed the sequence for the standard was ordered via gBlocks® Gene Fragments (Integrated DNA Technologies, Inc.). The standard contained amplicons for five potential pathogens, IHNV, *Renibacterium salmoninarum* (Rs), *Flavobacterium psychrophilum* (Fp), *Aeromonas salmonicida* (AS), and *Ichthyobodo necator* (IN) (Table I). IN was not used extensively in this study, because this pathogen is most abundant in the gills and mucus but not necessarily the spleen (Isaksen et al. 2012). Further, the amplicon for AS was subsequently used as the base sequence for the creation of a Single Tube Custom TaqMan® Gene Expression Assay (Applied Biosystems). As a result the sequence for AS on the standard is not the exact length of the amplicon created during PCR, but it was still specific to AS itself. Upon receipt of the 500ng of standard it was reconstituted in 500µl of TE buffer (10mM Tris pH 8, 1mM EDTA pH 8), resulting in a concentration of 1ng/µl. The length of the standard was 503 base pairs. Using the conversion of  $1.096 \times 10^{-21}$  g/bp, copy number was determined ("Creating Standard Curves with Genomic DNA or Plasmid Templates

for Use in Quantitative PCR - Quant\_pcr.pdf" 2016). After calculation, dilutions were created using molecular grade water to create standard stocks of various concentrations from the original TE Buffer stock of  $1.814 \times 10^9$  copies/ $\mu\text{L}$  to 1 copy/ $\mu\text{L}$

**Table I – Pathogen Primers, Standard Sequence, V<sub>H</sub> Family Sequences**

| Pathogen TaqMan Assays                  |  |            |                               |                               |                        |                   |
|---|--|------------|-------------------------------|-------------------------------|------------------------|-------------------|
| Disease                                 | Pathogen Name                                  | Assay Name | Forward Primer Sequence 5'-3' | Reverse Primer Sequence 5'-3' | Probe Sequence 5'-3'   | Reporter/Quencher |
| Infectious Hematopoietic Necrosis (IHN) | Infectious Hematopoietic Necrosis Virus (IHNV) | NTAQMAN1   | ATGACAAGCGCACTCAGA            | GAGGACTCCCCCTTGATG            | TCTCTGAGTCCAGTGAACG    | FAM/NFQ           |
| Bacterial Coldwater Disease (BCWD)      | <i>Flavobacterium psychrophilum</i> (Fp)       | BCWD       | TTTCTGCCACCTAGCGAATACC        | GGTAGCGGAACCGGAAATG           | CGTTCCTGAGCCAGA        | FAM/NFQ           |
| Bacterial Kidney Disease (BKD)          | <i>Renibacterium salmoninarum</i> (Rs)         | BKD1       | GTGACCAACACCCAGATATCCA        | TCGCCAGACCACCATTTACC          | CACCAGATGGAGCAAC       | FAM/NFQ           |
| Ichthyobodiasis                         | <i>Ichthyobodo necator</i> (IN)                | ICHT       | ACGAACTTATGCGAAGGCA           | TGAGTATTCACCTCCGATCCAT        | TCCACGACTGCAAACGATGACG | FAM/NFQ           |
| Furunculosis                            | <i>Aeromonas salmonicida</i> (AS)              | FURUN      | GTTGGTGCTTCTACTGCTG           | CACCTGATTGGTTTCAACTTCAAA      | CCGTCAGGCTCGCTTG       | FAM/NFQ           |

**GeneBlocks qPCR Standard PSTD1**

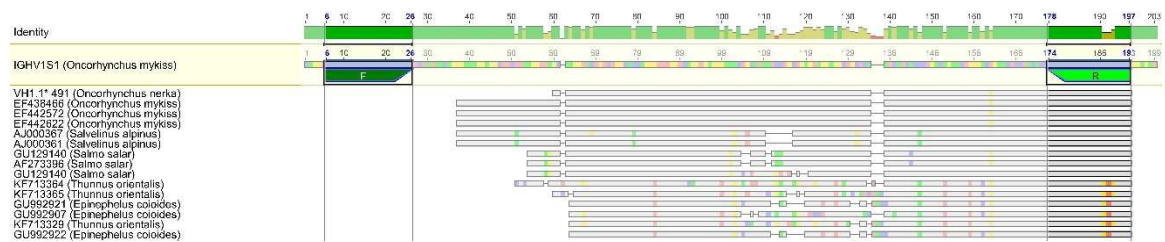
| Sequence 5'-3'   |
|--|
| TTTCTGCCACCTAGCGAATACCTACACCACCGCCATTTCCGCTTCCTGAGCCAGAGCATTCCGGTTCGCTACCTTTTATGACAAGCGCACTCAGAGAGACGTTCACTGGACTCAGAGACATCAAGGGGGAGTCTCTTTTACGAACTTATGCGAAGGCATTTCCAAGTATACCTTCATCAATCAAGAACTAAAGTGTGGAGATCGAAGATGATTAGATACCATTTGATCCAGCACTGCAAAACGATGACGCCCATGGATCGGGAGTGAATACTCATTTTGTGACCAACA<br>CCCAGATATCCATGCACCATGAGGAGCAACTCCGTTACTGGTAAATGGTGGTCTGGCGATTTTAAAGCACTGTCTGTACCTGCCAAGCGGTGGTGAAGTACTCTGAAGCCTGCTGATGTTGCTGCTGACGTTGGTCTCTA<br>TCACTGCTGGCCGTCAGGCTCGCTTGGTGTGAAAGTTGAAACCAATCAGGGTGAAGTAGC |

| VH Family Primer Sets |             |                         |                            |  |                 |
|-----------------------|-------------|-------------------------|----------------------------|--|-----------------|
| VH Family             | Primer Name | Forward Primer 5'-3'    | Reverse Primer 5'-3'       | Amplicon 5'-3'   | Amplicon Length |
| VH1.1                 | VH1.1IR.F   | GAATCCCAAACTGACCTGT     | TGCTGCTTGCTGTTGCTCT        | GAATCCCAAACTGACCTGTACATATTCTGGGTTTGGTGGCAACAATTGGAATGCTTGGATCAGACAGGCTCCTGGGAA                                 | 188             |
|                       | VH1.1IR.R   |                         |                            | AGGACTGGAGTGGGTTTCAGCTATTACAGCAGTGTGGTGGCAGCACTTACTACTCCAGTCTGTTTCAGGGTCAGTTCA<br>CCATCTCCAGAGACAACAGCAAGCAGCA |                 |
| VH2                   | VH2IR-G.F   | ACTGGATACGACAGCCAGCA    | TTTAGTGGTTTCCATACGTCCT     | ACTGGATACGACARCCAGCAGGAAATCTCTGGAGTGGATGGGTTWTTCTGGATTAGGCCTAGGTTACCATGCTAAAAG                                 | 108             |
|                       | VH2IR-G.R   |                         |                            | GTTTGAAGGACGTATGGAAACCACTAAA   |                 |
| VH5.1                 | MCVH5.1IR.F | ACCTCTGGGTTACATTAGT     | GACTCGATCCACCACTGTGTA      | ACCTCTGGGTTACATTAGTACCTATGCTTGGATGAAACAGGCTCCTGGGAAAGGACTTGAGTTTGTGCTTT  | 103             |
|                       | MCVH5.1IR.R |                         |                            | ATACACACTGGTGGATCGAGTC   |                 |
| VH9                   | VH9IR-B.F   | ATCCGTCAGAAACAGGGGAAA   | CACAATAATAAACAGCAGAGTCTTCA | ATCCGTCAGAAACAGGGGAAAGGTCTAGAGTGGATTGGACGGATTGACACTGGCACTGGCACTGTATTTGCTCAGTCCCT                               | 176/178         |
|                       | VH9IR-B.R   |                         |                            | CCAGGGCCAGTTCACCATCACCAGGACAACTCCAAAACCAAGTTGTATTTAGAGGTGAAAAGCCTGAAGACTGAAGACT<br>CTGCTGTTTATTATTGTG          |                 |
| VH10                  | VH10IR.F    | CCTGCAAAATCTCTGGGTTTCTA | TCATATCTAGAAGTATCTTGTCTCT  | CCTGCAAAATCTCTGGGTTTCTATATCTTCTTATTGTCTTGGATGGATACGACAAGCTGAAGGGAAAGCTTTGGAGCATG                               | 151             |
|                       | VH10IR.R    |                         |                            | TTGGTGTACATGCAGTAGTAGTACTACACTGTAGACTCACTAAAGAGCAAGATCACTCTAGATATGA  |                 |
| VH12                  | MCVH12IR.F  | CACCCAACCAAGCTCATTG     | TGGTCCAAATCCATTCAAGTG      | CACCCAACCAAGCTCATTGGATGTAAGGCTGGAGAGTCTTTGAGCATCACTGTAAGGTCTCTGGTTATTCTCTGACTGA                                | 152             |
|                       | MCVH12IR.R  |                         |                            | TAGCAGCAACGTCATGGAATAGGCTGGATCCGTCAGCTGTAGGAAAAGCACTTGAATGGATTGGACCA   |                 |

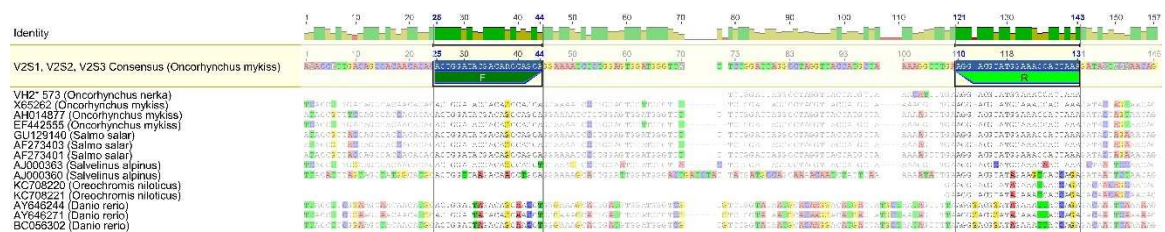
## Primer Design for V<sub>H</sub> Families

Primer design for SYBR Green Analysis began with previously established V<sub>H</sub> families in Rainbow trout as published in Castro et al. (2013) and available online at [imgt.org](http://imgt.org) (Lefranc et al. 2009). V<sub>H</sub> sequences were saved to Geneious for future use. To ensure compatibility with Sockeye salmon sequences the V<sub>H</sub> sequences were analyzed using NCBI BLAST. Sequences were analyzed manually for specific sites on the sequence that displayed significant homology between multiple species, including *Salvelinus alpinus*, *Salmo salar*, and *Thunnus orientalis* (Figure 5).

(A.)

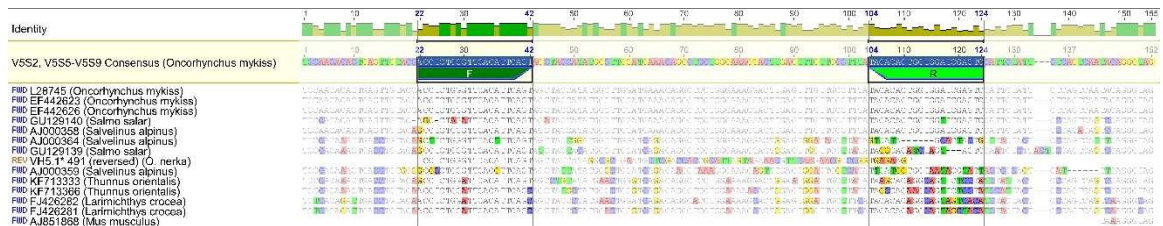


(B.)

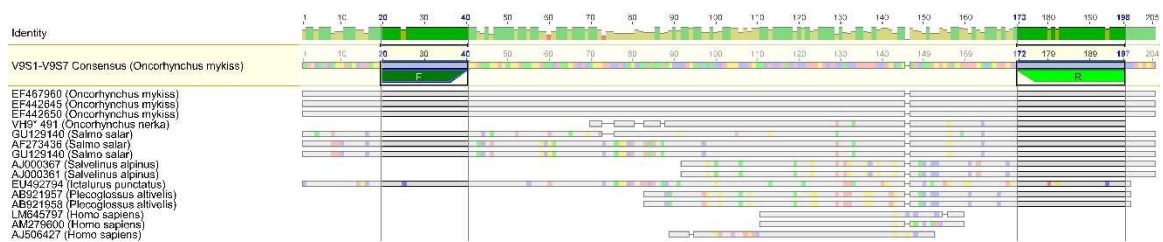


(C.)

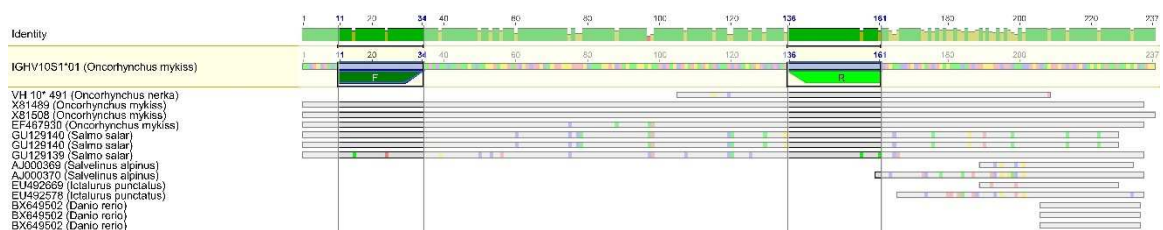




(D.)



(E.)



(F.)

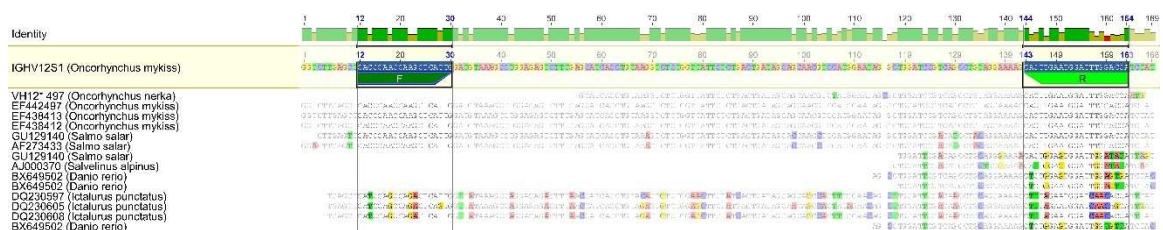
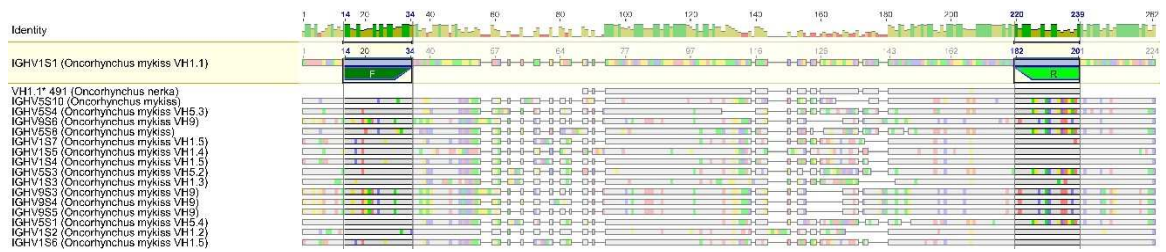


Figure 5: Designing primers for sockeye salmon  $V_H$  families. Sequence alignments of accepted *Oncorhynchus mykiss*  $V_H$  families with other species. Sequences are aligned according to the reference sequence at the top. Forward and reverse primers used for analysis are highlighted in green as “F” for the forward primer and “R” for the reverse

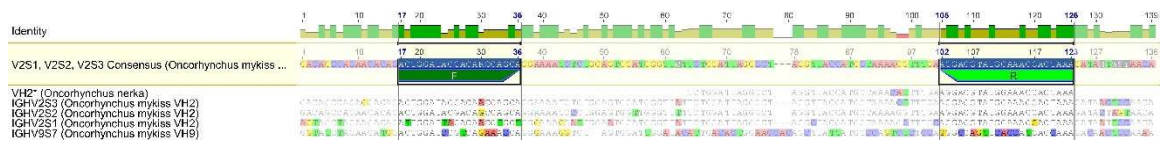
*primer. Disagreements to the reference sequence are highlighted in color (green indicates substitution of thymine relative to the reference, yellow indicates guanine, blue indicates cytosine, and red indicates adenine), while complete homology is shown as grey. Partial Oncorhynchus nerka sequences are included and labeled as such. For each partial sequence 5' and 3' ends were trimmed automatically using Geneious (Except  $V_H 5.1$  which was trimmed manually). O. nerka sample number is listed and corresponds to the fish of the same number listed in Supplemental Table I. (A.)  $V_H 1.1$  sequence analysis. (B.)  $V_H 2$  sequence analysis. (C.)  $V_H 5.1$  sequence analysis. (D.)  $V_H 9$  sequence analysis. (E.)  $V_H 10$  sequence analysis. (F.)  $V_H 12$  Sequence analysis.*

These homologous sequences were then evaluated for their ability to be used as primers through the use of the Primer3Plus tool (Untergasser et al. 2012, 3). Additionally, some trout  $V_H$  sequences were analyzed as a whole using the same Primer3Plus tool, whereupon the tool would suggest potential primers. These sequences were then analyzed in Geneious for the correct level of homology between species. Potential primer sequences were also analyzed for their uniqueness compared to other  $V_H$  families (Figure 6).

(A.)



(B.)

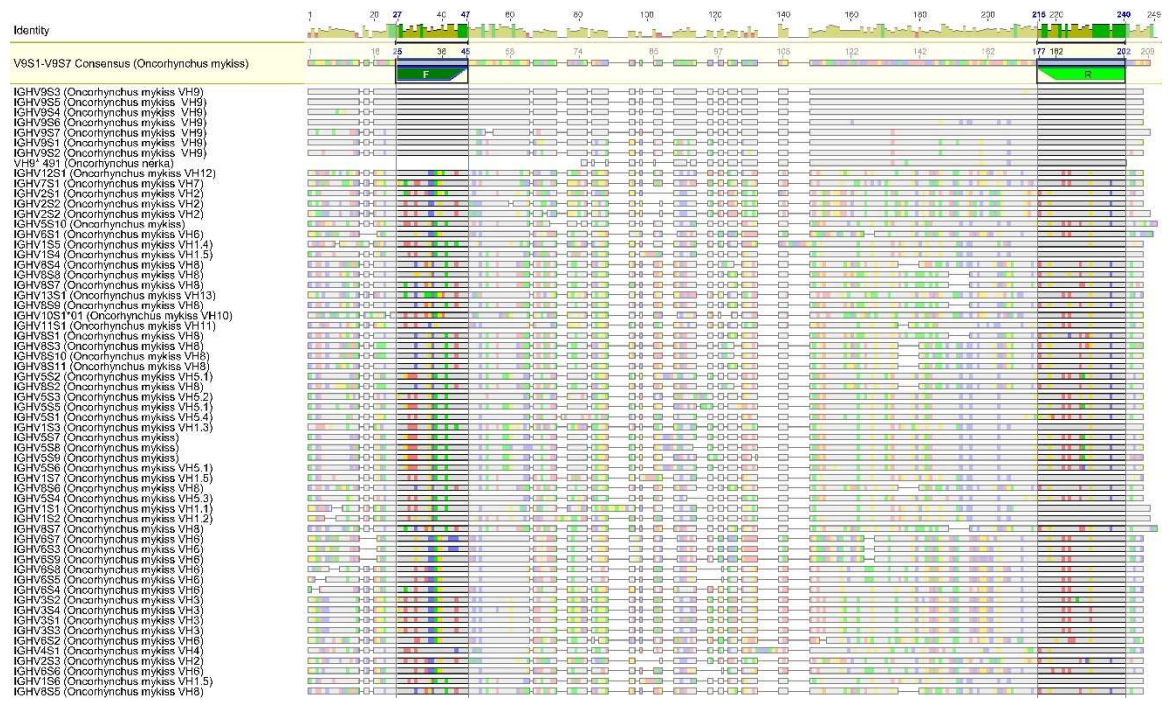


(C.)

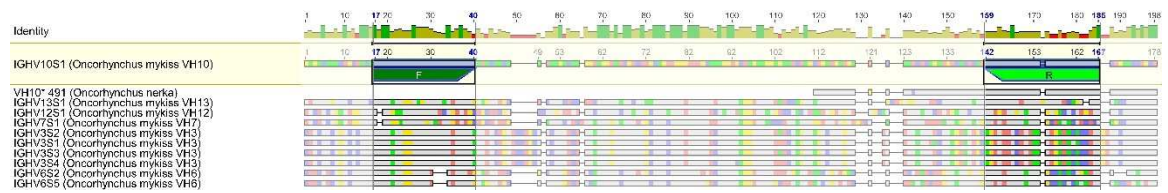




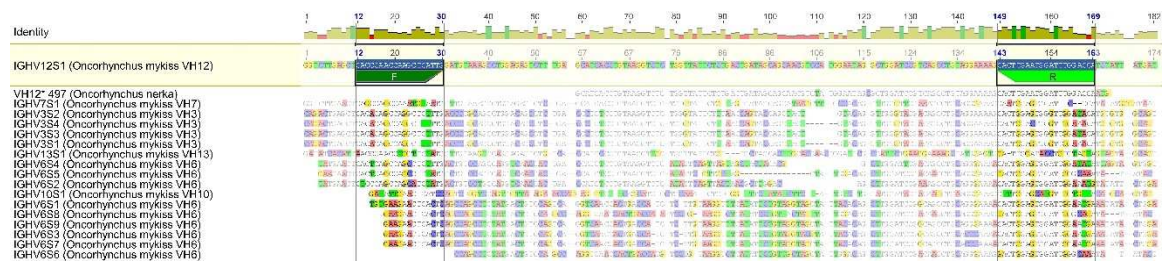
(D.)



(E.)



(F.)



**Figure 6:  $V_H$  family primer specificity.** Sequence alignments of specific *Oncorhynchus mykiss*  $V_H$  gene segments with other *Oncorhynchus mykiss*  $V_H$  gene segments. The reference sequence was the  $V_H$  gene sequence (or consensus of  $V_H$  gene sequences) corresponding to the  $V_H$  gene family in question. Forward primers are highlighted and labeled “F”, while reverse primers are highlighted and labeled “R”. Partial *Oncorhynchus nerka* sequences are included and labeled as such. For each partial sequence 5’ and 3’ ends were trimmed automatically using Geneious (except  $V_H$ 5.1, which was trimmed manually). O. nerka sample number is listed and corresponds to the fish of the same number listed in Supplemental Table I. Primer design focused on choosing sequences that were homologous for the specific  $V_H$  family in question, but that diverged for sequences of other  $V_H$  families. Of the 57  $V_H$  gene segments found in Rainbow trout only those that displayed a reasonable level of homology to the  $V_H$  family in question are shown. Those  $V_H$  gene segments not shown are more different than the gene segments used for comparison here. (A.)  $V_H$  1.1 sequence analysis. (B.)  $V_H$  2 Sequence analysis. (C.)  $V_H$  5.1 Sequence analysis. (D.)  $V_H$  9 Sequence analysis. (E.)  $V_H$  10 Sequence analysis. (F.)  $V_H$  12 Sequence analysis.

It was critical for accurate analysis that primer sets only amplify a single  $V_H$  family. Once potential primer combinations had been decided upon these primers were analyzed using the IDT Oligoanalyzer (*PrimerQuest® Program* (version 3.1) 2012). Each

primer set was analyzed for the Gibbs Free Energy necessary for formation of homodimers, as well as heterodimers between individual primers. Primer sets that produced homo- or heterodimers with a Delta G less than -7.00 kcal/mole were disregarded. Using the same tool, primer sets were analyzed for potential hairpin products and melt temperatures. Following this analysis primer sets were analyzed using the (*Multiple Primer Analyzer* 2016) for potential dimer formation. Significant dimer formation would disqualify potential primer sets, but single dimer formations using this tool did not immediately discount them. Potential primer sets were then entered into NCBI's Primer BLAST (Ye et al. 2012), searching all organisms in the nr (non-redundant) database. Results were compared to the desired organism specificity and amplicon size, any non-target amplification was noted. Primers designed are listed in Table I. Following this, primer sets were ordered as Value Oligos (Invitrogen) suspended in 100 $\mu$ L of molecular grade water. Upon receipt primer stocks were used to create aliquots of primers at a concentration of 3ng/ $\mu$ L before being stored at -80°C. Specificity of primers was tested by performing SYBR Green qPCR on a StepOne™ Real-Time PCR System 48 Well Instrument (Applied Biosystems) using the primers according to methods detailed in the next section. Upon completion the melt curve was analyzed for potential primer dimer products or off-target effects, and further analysis was conducted through 1.5-2% agarose gel electrophoresis (Agarose Fisher Scientific #BP160-100, TBE Buffer Thermo Scientific #B52) with the previously created PCR product. Electrophoresis was run at 120V for 10-20 minutes. Proper migration of bands of the expected size was observed.

### SYBR Green qPCR

Selected cDNA samples were removed from the -80°C freezer and allowed to thaw in ice. cDNA was diluted 1:4 using molecular grade water, resulting in a concentration of

12.5ng/μL. In an AirClean 600 PCR Work Station (AirClean® Systems, Inc.) a Master Mix was created consisting of 12.5μL of Power SYBR® Green PCR Master Mix (Applied Biosystems, #4367659), 9.5 μL of nuclease free water, and 1μL of each primer being used for each well for a total of 24μL per well. *Oncorhynchus*-specific primer sequences for membrane bound and secreted HCmu, as well as α-Tubulin can be found in Table II as described in Schouten et al (2013).. All primer sequences listed used a 60°C annealing temperature. Primer sequences for various V<sub>H</sub> families can be found in Table I.

**Table II – MemHCmu/SecHCmu/Tub Primers**

| Primer Name  | Primer Target              | Primer Sequence                  |
|--------------|----------------------------|----------------------------------|
| qtHCmu.F     | Secreted and Membrane HCmu | 5'-GCG CTG TAG ATC ACA TGG AA-3' |
| qtHCmu.sec.R | Secreted HCmu              | 5'-GCA AGT CAG GGT CAC CGT AT-3' |
| qtHCmu.mem.R | Membrane HCmu              | 5'-TTT CAC CTT GAT GGC AGT TG-3' |
| Ttub60.F     | Alpha-Tubulin              | 5'-CGT CCC CAG GTG TCC ACT-3'    |
| Ttub60.R     | Alpha-Tubulin              | 5'-GTA GGT GGG GCG CTC AAT-3'    |

After addition of all components the Master Mix was vortexed and centrifuged for 5 to 10 seconds using a benchtop microfuge to ensure homogenous distribution. 24μL aliquots of Master Mix were pipetted into a MicroAmp® Fast Optical 48-Well Reaction Plate (Thermo Fisher Scientific, Inc.). Next, 1μL of sample was added to the side of each specified well. A MicroAmp® 48-Well Optical Adhesive Film (Thermo Fisher Scientific, Inc.) was then used to seal the plate with a MicroAmp® Adhesive Film Applicator (Applied Biosystems). The plate was then spun down using a plate centrifuge and inserted into the 48 Well StepOne Instrument. Step One Software version 2.3 was used to create plate maps and identify wells and Master Mixes, as well as decide on the proper thermal conditions. For all SYBR Green qPCR assays samples were run at 95°C

for 10 minutes, followed by 40 cycles of 15 seconds at 95°C and 1 minute at 60°C. A melt curve analysis was performed for every SYBR Green experiment to ensure purity of PCR product. Each unknown sample was run in triplicate to ensure reproducibility of the results.

### TaqMan qPCR

Samples of spleen DNA were taken from the -80°C freezer and allowed to thaw on ice. Dilutions were calculated and created to standardize each sample at 50ng/μl. In the PCR work station the Master Mix was created according to the manufacturer provided protocol. For analysis of Fp, Rs, IN, and AS every well used required 10μl of TaqMan® Universal Master Mix II, no UNG (Applied Biosystems, #4440040) or TaqMan® Gene Expression Master Mix (Applied Biosystems, #4369016), 1μl of the desired assay, and a volume of RNase free water for a total of 20μL per well (including DNA volume). IHNV analysis required the use of the TaqMan® RNA-to-CT™ 1-Step Kit (Applied Biosystems #4392653), as sample RNA (rather than DNA) was used. 19 μl of master mix was added to specified wells of either a 48 well plate or 96 well plate depending on the size of the experiment. If the Stratagene Mx3005P Instrument (Agilent Technologies) was used a 96 well plate was used. If the StepOnePlus™ Real-Time PCR System 96 Well Instrument (Applied Biosystems) was used then the MicroAmp® EnduraPlate™ Optical 96-Well Fast GPLE Clear Reaction Plates with Barcode (Applied Biosystems #4483481) were used. Outside of the PCR work station if the concentration of the sample was 50ng/μL, 1μL of sample was added to the side of each specific well. If the concentration of the sample was 12.5ng/μL, 4μL of sample was added to the side of the specified well with corrected master mix volume. After all samples were added standards were retrieved from storage at -80°C.



Standards had been previously prepared in aliquots at varying concentrations (see earlier section). Standards available varied by a factor of 10 from 1 copy to 10,000,000 copies, however, after experimentation it was found that a range of 10 copies to 1,000,000 copies achieved accurate results while reducing risk of cross-contamination of standards. Standards were added to the sides of specified wells. Optical adhesive film was applied to the top of the plate using the film applicator. The plate was then centrifuged using the plate centrifuge. The centrifuged plate was then placed in the instrument and the run initiated. For detection of Rs and IHNv a Stratagene Mx3005P instrument (Agilent Technologies) was used. For Fp some plates were run on the same Stratagene instrument, while others were run on 48 or 96 well StepOne Instruments. AS detection was accomplished on either the 48 or 96 well StepOne Instrument. For IHNv quantification plates were run at 48°C for 15 minutes, 95°C for 10 minutes, and then 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. For quantification of other pathogens plates were run at 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. On both the Stratagene and StepOne instruments the associated software calculated the number of copies of the target present in each sample well based on the level of fluorescence produced from the known standard wells. All unknown samples were run in triplicate to ensure reproducibility of results. For certain samples (highlighted in red on Supplementary Table 1) no spleen tissue was available and anterior kidney tissue was used instead.

#### DNA Sequencing

To further confirm primer specificity, PCR products from V<sub>H</sub> amplifications of sockeye salmon were cleaned up using USB Affymetrix ExoSAP-IT PCR Product Clean-Up kit or GE Healthcare Illustra ExoProStar PCR cleanup kit according to the manufacturer's instructions. Cleaned-up sample was then used as a template in ABI BigDye v3.1

sequencing reactions. Reaction products were purified using Performa Gel-Filtration Cartridges (EdgeBio). Finally, samples were resuspended in ABI HiDi Formamide, and sequence products were resolved on an ABI3500 Genetic Analyzer and analyzed using ABI 5.4 Sequence Analysis Software. Methods for sequencing were performed according to instructions from the manufacturer. Partial sequences were obtained using both forward and reverse primers specific to the  $V_H$  family being analyzed.

#### Calculation of Fold Change

For evaluation of  $V_H$  families the fold difference from a selected sockeye salmon from the Mouth of the Kenai was used as a relative form of comparison. This reference fish was selected as having measured values closest to the average of all MoK fish, and the MoK site was used as it should theoretically represent the broadest array of antibody fingerprints. Each SYBR Green qPCR run produced results in the form of a  $C_T$  value. These  $C_T$  values represented the cycle at which the fluorescence from each well passed a predetermined threshold ("Real-Time PCR: Understanding Ct Application Note" 2016). The  $C_T$  values were then used to calculate the fold change relative to the reference fish, taking into account the  $C_T$  value of different endogenous controls. When calculating the fold change of SecHCmu or MemHCmu,  $\alpha$ -tubulin was used as the endogenous control. When calculating the fold change of various  $V_H$  families the endogenous control used was the whole of the secreted heavy chain mu transcripts, as the  $V_H$  families should be a subset of this wider range of antibodies (Table I). SecHCmu was used as the control rather than MemHCmu because during the spawning stage the majority of all secreted antibody comes from LLPCs (which are cortisol resistant), while transcripts from mature B cells (expressing membrane IgM) are less abundant. Fold change was calculated according to the formulas in Figure 7 (Livak and Schmittgen 2001). The fold change calculated this way was then used for subsequent statistical analyses.

$$\Delta C_T = C_{T, Target} - C_{T, SecIgM}$$

$$\Delta(\Delta C_T) = \Delta C_{T, Unknown} - \Delta C_{T, Control}$$

$$Fold\ Change = 2^{-\Delta(\Delta C_T)}$$

*Figure 7: Formulas used to calculate fold change from collected  $C_T$  values according to Livak and Schmittgen (2001). When calculating  $V_H$  fold change the  $C_T$  value of the sample for SecHCmu was subtracted from the  $C_T$  value for the  $V_H$  family in question. This value is known as the  $\Delta C_T$ . Next, the  $\Delta C_T$  for that specific  $V_H$  family with regard to the reference sample was subtracted from the  $\Delta C_T$  value of the unknown sample. This value is known as the  $\Delta(\Delta C_T)$ . Finally, the fold change is calculated for the unknown sample by raising 2 to the  $-\Delta(\Delta C_T)$ .*

### Statistical Analyses

Data from the analysis of  $V_H$  families were taken in the form of fold change adjusted for secreted heavy chain mu transcripts, relative to a control fish. Data from the pathogen experiments were taken in the form of copy number of the target amplicon through the use of a standard curve. MemHCmu and SecHCmu expression was measured as fold change adjusted for  $\alpha$ -tubulin transcripts, relative to a control fish. To compare differences in fold change of individual  $V_H$  families an ANOVA was performed followed by a Tukey-Kramer analysis using R (Lau 2013). To evaluate the selected antibody repertoire using available  $V_H$  family data a MANOVA was performed, followed by 15 selected Hotelling's T-Squared Tests (Curran 2013). To account for multiple comparisons a Bonferroni correction was applied in the case of the Hotelling's T-Squared Tests. In all cases the statistical programming environment R was used (R Core Team 2015). Locations were analyzed for independence from infection rate using a permutation test found in the "coin" package in R (Hothorn et al. 2015).

To ensure even sample numbers five samples were randomly removed from the EFG site samples using the random number generator in R (“How to Generate a Random Number in R | inside-R | A Community Site for R” 2016). To better visualize the data it was input into PAST, a program that allows for a variety of multivariate statistical analyses (Hammer, Harper, and Ryan 2001). A non-metric dimensional scaling (NMDS) was performed to produce a 2-dimensional representation of the data and allow for the visualization of pattern groupings that takes into account all of - or a subset of - the variables.  $V_H$  family data were shown via a strip chart created in the R environment using the program “ggplot” (“ggplot2 Stripchart (Jitter) : Quick Start Guide - R Software and Data Visualization - Documentation - STHDA” 2016), (“Plotting Means and Error Bars (ggplot2)” 2016), (“Summarizing Data” 2016), (Wickham, Chang, and RStudio 2016). Additionally, the correlation between variables was calculated using R and visualized using the program “corrplot” (“R: A Visualization of a Correlation Matrix.” 2016; “An Introduction to Corrplot Package” 2016; Wei and Simko 2016),.

## Results

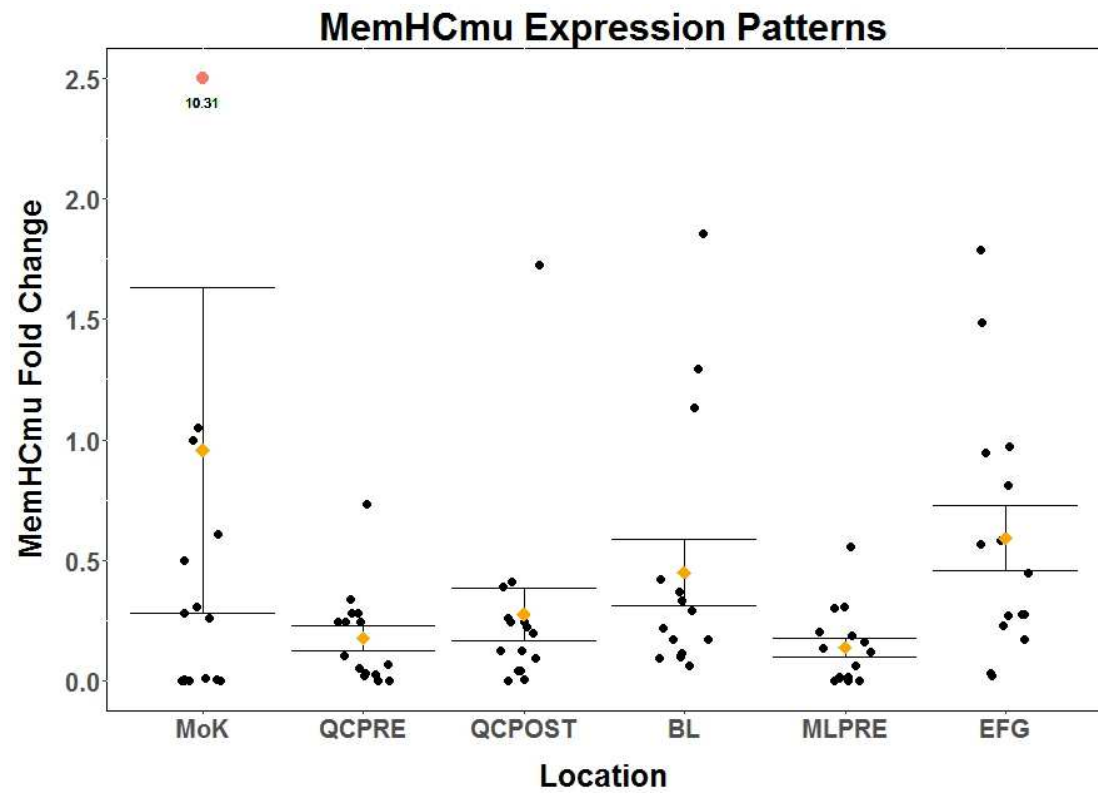
To assess the potential differences between the antibody usage patterns of Sockeye salmon, five different sites were studied, from three different salmon runs (Figure 3). Fifteen fish from each site were analyzed. The Kenai Peninsula run included two sites, Mouth of the Kenai (MoK), where the fish entered the Kenai River, and Quartz Creek (QC), where fish spawned. Quartz Creek was visited twice, with the goal to collect both prespawners (QCPRE, early) and post-spawners (QCPOST, late). The Copper River included two sites, EFG (very early pre-spawners, a few miles from their spawning ground), and Mentasta Lake, a separate spawning site. The Resurrection Peninsula run included one spawning site, Bear Lake.

Three specific components of the humoral immune response were studied. First the relative levels of MemHCmu and SecHCmu were measured. Next, specific V<sub>H</sub> family usage patterns were determined. Finally, the presence and pathogen load of 5 different pathogens were determined for each fish. The relationships between these values were then compared and contrasted between sites. The Quartz Creek site was useful to compare immune patterns between pre-and post-spawners within the same site. For purposes of analysis, pre and post-QC are considered two different “sites”.

### Membrane Bound and Secreted HCmu

Measurements of MemHCmu and SecHCmu immunoglobulin heavy chain mu transcripts began in 2009 and were continued here. Using qPCR, the expression of MemHCmu and SecHCmu was calculated relative to a control *Oncorhynchus nerka* sample from the Mouth of the Kenai River (#349, see Supplemental Table I). Expression data were then plotted as fold-change on a strip chart according to sampling site (Figure 8).

(A.)



(B.)

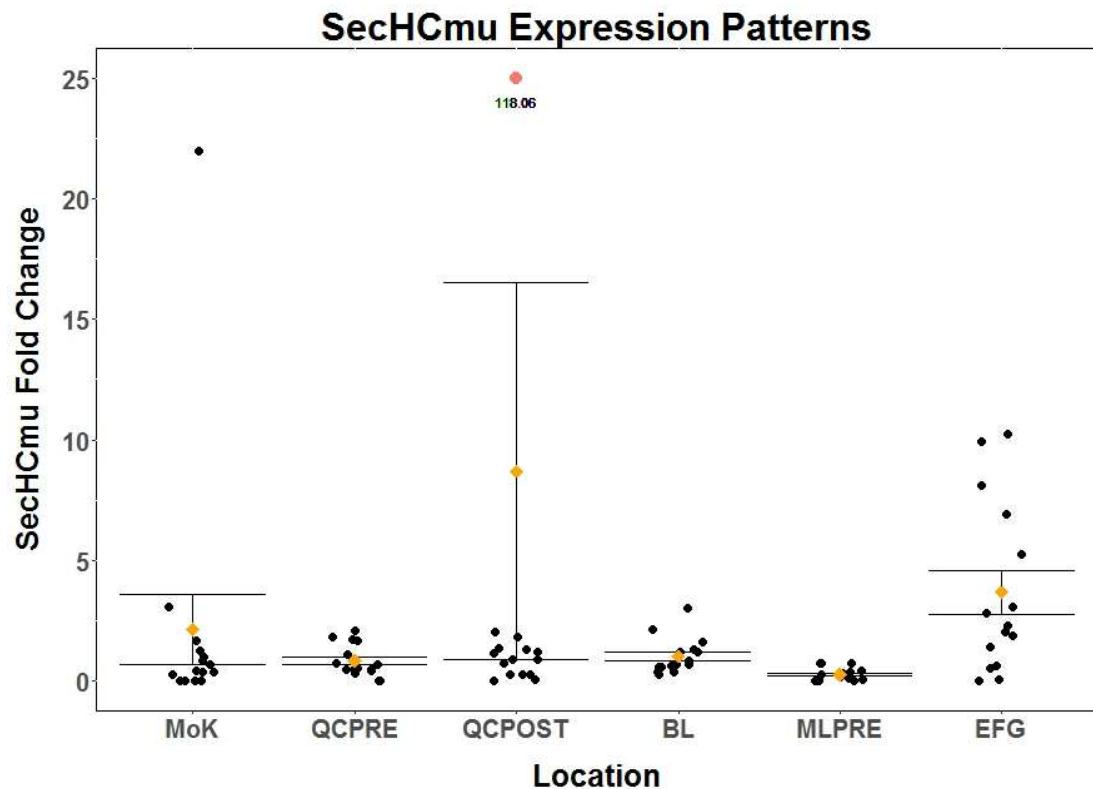


Figure 8: Strip charts showing relative expression of MemHCmu (A.) and SecHCmu (B.) relative to the reference fish from the Mouth of the Kenai (fish #349). The Y-axis shows fold change. The X-axis shows the locations sampled. “MoK”, Mouth of the Kenai, “QCPRE”, Quartz Creek pre-spawned, “QCPOST”, Quartz Creek post-spawned, “BL”, Bear Lake, “ML”, Mentasta Lake, and “EFG”, East Fork Gulkana. Standard error bars are shown and the average value is indicated in orange. The Y-axis has been adjusted to represent the majority of data, while outliers are shown on the top with fold change specified.

As the reference sample from which all relative fold changes were calculated was a MoK fish it is reasonable to think that the majority of MoK fish should demonstrate fold changes close to 1.0. This was not the case; likely due to the level of variation visible at

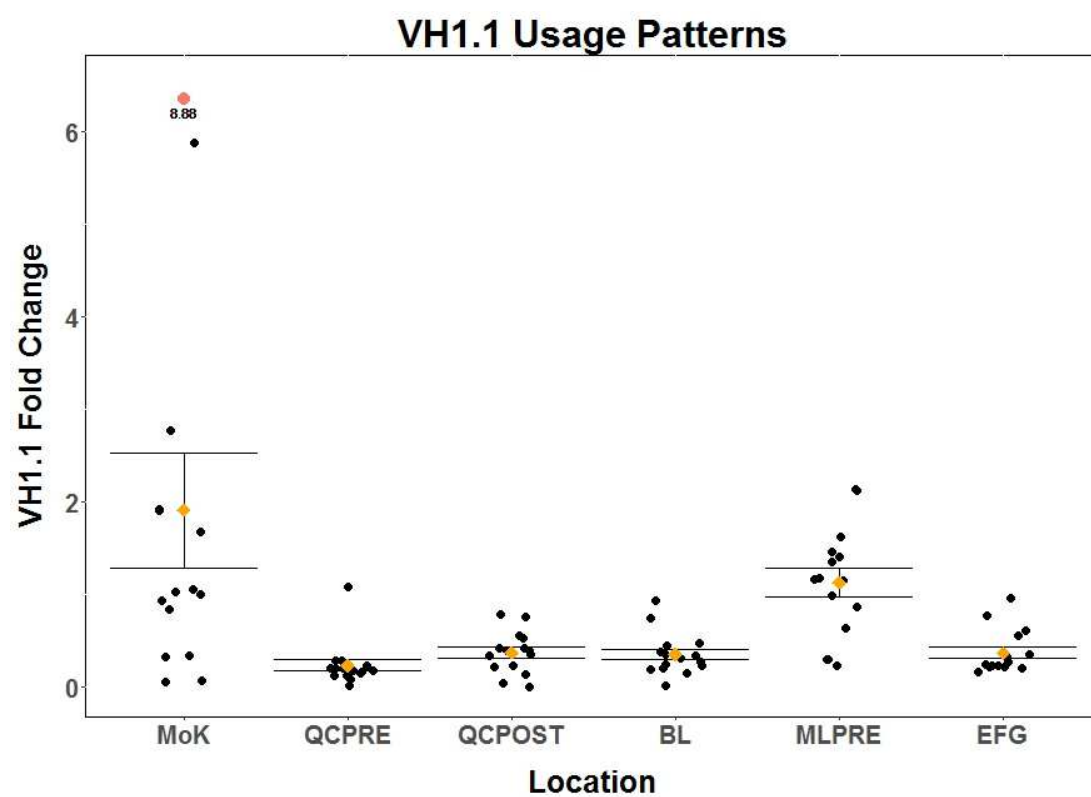
the MoK site. While the reference sample chosen (#349) had measured values closest to average for the entire site, there were still differences between the reference fish and many of the other MoK fish. Interestingly, for the Kenai run, the average expression level for MemHCmu was 3-5 times higher at the beginning of the run (MoK), compared to their expression at the spawning site (QC). Further, for the Copper River run, EFG had almost 5 times higher SecHCmu expression compared to most other sites measured, although these differences were not significant. Lastly, SecHCmu expression was very evenly distributed through the locations analyzed. Hence, secreted IgM levels seem to be maintained at very similar levels in spawning fish from different runs.

#### Analysis of V<sub>H</sub> Families

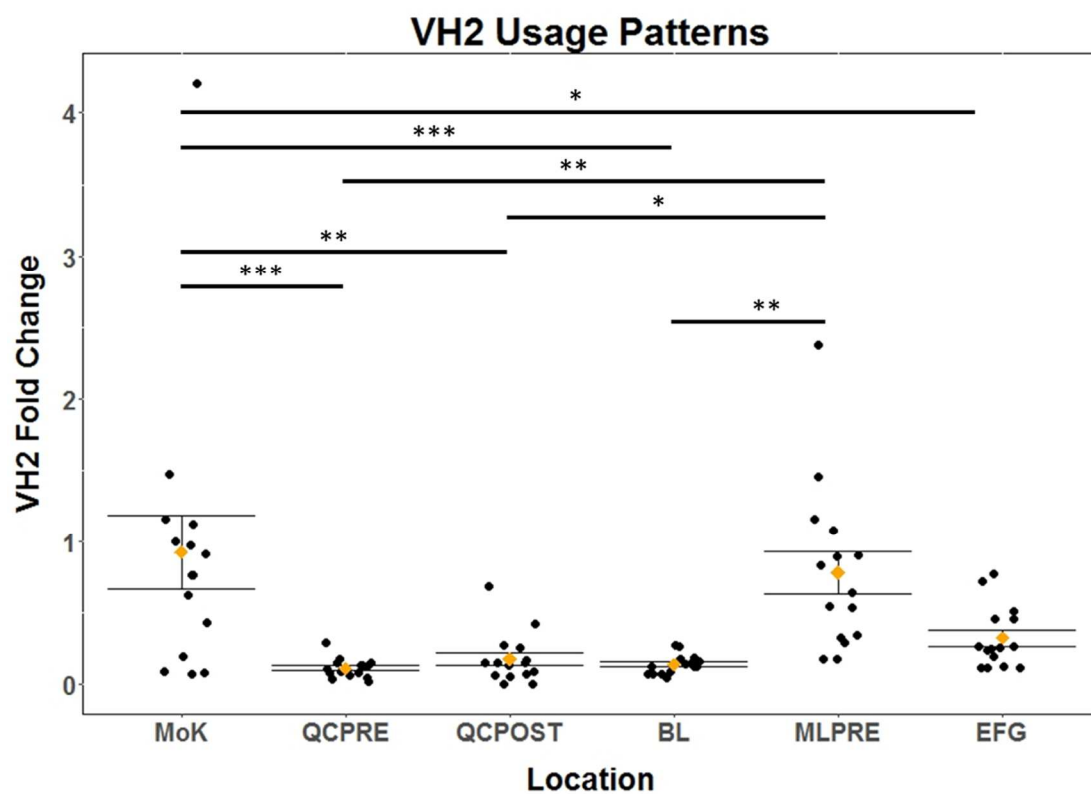
Relative gene expression of six different V<sub>H</sub> families (V<sub>H</sub>1.1, V<sub>H</sub>2, V<sub>H</sub>5.1, V<sub>H</sub>9, V<sub>H</sub>10, and V<sub>H</sub>12) was measured for each fish using qPCR. The anterior kidney was used as the tissue for V<sub>H</sub> usage determination, as this tissue is the primary site of LLPC residence, the cell population of primary interest to us. Expression of each individual V<sub>H</sub> family was then plotted on a strip chart according to sampling site. Figure 9 shows the comparisons of individual V<sub>H</sub> families across the six sites.



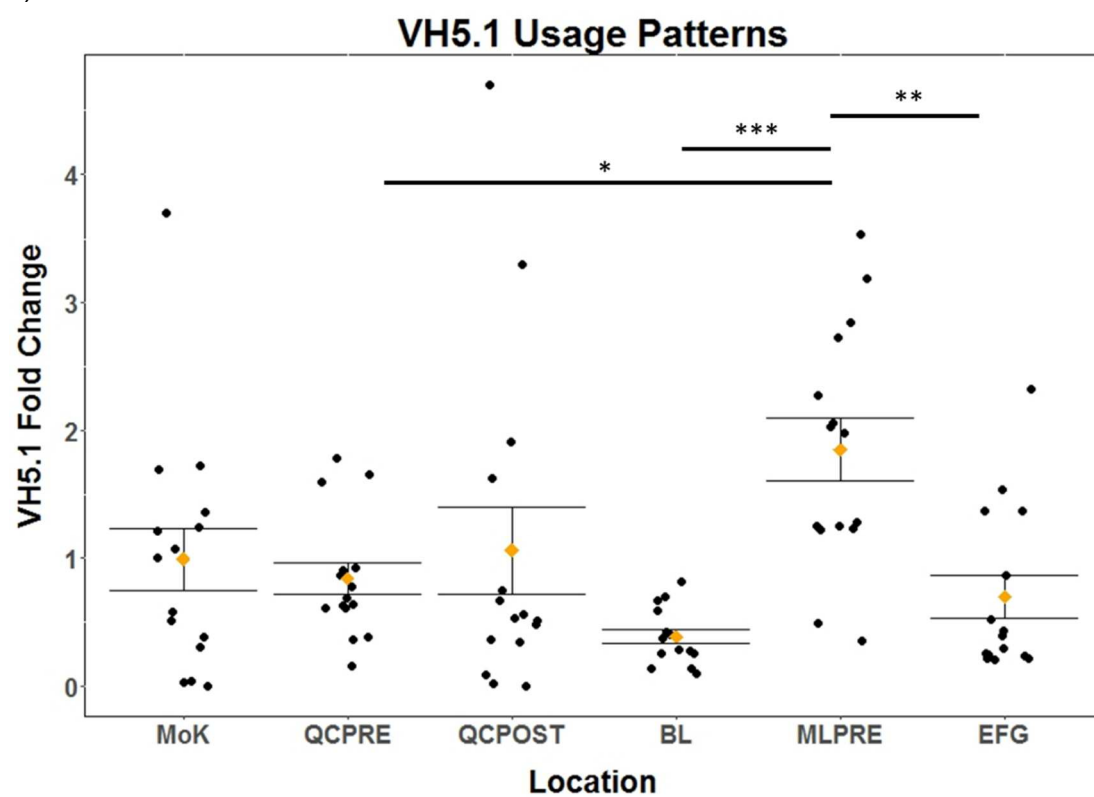
(A.)



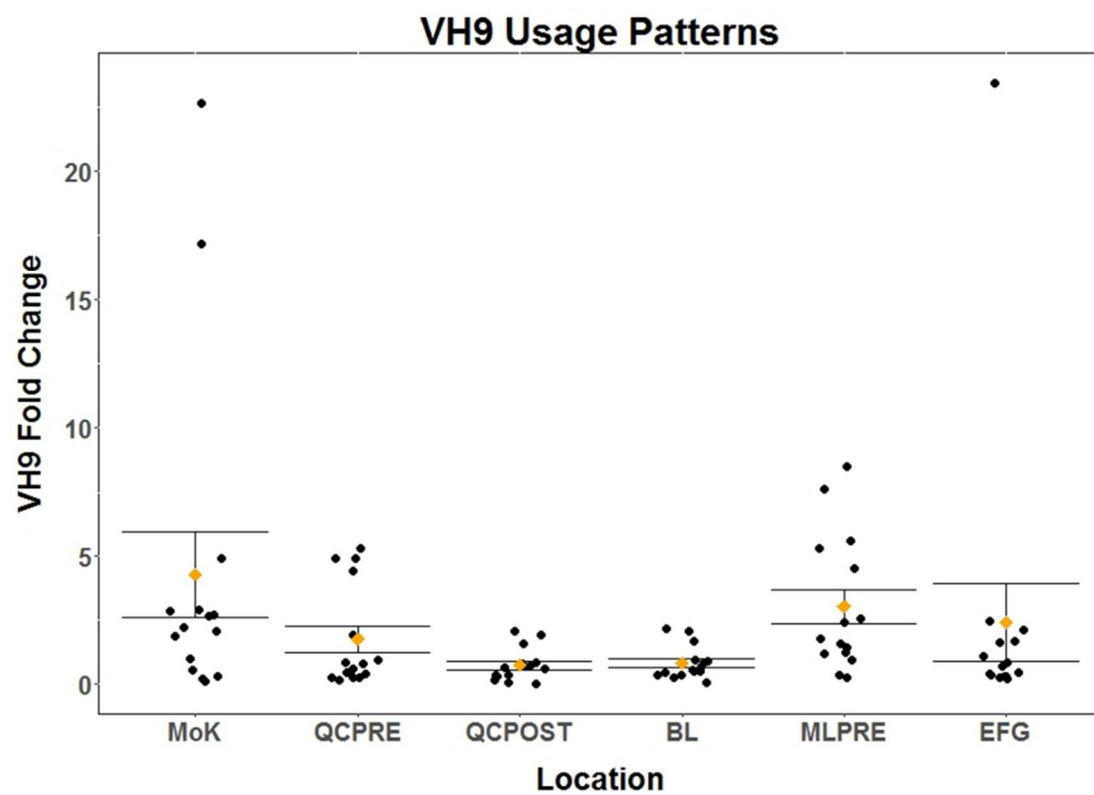
(B.)



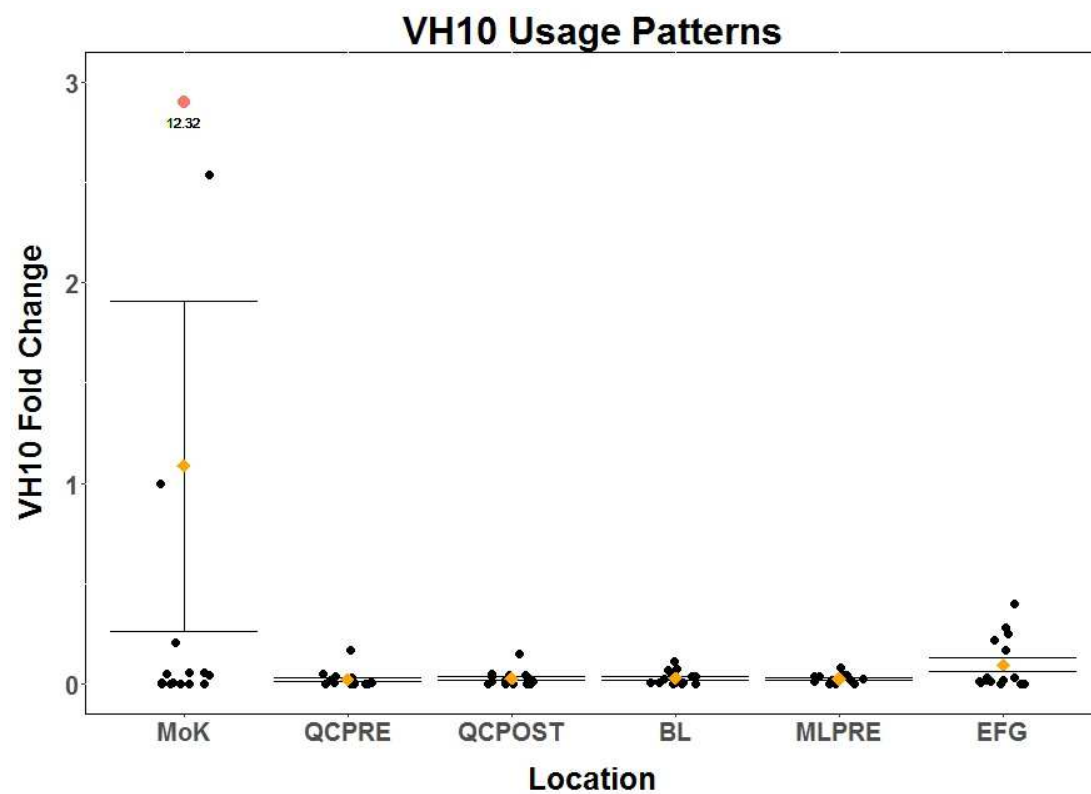
(C.)



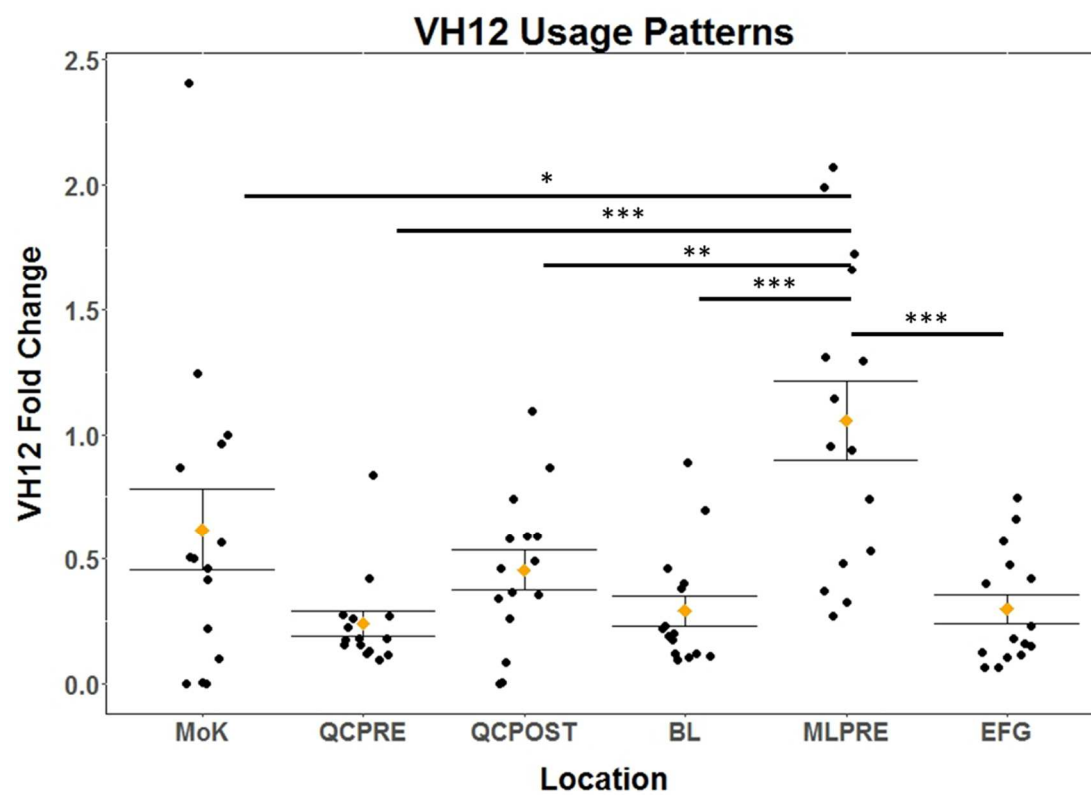
(D.)



(E.)



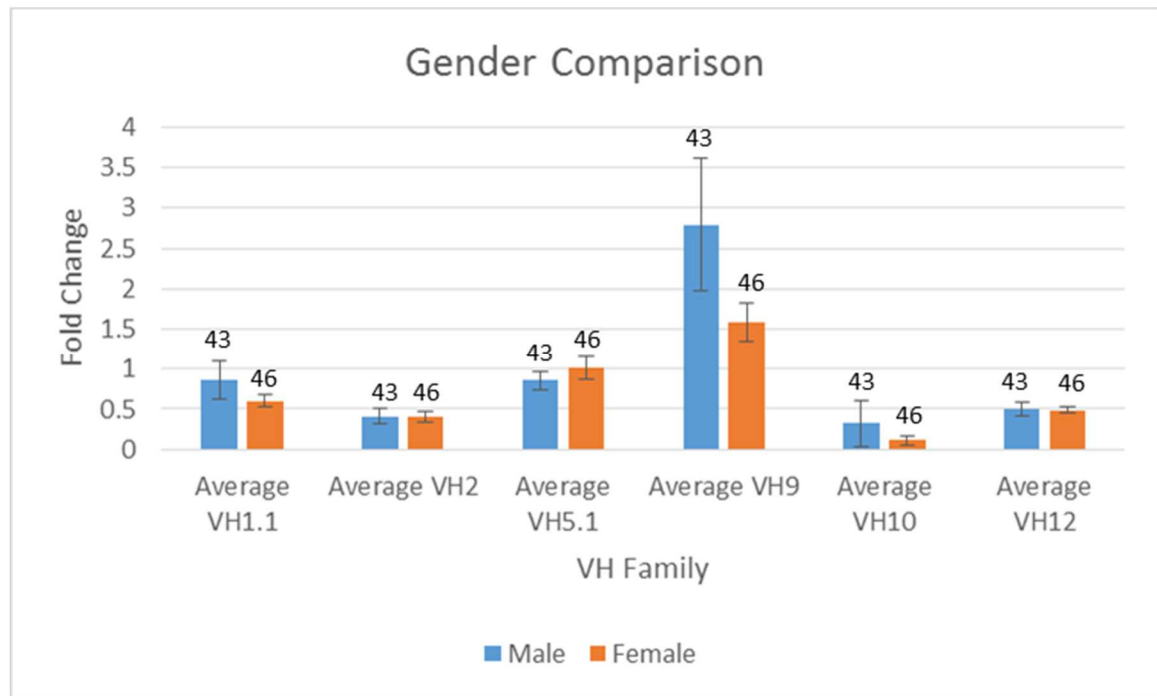
(F.)



*Figure 9: Strip charts showing the distribution of  $V_H$  fold change data according to location. The fold change relative to the Mouth of Kenai (MoK) control fish (#349) is shown on the Y-axis, while the geographic location sampled is shown on the X-axis. “MoK” is Mouth of the Kenai, “QCPRE” is Quartz Creek pre-spawned, “QCPOST” is Quartz Creek post-spawned, “BL” is Bear Lake, “ML” is Mentasta Lake, and “EFG” is East Fork Gulkana. The standard error bars are shown for each location and the mean is shown in orange for each location. Significant relationships are labeled with bars connecting the two sites found to be different and asterisks to denote level of significance. One asterisk indicates  $p < 0.05$ , two asterisks indicate  $p < 0.01$ , three asterisks indicate  $p < 0.001$ . (A.)  $V_{H1.1}$  usage. (B.)  $V_{H2}$  usage. (C.)  $V_{H5.1}$  usage. (D.)  $V_{H9}$  usage. (E.)  $V_{H10}$  usage. (F.)  $V_{H12}$  usage. The Y-axis has been adjusted to better view trends in the data, and outliers are shown on top with fold change specified.*

Varying levels of significance can be seen for each  $V_H$  family and between each site. For example, there are significant differences between sites for  $V_{H1.1}$ ,  $V_{H2}$ ,  $V_{H5.1}$ , and  $V_{H12}$ , while there are no significant differences between sites for  $V_{H9}$  and  $V_{H10}$ . Interestingly, a fair portion of the significance observed appears to be associated with differences between sites from different runs.  $V_{H2}$  and  $V_{H1.1}$  appeared to show the most significant differences between spawning sites. The significant comparisons for  $V_{H2}$  appear to be somewhat widely distributed in terms of the comparisons between sites. In contrast the significance associated with  $V_{H5.1}$  and  $V_{H12}$  is centered on the ML site. In fact, every significant comparison seen in the  $V_{H12}$  analysis involves ML. Similarly, for  $V_{H1.1}$  every significant comparison is driven by the inclusion of MoK. It should be noted that although not significant, average  $V_{H1.1}$  expression for ML fish was at least 2x higher

compared to other sites. There were no significant differences observed across sampling sites for V<sub>H</sub>9 and V<sub>H</sub>10.



*Figure 10: Bar graph showing the average fold change relative to a control fish for each V<sub>H</sub> family, divided by gender. Standard error bars are shown, and the N value is shown above each bar. Fold change was calculated relative to a MoK reference fish. The Y-axis shows fold change, the X-axis shows V<sub>H</sub> family analyzed. Blue indicates male, orange indicates female. Total sample size was 89 (one fish analyzed did not have gender recorded).*

Next, potential differences due to gender were evaluated (Figure 10). When comparing all V<sub>H</sub> families according to gender no significant relationships were observed. When sites were analyzed individually for potential differences between genders, several significant relationships were observed (Table III).

**Table III – Summary of Comparisons Between Genders per Site**

| VH Family | MoK          |                |         | QCPRE        |                |         | QCPOST       |                |         | BL           |                |         | MLPRE        |                |         | EFG          |                |         |
|-----------|--------------|----------------|---------|--------------|----------------|---------|--------------|----------------|---------|--------------|----------------|---------|--------------|----------------|---------|--------------|----------------|---------|
|           | Male Average | Female Average | P-Value | Male Average | Female Average | P-Value | Male Average | Female Average | P-Value | Male Average | Female Average | P-Value | Male Average | Female Average | P-Value | Male Average | Female Average | P-Value |
| VH1.1     | 2.62         | 1.13           | 0.24    | 0.16         | 0.30           | 0.29    | 1.51         | 0.39           | 0.78    | 0.29         | 0.39           | 0.38    | 1.02         | 1.22           | 0.56    | 0.44         | 0.29           | 0.25    |
| VH2       | 1.04         | 0.85           | 0.71    | 0.08         | 0.14           | 0.06    | 0.13         | 0.22           | 0.36    | 0.09         | 0.17           | 0.02    | 0.57         | 0.96           | 0.20    | 0.41         | 0.22           | 0.07    |
| VH5.1     | 0.76         | 0.85           | 0.79    | 0.64         | 1.01           | 0.13    | 1.03         | 1.09           | 0.92    | 0.37         | 0.40           | 0.78    | 1.38         | 2.26           | 0.07    | 0.92         | 0.45           | 0.17    |
| VH9       | 6.25         | 1.92           | 0.20    | 1.68         | 1.84           | 0.89    | 0.62         | 0.86           | 0.49    | 0.64         | 0.96           | 0.37    | 2.79         | 3.22           | 0.77    | 3.84         | 0.80           | 0.32    |
| VH10      | 1.57         | 0.59           | 0.56    | 0.04         | 0.01           | 0.25    | 0.02         | 0.03           | 0.47    | 0.03         | 0.03           | 0.96    | 0.02         | 0.03           | 0.47    | 0.06         | 0.14           | 0.33    |
| VH12      | 0.58         | 0.70           | 0.72    | 0.20         | 0.28           | 0.37    | 0.42         | 0.49           | 0.67    | 0.31         | 0.29           | 0.87    | 1.07         | 1.04           | 0.94    | 0.41         | 0.18           | 0.05    |

*Table III: Locations are listed along the X-axis, V<sub>H</sub> families analyzed are listed along the Y-axis. Results of Welch's t-tests between male and female fish with regards to individual V<sub>H</sub> families at individual locations analyzed are listed under P-value. Average fold changes for each V<sub>H</sub> family are also listed. Using an  $\alpha=.05$  threshold those values that were found to be significant are highlighted in green.*

The only significant gender differences occurred at the BL site for the V<sub>H</sub>2 family, and at the EFG site for V<sub>H</sub>12.

Although analysis of individual V<sub>H</sub> family usage is potentially useful, it does not necessarily address the original research question, which is whether the overall V<sub>H</sub> usage *patterns* differ between sites. This overall usage pattern is referred to here as the “antibody fingerprint”. To better assess the antibody fingerprint of each site, the averages of each site were converted into pie charts to show the relative distribution of each V<sub>H</sub> family per site (Figure 11). These pie charts are useful for evaluating the relative

V<sub>H</sub> usage patterns of individual sites, but they cannot be compared directly as they use only relative information.

To quantify potential differences between entire antibody fingerprints at once a MANOVA was conducted. The MANOVA essentially compares each pie chart to each other while taking into account the actual fold changes between sites. The MANOVA gave a p-value of  $8.909 \times 10^{-7}$ , indicating a high level of significance. However, this only means that between all of the sites analyzed there is at least one comparison of two sites that produces a significant difference; it cannot specify where that significant difference might be. To approach the question of which specific comparisons may be significant a series of Hotelling  $T^2$  tests were performed. The Hotelling  $T^2$  test is analogous to a two sample T-test with univariate data, but using multivariate data. In other words, the Hotelling  $T^2$  test compares individual pie charts to each other to see if, when all of the V<sub>H</sub> families are taken into account there may be significant differences that could not be observed through analysis of individual V<sub>H</sub> families. However, to avoid the problem of multiple comparisons associated with using multiple tests on the same data we performed a Bonferroni correction, meaning that rather than using a threshold of  $\alpha=.05$  we instead divided this generally accepted value by the number of comparisons we were making. Fifteen comparisons were made, as shown in Table IV. This means that our threshold value would be 0.0033, thus a comparison would only be considered significant if the Hotelling  $T^2$  test returned a p-value less than 0.0033. P-values for the fifteen selected comparisons are listed in Table IV.

Figure 11

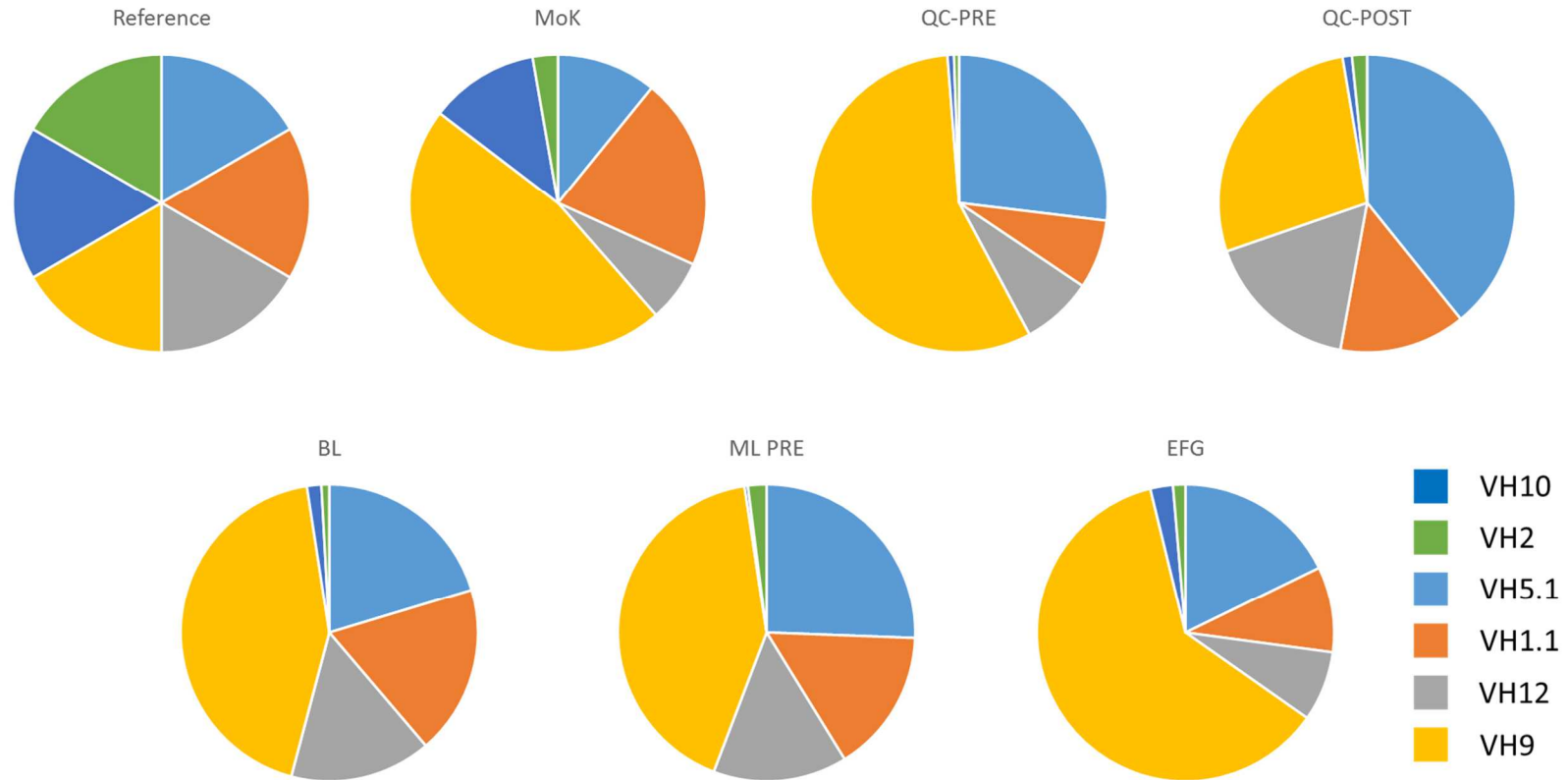


Figure 11: Pie charts showing relative usage patterns of each  $V_H$  family at each geographic location. Sites are listed according to their title, with the reference first followed by MoK, QC-PRE, QC-POST, BL, ML, and EFG. Light blue represents  $V_{H5.1}$ , orange represents  $V_{H1.1}$ , grey represents  $V_{H12}$ , represents  $V_{H9}$ , dark blue represents  $V_{H10}$ , and green  $V_{H2}$ . The "Reference" pie chart is included and demonstrates what the reference fish's (fish #349) antibody fingerprint would look like as all antibody expression levels would be 1.0. Each pie chart can also be seen as the measured deviation from this reference sample.



| Comparison         | P-Value |   |
|--------------------|---------|---|
| MoK vs QC-Pre      | 0.01    | <i>Table IV: Results of Hotelling <math>T^2</math> tests analyzing the potential differences between measured <math>V_H</math> families across different sites. Comparisons made are listed on the left, p-values for each comparison are listed on the right. A Bonferroni correction was applied, resulting in <math>\alpha=0.00333</math>. Those comparisons found to be significant after correction are shaded in green.</i> |
| MoK vs QC-Post     | 0.03    |   |
| MoK vs ML          | 0.02    |   |
| MoK vs BL          | 0.01    |   |
| MoK vs EFG         | 0.13    |   |
| QC-Pre vs QC- Post | 0.18    |   |
| QC-Pre vs ML       | 0.0002  |   |
| QC-Pre vs BL       | 0.001   |   |
| QC-Pre vs EFG      | 0.00002 |   |
| QC-Post vs ML      | 0.001   |   |
| QC-Post vs BL      | 0.48    |   |
| QC-Post vs EFG     | 0.0026  |   |
| ML vs BL           | 0.0001  |   |
| ML vs EFG          | 0.002   |   |
| BL vs EFG          | 0.01    |   |

Importantly, even with the low  $\alpha$  necessary for significance approximately half of the comparisons indicated significant differences (or unique “antibody fingerprints”) between sites. Interestingly, the majority of those comparisons that were not significant included MoK, while all significant differences were comparisons between spawning locations. In addition, the antibody fingerprint at the BL spawning site was surprisingly similar to that of the QC-Post site. Further, and in support of the Immunological Imprinting Hypothesis, the antibody fingerprints of pre- and post-spawned fish from the same site (QC) were not significantly different. In summary, when the entire antibody fingerprint was taken into account there were significant differences in  $V_H$  usage patterns between fish from different spawning sites.

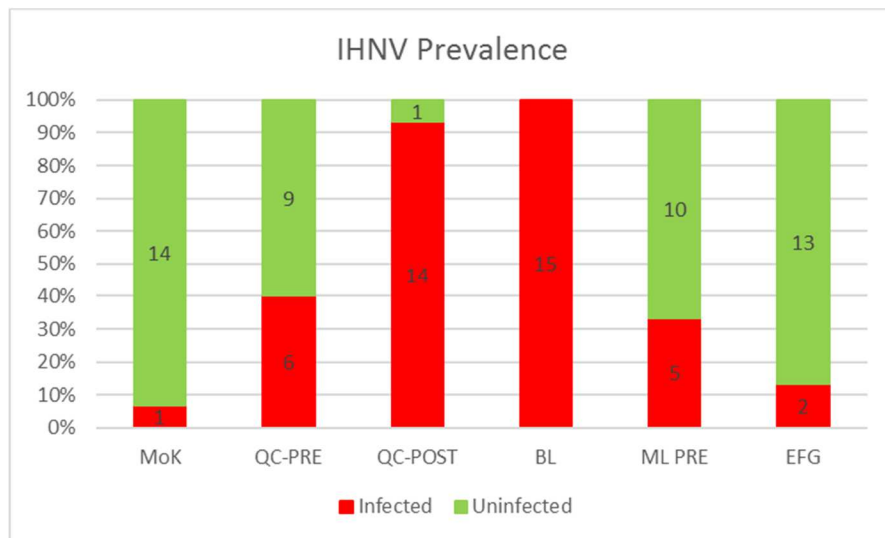
### Pathogen Analysis

The “antibody fingerprint”, as measured by  $V_H$  family usage, is reasonably assumed to be affected by the unique set of pathogens at different sites, known as the

“pathogen fingerprint”. To investigate the “pathogen fingerprint”, as well as how  $V_H$  usage may be impacted by pathogen infection, fish were analyzed for the presence and copy number of four relevant fish pathogens. Infectious Hematopoietic Necrosis Virus (IHNV), *Flavobacterium psychrophilum* (Fp), *Renibacterium salmoninarum* (Rs), and *Aeromonas salmonicida* (AS) were analyzed using TaqMan qPCR for their presence and load. Spleen was chosen for these four pathogens because it is a common immune organ used for pathogen detection (Strepparava et al. 2014; Soto et al. 2010; Marancik and Wiens 2013; Bowers, Lapatra, and Dhar 2008). Visual presence or absence of generic worms was recorded during the time of collection.

The analysis for presence and levels of IHNV RNA (targeting the N gene) is shown in Figure 12.

(A.)



(B.)

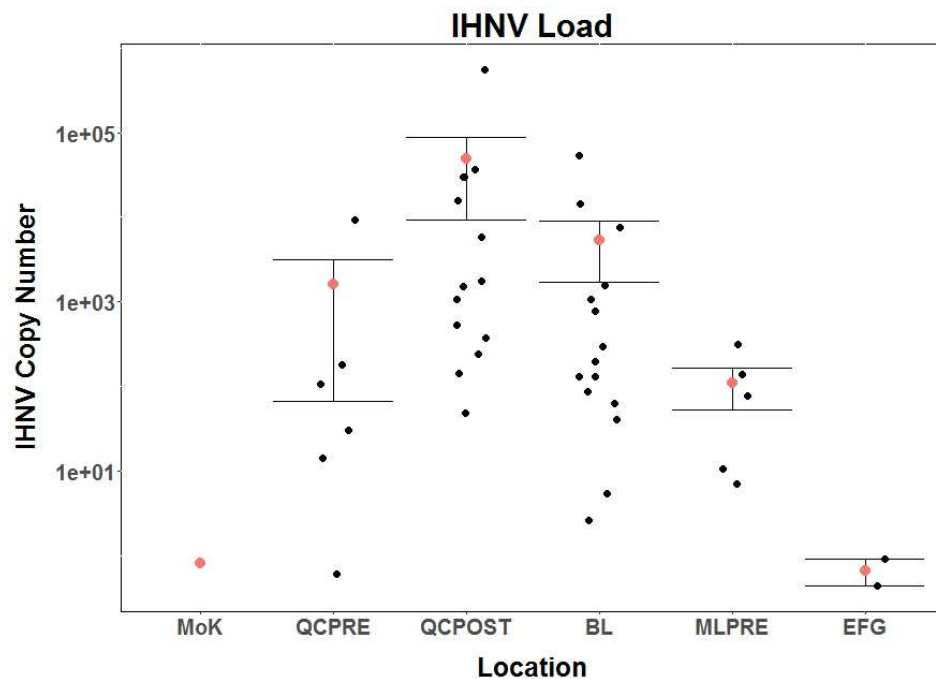
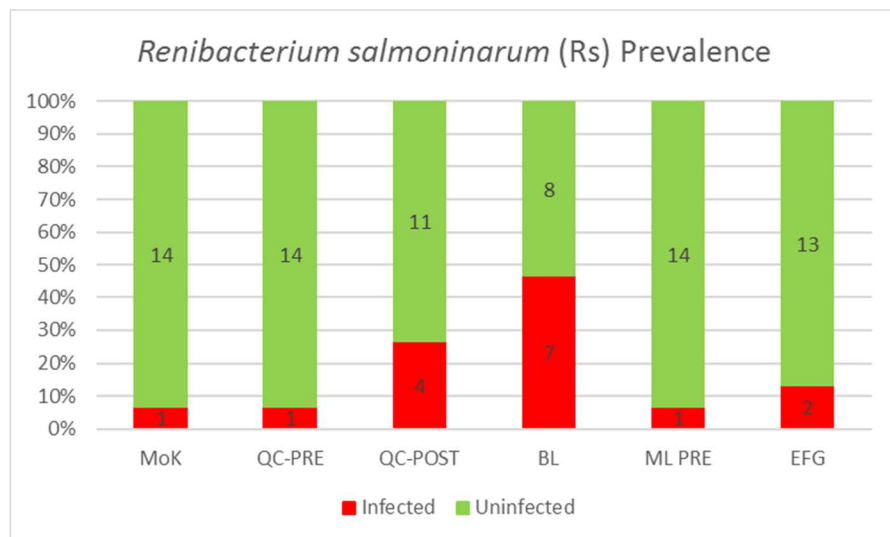


Figure 12: (A.) IHN Infection. Bar graph showing the prevalence of infection in analyzed fish. The collection site analyzed is listed along the X-axis and the relative percentage is listed along the Y-axis. 15 fish from each site were analyzed, the percentage infected is shown in red with the number infected listed within the bar. (B.) Strip chart showing the measured copy number of IHN in those fish that were infected. Means are shown in orange. Standard error bars are used showing the standard error of the mean. Number of copies of the given pathogen is listed on the Y-axis in log scale while the location analyzed is listed on the X-axis.

Of all of the pathogens analyzed the highest number of copies was found for IHN. Such a result may be expected given the nature of viral replication and the production of a large number of genomic copies, both those present in viable virions and those that are not. This highlights a potential weakness in the quantification of pathogens via qPCR given that the presence of genomic material does not imply infectivity or viability.. TaqMan analysis in search of IHN found N gene expression at

each site analyzed, though with drastically different frequencies (Figure 12A). For example, at MoK only a single fish was found to show any evidence of IHNV, but at BL every single fish analyzed was infected with the virus. Apart from the frequency of infection between fish varying, there was a fair amount of variation in the number of copies found in each fish. While BL clearly had the highest percentage of fish infected with IHNV, fish from the QCPOST site had almost 10-fold higher average copy number per infected fish compared to BL. One “outlier” fish from QCPOST demonstrated the highest measured level of IHNV presence from all 90 fish studied, with 557,000 copies found, a whole order of magnitude higher than in any other IHNV-infected fish. Fish entering the river at MoK had a lower average percentage of infection and a lower average IHNV load compared to QC-PRE, while QC-POST was highest for both.

(A.)



(B.)

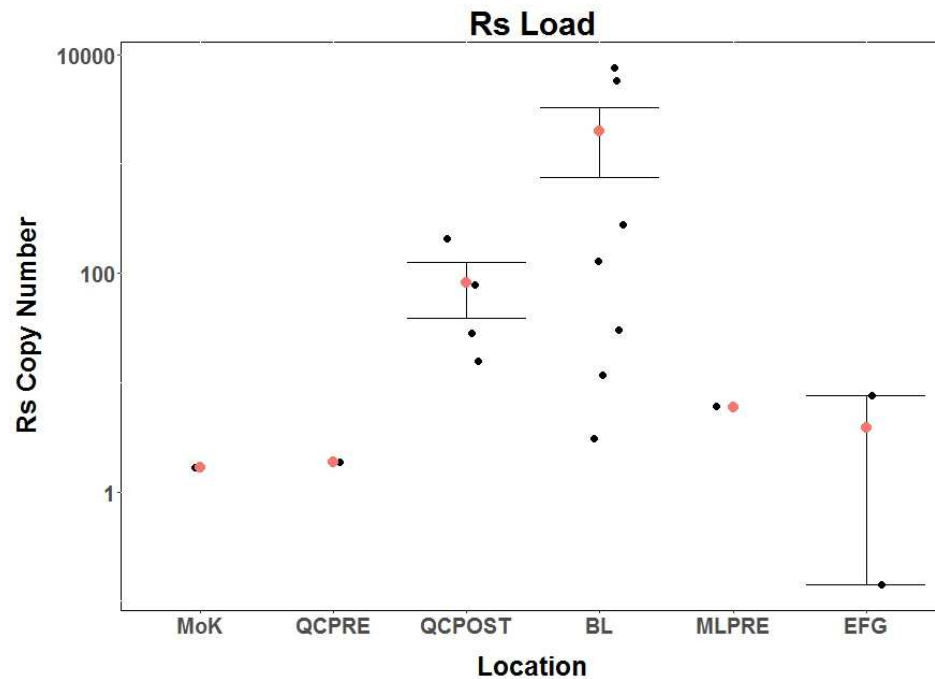
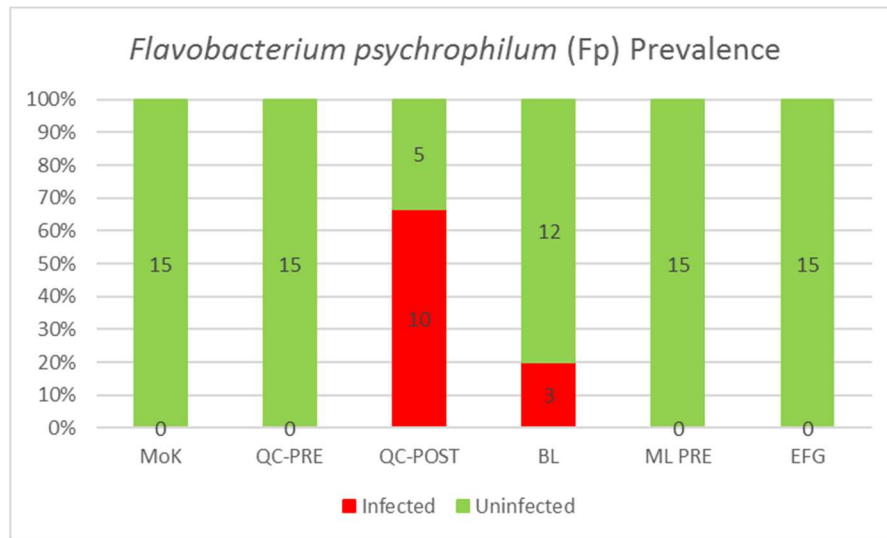


Figure 13: (A.) *Rs* Infection. Bar graph showing the prevalence of infection in analyzed fish. The collection site is listed along the X-axis and the relative percentage is listed along the Y-axis. 15 fish from each site were analyzed. The percentage infected is shown in red with the number infected listed within the bar. The percentage uninfected is shown in green with the number uninfected listed within the bar. (B.) Strip chart showing the measured copy number of *Rs* in infected individuals. Means are shown in orange. Standard error bars are shown, indicating standard error of the mean. Number of copies of *Rs* is listed on the Y-axis in log scale while the locations are listed on the X-axis.

Analysis of *Rs* was performed next (Figure 13). Patterns for *Rs* were similar to those for IHNV, especially for BL and QC-post, which had the highest rates and loads of both pathogens. Notably, the prevalence of *Rs* at all of the sites was low compared to IHNV, with several sites (MoK, QCPRE, and ML) having only a single infected individual. Additionally, actual copy numbers of *Rs* were much lower than those observed for IHNV,

with the highest level detected being approximately 8000 copies. This difference in copy number likely can be explained by the difference in pathological agent. There were a large number of fish that had a relatively low level of Rs, but as the assay is specific, even a low level of detection should be interpreted as being infected.

(A.)



(B.)

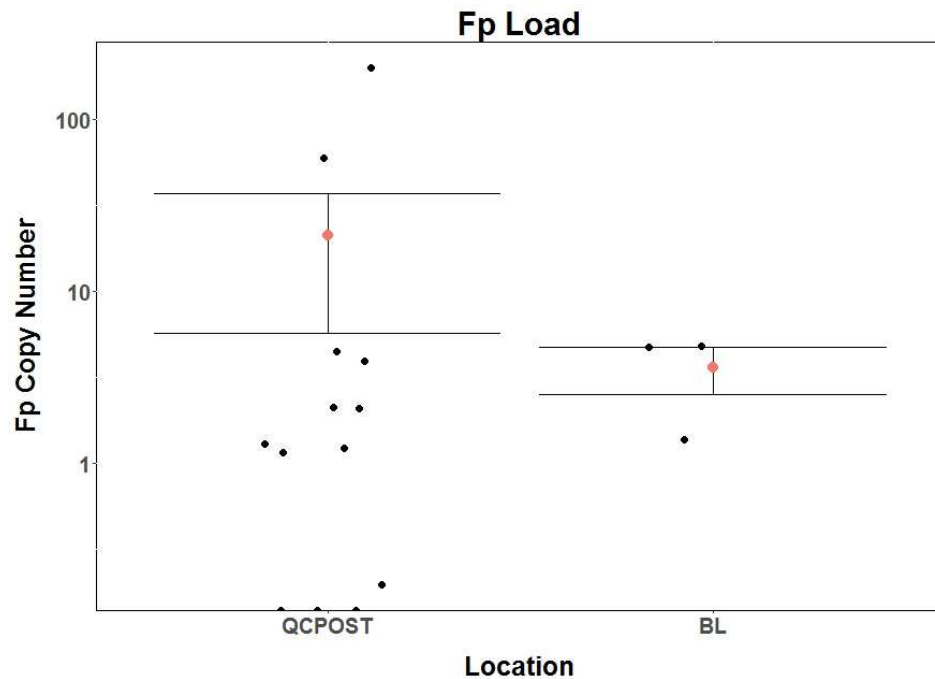


Figure 14: (A.) *Fp* infection. Bar graph showing the prevalence of infection in analyzed fish. The collection site is listed along the X-axis and the relative percentage is listed along the Y-axis. 15 fish from each site were analyzed, the percentage infected is shown in red with the number infected listed within the bar. The percentage uninfected is shown in green with the number of uninfected listed within the bar. (B.) Strip chart showing the measured copy number of *Fp* in infected individuals. Means are shown in orange. Standard error bars are shown, indicating standard error of the mean. Number of copies of the given pathogen is listed on the Y-axis in log scale while the location analyzed is listed on the X-axis.

Analysis of *Fp* was conducted next (Figure 14). Of the pathogens analyzed thus far *Fp* demonstrated by far the lowest frequency across all sites, as well as the lowest average number of copies in infected individuals. Of note here is that qPCR results indicated inconsistency between wells from the same sample. This can in part be

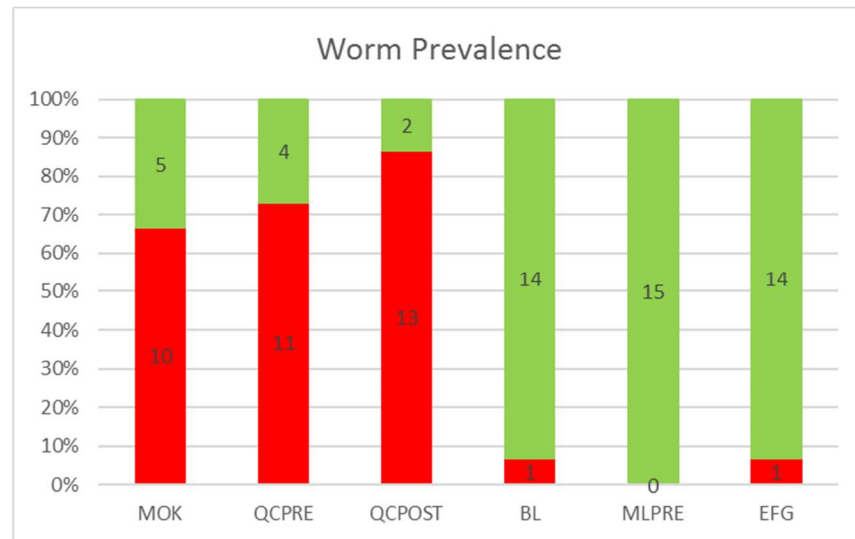
attributed to the low copy numbers of the target itself given that a small number of copies has a greater chance of being unequally distributed between replicates according to the Poisson distribution. Only two sites demonstrated any level of infection with Fp; QCPOST and BL. Interestingly, QCPOST and BL thus far have been positive for every pathogen tested and in every case one of the two has been the site with the highest number of copies and relative frequency of infected fish. Clearly QCPOST and BL fish have a greater propensity for infection with a number of pathogens compared to other sites studied here.

Analysis of the fourth pathological agent, AS, was performed for all sites and samples; however, there were no indications of its presence in any of the fish tested. While the TaqMan qPCR assay is suspected to have worked as the standard still amplified appropriately, there were no experimental samples that amplified at all, hence we could not verify that the assay could detect the pathogen in our fish; potentially due to differences in AS strain.

Lastly, collected fish were examined for the presence of worms, without identifying the species, though they are potentially members of the genus *Anasakis*. The percentage of fish with visible worms was then graphed for each site, as shown in Figure 15.



(A.)



(B.)



Figure 15: (A.) Bar graph showing the presence of worms in analyzed fish. The collection site is listed along the X-axis and the relative percentage is listed along the Y-axis. 15 fish from each site were analyzed, the percentage infected is shown in red with the number infected listed within the bar. The percentage uninfected is shown in green with the number of uninfected listed within the bar. Data gathered for this analysis were measured purely as a “yes” or “no” for the presence or absence of worms, respectively. (B.) Photograph of worms in question in the peritoneal cavity of Sockeye salmon, although not the only type of worm observed.

All pathogen infection prevalences were analyzed using a permutation-based independence test found in the “coin” package of R (Hothorn et al. 2016; Hothorn et al. 2015) (Zeileis et al. 2008). Results of these independence tests are shown in Table V. Additionally, copy numbers (for both infected and uninfected fish) were analyzed using an ANOVA for any significant differences between sites, and once again a MANOVA was used to determine if all three pathogens vary significantly between sites.

**Table V – Results of Pathogen Statistical Tests**

|               | ANOVA Between All Sites For<br>Copy Number of Listed Pathogen | Independence Test of Infection Rates Across<br>all Sites for Listed Pathogens | MANOVA Comparing IHN, Rs, and<br>Fp Across All Sites |
|---------------|---|---|--|
|               | P-Value   | P-Value   | P-Value  |
| IHN           | 0.0635  |   | 0  |
| Rs            | 0.0571  |   | 0.0271   |
| Fp            | 0.118   |   | 0  |
| Worms         | N/A   |   | 2.00E-04   |
| All Pathogens |   |   | 0.02862  |

Table V: P-values are listed for a variety of statistical tests performed. Under the column “ANOVA” an ANOVA was performed for each pathogen to analyze whether there was a significant difference in copy number between any of the sites when looking at a single pathogen. Under “Independence Test” a permutation based Monte Carlo Independence

Test using 10,000 resamplings was performed to determine whether the percentages of infection (e.g. Figure 12A) differ significantly from what would be expected if probability of infection was independent of location. Under “MANOVA” a MANOVA was performed looking at the copy numbers of all of the pathogens analyzed to determine if there was variation between locations. Those values deemed significant according to an  $\alpha=.05$  are highlighted in green.

No significant differences were detected between sites with regards to the copy number of each individual pathogen measured. However, when all of the pathogens are taken together and compared between sites there is a significant difference (via MANOVA); between which two sites this difference exists is unknown. When looking at the prevalence of infection between sites it is clear from the independence tests that the presence of a given pathogen is at least in part dependent upon the location the fish was sampled in. A significant p-value in this test indicates that the assumption of independence of site and infection with each given pathogen is incorrect.

The same test for independence was conducted for each individual site (Table VI).

**Table VI – Results of Site-Specific Independence Tests for Pathogen Infection**

| Pathogen | MoK p-value | QCPRE p-value | QCPOST p-value | BL p-value | MLPRE p-value | EFG p-value |
|----------|-------------|---------------|----------------|------------|---------------|-------------|
| IHNV     | 0.0031      | 0.98          | 0.0007         | 0.00006    | 0.74          | 0.021       |
| Rs       | 0.74        | 0.74          | 0.88           | 0.009      | 0.74          | 0.99        |
| Fp       | 0.38        | 0.38          | 0.000000002    | 0.97       | 0.38          | 0.38        |
| Worms    | 0.12        | 0.024         | 0.0003         | 0.024      | 0.0033        | 0.024       |

Table VI: P-values are listed for a permutation based Monte Carlo Independence Test using 10,000 resamplings of infection rate data for IHNV, Rs, Fp, and worm presence. P-

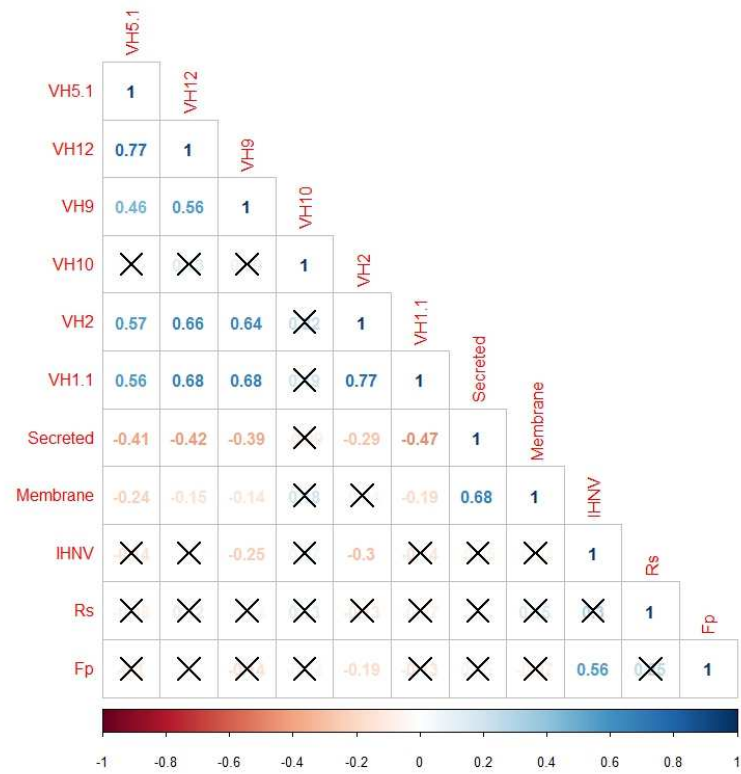
*value indicates the probability that the rates of infection recorded would be observed given independence between site and infection. Pathogen tested is listed under the column “pathogen”, while the p-values are listed for each pathogen at each labeled site. P-values found to be significant at  $\alpha=.05$  are highlighted in green.*

QCPOST and BL demonstrated the lowest p-values, indicating that presence at either of these sites is (for the most part) not independent of infection with the measured pathogens.

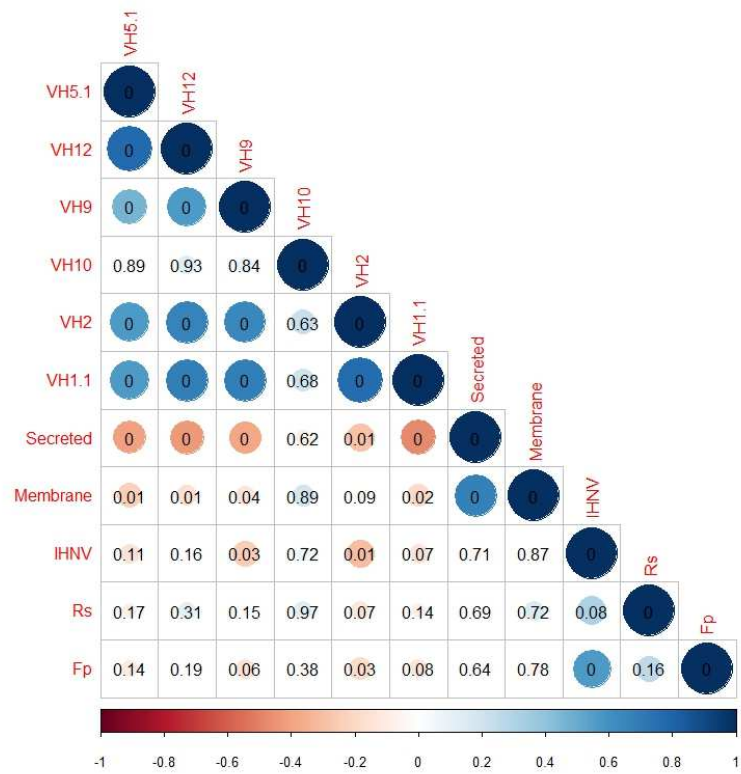
### Overarching Analysis

The data gathered from these various experiments could potentially be viewed together to draw conclusions about the differences that may exist between sampling locations. This overall analysis was accomplished in several ways. First, the data was plotted onto a correlation table (Figure 16) demonstrating the calculated correlations between each of the variables measured.

(A.)



(B.)



*Figure 16: Correlation table showing the relationships between multiple variables.*

*Variables include V<sub>H</sub>5.1, V<sub>H</sub>12, V<sub>H</sub>9, V<sub>H</sub>10, V<sub>H</sub>2, V<sub>H</sub>1.1, SecHCmu and MemHCmu fold change measured relative to an O. nerka reference sample (#349). Variables measured*

*also includes the number of copies of pathogens measured (IHNV, Rs, and Fp). (A.)*

*Variables are listed along the X and Y axes, the square present at the convergence of two variables represents the correlation between the two. Therefore the squares*

*between the same variable will have a complete positive correlation (1.0). Correlations*

*are listed numerically between -1.0 and 1.0, with -1.0 indicating a complete negative*

*correlation and 1.0 indicating a complete positive correlation. The values in between*

*then demonstrate correlations of various strengths (r-values). Positive correlations are*

*shown in varying shades of blue while negative correlations are shown in varying shades*

*of red according to the scale seen at the right of the table. X's indicate non-significance*

*of the listed correlation. Correlation values were calculated according to the non-*

*parametric Spearman's Rho method. (B.) P-values are listed for each of the correlations*

*made in (A). Additionally, size and color of circle indicates the strength and direction of*

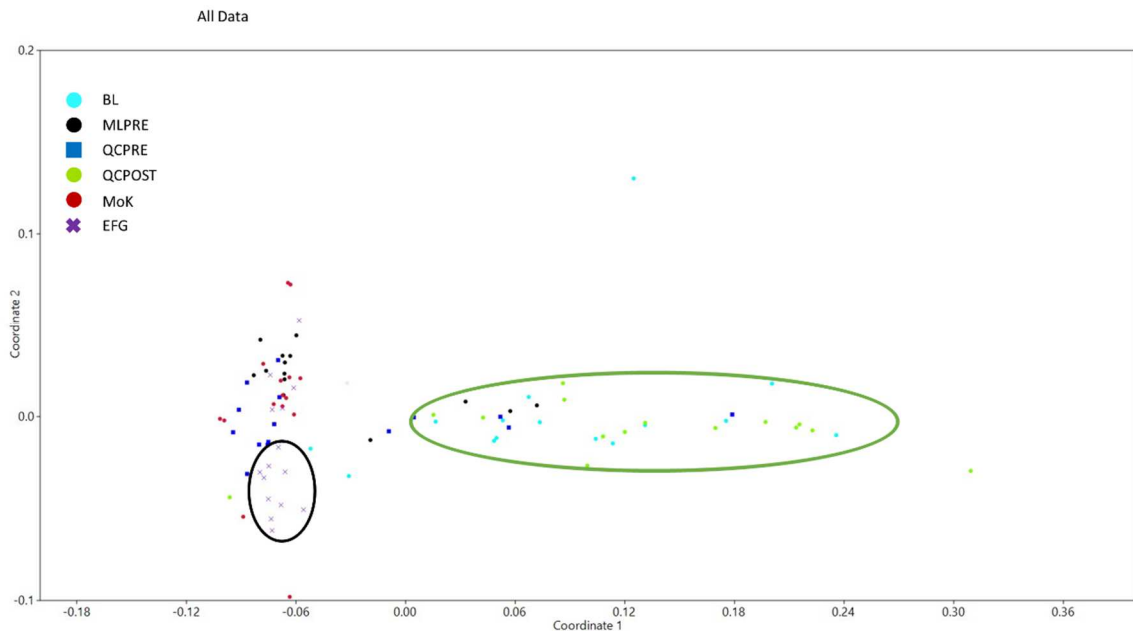
*the correlation.*

When first observing the correlation table it is interesting to note that correlations including the pathogens analyzed were rarely considered significant. There appears to be a slight negative correlation between infection with IHNV and usage of V<sub>H</sub>9 (-0.25), as well as between IHNV and Fp and V<sub>H</sub>2 (-0.3 and -0.19 respectively), but while significant ( $p=.03$ ,  $p=.01$ , and  $p=.03$  respectively), they are weak correlations. There is a moderate positive correlation between infection with BCWD and infection with IHNV (.56), which agrees with previous observations made by others (Greg Wiens, personal communication; ("Research - Wargo Lab" 2016)). With the exception of V<sub>H</sub>10 every V<sub>H</sub> family appears to be positively correlated with the usage of every other V<sub>H</sub> family to

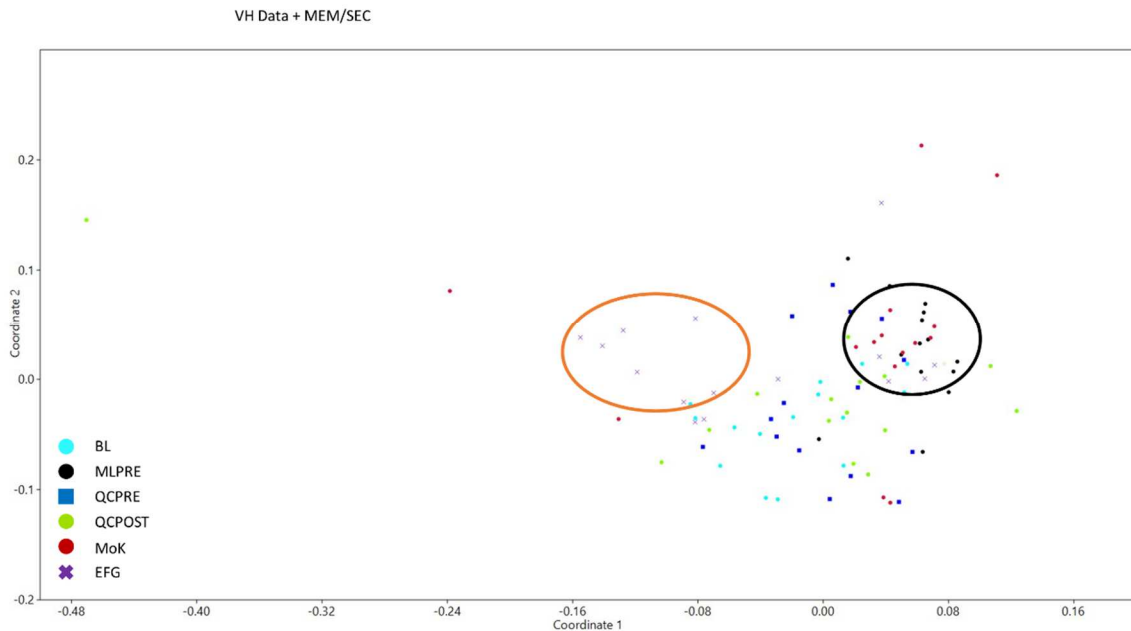
differing degrees ( $r=0.56$ - $r=0.77$ ). Additionally, there are slight negative correlations occurring between  $V_{H5.1}$  ( $r=-0.41$ ),  $V_{H12}$  ( $r=-0.42$ ),  $V_{H9}$  ( $r=-0.39$ ),  $V_{H2}$  ( $r=-0.29$ ), and  $V_{H1.1}$  ( $r=-0.47$ ) with SecHCmu. A similar correlation can be seen between these  $V_H$  families and MemHCmu with the exception that there is no significant relationship between MemHCmu and  $V_{H2}:V_{H5.1}$  ( $r=-0.24$ ),  $V_{H12}$  ( $r=-0.15$ ),  $V_{H9}$  ( $r=-0.14$ ),  $V_{H1.1}$  ( $r=-0.19$ ). Finally, it appears that there is a moderate correlation between MemHCmu and SecHCmu ( $r=0.68$ ,  $p=2.8 \times 10^{-13}$ ).

The data were then analyzed through the use of non-metric dimensional scaling (NMDS); a multivariate ordination technique that simultaneously takes data from multiple variables and plots them on a single 2-dimensional graph for aid in visualization. For this study NMDS was performed with two different sets of data. One set included all data gathered, while the other set limited the data to information on the fold changes of various  $V_H$  families, as well as MemHCmu and SecHCmu. The NMDS analyses can be seen in Figure 17.

(A.)



(B.)



*Figure 17: Non-metric Dimensional Scaling (NMDS) analysis showing the relationships between many variables. Variables measured include  $V_{H5.1}$ ,  $V_{H1.1}$ ,  $V_{H2}$ ,  $V_{H9}$ ,  $V_{H10}$ ,  $V_{H12}$  fold change corrected for  $SecHCmu$  from an *O. nerka* reference (fish #349).  $MemHCmu$  and  $SecHCmu$  corrected for  $\alpha$ -tubulin from an *O. nerka* reference (fish #349). Number of copies of *IHN*V, *Rs*, and *Fp*, as well as the presence or absence of worms. Data was analyzed using a Bray-Curtis similarity index. Sites are identified according to the legend accompanying both charts. Areas of particular interest are circled in black and green. (A.) Analysis including all variables listed above. Shepard stress plot produced a value of 0.07943, within the acceptable range. (B.) Analysis including only  $V_{H5.1}$ ,  $V_{H1.1}$ ,  $V_{H2}$ ,  $V_{H9}$ ,  $V_{H10}$ ,  $V_{H12}$ ,  $MemHCmu$ , and  $SecHCmu$ . Shepard stress plot produced a value of 0.1661, within the acceptable range.*

NMDS plots are often useful for the visualization of how variables cause data points to group together relative to each other. Generally, those points observed closer together demonstrate data that are more similar to each other than those points that are



seen further away. While there is some degree of subjectivity in this type of analysis, there is still value in being able to visualize relationships like this. In the current study we are interested in evaluating the similarities and differences between locations. As such, the data used in this analysis were identified according to location. Each individual color and/or shape indicates a different site analyzed. Circles surround groupings of samples that correspond with individual sampling sites. The closer the individual points are grouped the more alike they are to each other, demonstrating potential similarities that exist between samples within sites, and differences that exist between sites.

## DISCUSSION

While the mechanistic understanding of salmon homing has improved in past years (Ueda 2011), the scientific community is in many ways no closer to understanding why such a costly and challenging life history exists. The Immunological Imprinting Hypothesis provides an explanation for this evolved behavior. While it is likely that many physical and biological aspects of spawning sites vary, pathogens would contribute a great deal toward their uniqueness. Just as distinct pathogens elicit different antibody responses, distinct sites will presumably elicit different patterns of antibody responses. If anadromous fish were encountering these “pathogen fingerprints” for the first time as adults, it would be reasonable to expect that the fish would produce a novel immune response. In reality, however, the ability to produce novel immune responses is limited due to hormonal changes in migrating salmon. It is advantageous then that anadromous fish return to sites that they have inhabited before (their natal grounds), and thus return to pathogens that they have been exposed to as juveniles. As a result of this previous exposure returning fish will have developed immunological memory against the pathogen fingerprint unique to their spawning sites, giving them and their offspring a survival advantage if they return to their natal body of water. As such, we have hypothesized that the mechanism of chemical imprinting might have been driven by the advantage of proper immunological “imprinting” (Zwollo 2012).

The focus of this thesis research was to investigate the Immunological Imprinting Hypothesis. Through qPCR-based analysis of both the antibody usage patterns (“the antibody fingerprint”), as well as pathogen infection patterns (a measure of “the pathogen fingerprint”), several significant differences between fish, dependent on their specific spawning site were shown. The antibody fingerprints were determined in the anterior kidney, while the pathogen fingerprints were investigated using the spleen.

### Patterns of Membrane and Secreted Heavy Chain Mu Expression

Long-lived plasma cells stored in the anterior kidney are thought to be maintained throughout the life span of salmon, including the spawning journey (Schouten et al. 2013). Continuing from this work, in the present study the levels of membrane heavy chain mu transcripts (MemHCmu) and secreted heavy chain mu transcripts (SecHCmu) were analyzed. MemHCmu expression was quite variable between fish and there were no significant differences between sites. SecHCmu expression was relatively constant with no significant differences between sites, although EFG had 3-4 fold higher average SecHCmu expression relative to the other sites.

Based on our previous study (Schouten et al. 2013), we had expected to see a relative decrease in MemHCmu expression as fish approach their spawning site, while simultaneously observing little to no change in SecHCmu expression. The lack of significant differences between SecHCmu expression at different sites was as expected; however, we did not observe the expected decrease in MemHCmu, possibly the result of variation produced as a result of sampling extremely outbred populations of fish. Analysis of the ratio of SecHCmu to MemHCmu transcripts did not show any significant differences between sites.

### Differences in V<sub>H</sub> Family Usage Varies Between Families and Locations

To determine the versatility inherent in immunoglobulin production we examined the V<sub>H</sub> family usage of individual fish at different sampling sites. Usage of six V<sub>H</sub> families was examined between sites; V<sub>H</sub>1.1, V<sub>H</sub>2, V<sub>H</sub>5.1, V<sub>H</sub>9, V<sub>H</sub>10, and V<sub>H</sub>12. Not every site showed significant differences in V<sub>H</sub> usage for each V<sub>H</sub> family, but this may be expected given that V<sub>H</sub> family usage is a very broad way to analyze the antibody fingerprint. In the

future we may find additional differences by examining specific  $V_H$  sequences within each family, and/or  $D_H$  or  $J_H$  segment usage. As it stands, analysis of  $V_H$  families may be too broad to observe all of the differences that may exist between immune responses at different sites.

Clear differences were found between sites for four of the six  $V_H$  families analyzed;  $V_H1.1$ ,  $V_H2$ ,  $V_H5.1$ , and  $V_H12$  (Figure 9). Many of the differences observed were focused around comparisons to a single location, but for different  $V_H$  families at different sites. For example,  $V_H5.1$  usage differed significantly between ML and QCPRE, BL, and EFG, while  $V_H1.1$  usage differed significantly between MoK and QCPRE, QCPOST, BL, and EFG. Additionally, for  $V_H12$ , ML was significantly different from every other location sampled.  $V_H2$  showed the most variability in usage between sites.

While several of the sampling sites were part of the same spawning run (MoK, QCPRE, QCPOST), BL, ML, and EFG were part of completely different runs. BL, while geographically somewhat close to the Kenai run, is actually part of a different spawning run that enters from Resurrection Bay. This means that fish could theoretically be exposed to different environmental variables. Likewise, ML and EFG are both part of the Copper River run, located several hundred miles from the Kenai run. Hence, a larger distance between locations might be expected to produce different environments, leading to different  $V_H$  patterns.

The  $V_H2$  usage patterns were noteworthy for having the most significant relationships of any single  $V_H$  family analysis, with many of the differences observed between MoK and ML. Further,  $V_H2$  usage also showed significant differences between genders, while most other  $V_H$  families did not. Such a difference may indicate a difference in the type of pathogens that affect males vs. females or a potential difference in the response to pathogens. Additionally, differences in  $V_H$  usage between males from

QCPRE and QCPOST exist (Table III). When this is tested a significant difference is evident in  $V_H12$  usage ( $p=.023$ ).

There was only one instance where a single sampling site was significantly different from every other sampling site; the aforementioned  $V_H12$  comparison driven by ML. From a broad view it may be surprising that ML was significantly different from EFG as both are part of the same spawning run. However, EFG and ML are actually many miles apart, meaning that in many ways this comparison is like a comparison between two spawning sites.

MoK demonstrated significant differences relative to many locations for several  $V_H$  families. As the beginning of a spawning run this pattern can be attributed to the idea that there are technically fish from many different eventual spawning locations present. Each of these individual spawning locations should theoretically have a very specific fingerprint of  $V_H$  usage. However, at the start of a spawning run when viewed at the population level these differences would result in a much wider variation than would be observed at the individual spawning sites themselves. In fact, when looking at the individual comparisons of the  $V_H$  families between sites it is interesting to note that the standard error bars shown are almost always highest for the MoK site, indicating the most variation of any sampling site.

The  $V_H2$  usage patterns were noteworthy for having the most significant relationships of any single  $V_H$  family analysis. Neither  $V_H10$  nor  $V_H9$  usage patterns produce significant comparisons between any two sites. This could mean that  $V_H10$  and/or  $V_H9$  are so critical to the immune response that they are necessary for survival of all fish. Another possibility is that there is a single, ubiquitous pathogenic agent at all of the sampling sites analyzed, meaning that across all of the sites analyzed there would be a similar production of these families to combat it. Conversely, usage of  $V_H9$  and

V<sub>H</sub>10 could be of little consequence in terms of a response to pathogens and its usage doesn't matter in the sites we sampled. Certain individuals within sites produce these V<sub>H</sub> families at enhanced rates, but with the sample size used in this study I was unable to establish whether those individuals represented real differences. For example, V<sub>H</sub>10 appeared to have the lowest levels of usage of any of the V<sub>H</sub> families analyzed; except in a few fish. These fish were exclusively found at MoK and were at least 1-2 fold higher than the rest of the sites. High production of V<sub>H</sub>10 could potentially be a critical part of the antibody fingerprint from (a) different site(s) not included in this study.

In addition to standard error being highest at MoK, standard error appears to decrease when moving to the QCPRE sampling location, and again in post-spawned fish from QC (QCPOST). It is possible that this narrowing of the standard error represents the gradual focusing of the hypothesized antibody fingerprint. MoK would potentially have the highest variability. The majority of fish designated QCPRE are likely to have successfully navigated back to their natal stream, but there are potentially fish sampled here that have reached the wrong destination. These "stray" fish in turn are likely to have antibody fingerprints different from those fish that hatched in QC, and thus are less able to deal with the pathogens present. When sampled these stray fish may produce the variability observed. QCPOST fish are those that have successfully spawned. Fish categorized as QCPOST exhibited the least variability in V<sub>H</sub> usage compared to QCPRE and MoK because they returned to the correct spawning site. "Stray" fish included in the QCPRE population may have died before reproducing due to their lack of an effective antibody fingerprint, leaving behind those fish that had successfully navigated to their natal stream and possessed an effective antibody fingerprint. This pattern is most easily observed when looking at V<sub>H</sub>9 usage, but is still somewhat visible when looking at V<sub>H</sub>1.1 and V<sub>H</sub>10. Alternatively, fish categorized as QCPRE may be able to physically use a

broader range of  $V_H$  families, while those from QCPOST may not physically be able to produce a varied  $V_H$  repertoire.

#### The Prevalence of Pathogen Infection is Dependent in Part on Location

Every individual analyzed for  $V_H$  usage patterns was also analyzed for the presence of four pathogens, IHNV, Fp, Rs, and AS. Additionally, the presence or absence of parasitic worms in the body cavity was determined and recorded. All pathogens analyzed were found at the highest percentages in fish at QCPOST and BL. While it might be expected that QCPOST would have a high level of infection with a variety of pathogens given that these fish are likely to have been exposed to the pathogens along the Kenai run for the longest period of time, and are dying, the percentages of infection at BL are somewhat of a surprise. BL, while a spawning site, is also a managed hatchery utilizing artificial fertilization methods. Previous studies have noted that hatchery practices can lead to decreased genetic variation, which in turn can potentially lead to increased susceptibility to pathogens (Peters and Turner 2008). A lack of genetic diversity at BL then could explain the higher than average rates of infection. However, as BL is the only site sampled from this particular spawning run it is also possible that the prevalence of the analyzed pathogens is perhaps naturally higher than in other runs.

Another potential explanation for BL's high rates of infection pertains to the methods used to rear the fish. BL hatchery fish are actually raised using water from the Kenai River. This means that juveniles at the hatchery may be forming an antibody fingerprint effective against Kenai run pathogens, but not against pathogens found in their natural spawning site: BL. As a result, when these fish return to spawn it is possible that the lack of a proper antibody fingerprint results in abnormally high rates of infection.

Infection rates. In the Kenai run there was a clear increase in pathogen infection rate for fish at their spawning site (QC). IHNV infection rates were lowest at MoK, increased for prespawning fish, and were the highest in post-spawned fish. This trend could be caused by prolonged exposure to pathogens, or simply selection for individuals tolerant toward infection. This approaches the idea of tolerance vs. resistance. Tolerance is the capability of a fish to survive with a pathogen load, while resistance is the active reduction of the pathogen load. While arguably the most obvious way to combat infection is through resistance, tolerance can also be a successful strategy. What's more, tolerance has been proven to be heritable, as well as influenced by environmental conditions (Blanchet, Rey, and Loot 2010). As fish approach their spawning site they may come into closer contact with other fish who may already be infected and thus infection is all but unavoidable.

Both sites from the Copper River run; EFG and ML, have relatively low levels of infection for all pathogens analyzed. As at least one of these sites (ML) is a spawning site one might expect there to be a reasonable level of infection given the trend seen from the Kenai run; however, this does not appear to be the case.

Of the pathogens analyzed IHNV was found at every site, although infection levels varied widely between sites, but including MoK; confirming previous studies suggesting that IHNV infection is evident in fish from saltwater (Traxler et al. 1997; Traxler et al. 1993). Rs was also found at every site, although at lower rates, and Fp was found at only two sites (QCPOST and BL). While the reason for differing infection rates is unclear, what is clear is that probability of infection is in part dependent on the site the fish was sampled from. Analysis of site-specific independence of infection rates suggests that QCPOST and BL are most likely to represent a dependence of infection on location. This likely occurs as a result of the high infection rates at these sites when



compared to other sampling sites. Lack of significance at other sites may simply indicate a need for an increased number of sampling sites and samples.

Pathogen loads. Pathogen loads in many ways mirrored pathogen infection rates. For the Kenai run, IHNV copy number increased as fish approached spawning (QCPOST). Copy numbers are highest at QCPOST and BL, just as the infection rates are highest at these two sites. IHNV copies in ML and EFG fish, while mirroring the infection rate (ML being higher than EFG), are quite different despite being from the same run. This in turn reflects the distances between, and the differences possible, between two sites from the same spawning run. Copy numbers of Rs appear in many ways to follow what is seen with IHNV. The highest number of copies seen was from BL, followed by QCPOST, with most other spawning sites demonstrating relatively low copy numbers. Fp continues this trend to an even greater degree. Fp was observed at the lowest level of any pathogen analyzed (except As). Fp was only seen in QCPOST and BL and the number of copies rarely exceeded 100 (Figure 14B). It is possible that the Fp loads at these two locations indicates a geographic bias for its distribution; perhaps Fp has not reached the other locations yet. QCPOST demonstrates clear infection with Fp, but there are no samples with Fp infection from either MoK or QCPRE. One might think that if the Fp loads are so high at QCPOST we might observe it in fish earlier in the run. Possibly Fp exists in a reservoir organism at the QC site (such as Rainbow trout or Dolly Varden), thus when fish arrive they become infected. Alternatively, infection may occur early on in the run but the bacteria may not have reached detectable levels before the fish have spawned. The three sites along the Kenai run broadly demonstrated a relationship between distance travelled and pathogen loads. As fish approached QCPOST pathogen loads increased; however, The Kenai run includes only three

collection sites, and of those only two are geographically distinct so any true correlation is not clear.

Parasitic worms. Analysis of visible parasitic worm infestation demonstrated a clear divide between those sites that are part of the Kenai run and those that are not. All of the sites from the Kenai run demonstrated a very high rate of infestation with parasitic worms while those sites sampled from other runs did not (Figure 15). Conditions in the Kenai run may simply be more conducive to worm growth and infestation than other runs. Alternatively, a necessary intermediate host may not be available in the spawning runs where we do not observe worms. In either case, proper identification would be necessary before investigation of the cause of this disparity could take place.

It is very interesting to note that MoK had a relatively high rate of infection despite being the first site of a run. All of the fish that enter the Kenai run enter from the Mouth of the Kenai, and all fish sampled from MoK are coming directly from the ocean, leading to the question of where the worms actually come from. Worm infection may happen earlier in life, meaning that the fish must live with a persistent worm infection for the majority of their lives. This in turn may help to explain why the rates of infection of various pathogens are so high from sites along the Kenai run; chronic worm infections may wear down the immune system and make the animals susceptible to attack from pathogens (Petney and Andrews 1998; Fenton 2013). However, both QCPOST and BL demonstrated the highest rates of infection with all of the pathogens analyzed, but QCPOST has a high rate of worm infection while BL does not.

#### Total V<sub>H</sub> Usage Analysis

Analysis of individual V<sub>H</sub> family usage doesn't necessarily investigate the *patterns* present at the sampling sites. To more accurately visualize the overall differences in V<sub>H</sub>

usage patterns between sites the relative fold changes of each  $V_H$  family were plotted on pie charts (Figure 11). Each pie chart only takes into account the fold change of the  $V_H$  family relative to the other  $V_H$  families from that individual site, but since the fold changes are calculated from a single reference fish the values are standardized to some degree. Regardless of the actual level of usage  $V_H9$  was always used at the highest relative frequency, at times constituting more than half of the antibody fingerprint. Additionally,  $V_H2$  usage appears to be reasonably low across all sites.  $V_H10$  varies widely between sites, constituting a fairly large portion of the MoK fingerprint, while being almost non-existent at any of the other sites. This suggests that  $V_H10$  may have greater use against saltwater pathogens relative to freshwater pathogens.  $V_H12$  constitutes a similar proportion of the antibody fingerprint across all sites, and the same can be seen with  $V_H1.1$ , although with more variation.  $V_H5.1$  also differs greatly between sites, most notably within the sites of the Kenai run. At the mouth of the Kenai run  $V_H5.1$  represents a relatively small portion of the fingerprint. This portion increases at QCPRE and again in the post-spawned fish of QC. This may demonstrate the importance of  $V_H5.1$  retention as fish approach the QC spawning site.

While the pie charts do not inherently include fold changes between sites, the calculations involving MANOVA and Hotelling- $T^2$  tests do (Table IV). Significant differences were found between many of the spawning sites analyzed, but not between spawning sites and MoK. The majority of comparisons involving two spawning sites were significantly different when all  $V_H$  family usages are taken into account. This indicates that the fish from these sites produce antibody responses that differ significantly from one another. MoK may not produce significant comparisons because it represents the mouth of a spawning run, and thus has a much broader overall  $V_H$  usage. A lack of a significant difference between QCPRE and QCPOST was expected as these fish are at

the same spawning site. If  $V_H$  usage is largely dictated by environmental variables these fish should have close to the same antibody fingerprint. The remaining two spawning site comparisons that were not significantly different both included BL. While the comparison between BL and EFG was not significantly different ( $p=.0097$ ), indicating a reasonable similarity, the comparison between BL and QCPOST displayed the highest p-value of any comparison ( $p=0.48$ ), indicating that the two sites are likely very similar in the  $V_H$  patterns measured.

Probability of infection with a pathogen was at least partially dictated by spawning location (Table V); certain sites had a higher chance of infection than others. However, this did not necessarily result in a shift in  $V_H$  usage patterns. There are a number of reasons why differences observed in the antibody fingerprint may not be observed in the analysis of the pathogen fingerprint. Most obviously, it is possible that the pathogens analyzed were not representative of the entire pathogen fingerprint of a given site. The pathogens analyzed in this study, while intended to include certain well-known infectious agents, are by no means a complete measure of the pathogen fingerprint of a given spawning site. While changes to the antibody fingerprint can be ascertained via analysis of variation in the same  $V_H$  families, analysis of the pathogen fingerprint through analysis of infection patterns must single out specific pathogens for analysis. This means that it is more likely to observe differences at the level of antibody fingerprint than at the level of pathogen fingerprint when using this approach.

In the future it may be worthwhile to conduct RAPD-PCR analysis of viral communities at different sites. RAPD-PCR uses a non-specific primer to amplify any viral sequences that are present in a sample. These amplified sequences are then run on a gel and a distinct pattern is observed based on which viral sequences are amplified, and how well those sequences are amplified. While this method does not specifically identify

viruses present, it could illustrate differences that exist between sites. This procedure would likely be performed on water samples from the selected site, giving a more direct evaluation of what viruses are present. This method has been used to effectively identify changes in viral community structure in freshwater in the past (Williamson et al. 2014). However, this method evaluates a broad swath of the viral community, meaning that it is not limited to those viruses that are pathogenic towards fish. Therefore much of the variation observed may not directly affect the salmon.

Incorporation of the pathogen data becomes especially interesting in evaluating the QCPOST vs. BL comparison. As previously mentioned the QCPOST vs. BL  $V_H$  usage comparison produced a p-value of 0.48, indicating that the two sites are seemingly very similar in antibody composition. QCPOST and BL were also found to have the highest infection rates and loads for the pathogens analyzed out of all of the sampling sites. It is possible that the high rates of infection with these specific pathogens are associated with the specific antibody fingerprints at the two sites. In other words, a similarity in pathogen fingerprint could lead to a similarity in antibody fingerprint. As mentioned previously, if we aren't looking at the correct array of pathogens then we are likely to encounter difficulties in attempting to define a pathogen fingerprint. At these two sites the pathogens analyzed may represent a significant portion of the pathogen fingerprint. So theoretically similarities in pathogen fingerprint may produce similarities in antibody fingerprint. Conversely, the differences observed in antibody fingerprint may represent differences in the pathogen fingerprints that can't be seen looking for these specific pathogens. The pathogen fingerprints of these two sites likely are not identical, they just demonstrate similarities when these select pathogens are the measure of the difference.

#### Grouping According to Pathogen Presence

NMDS analysis using all data gathered (Mem/Sec/V<sub>H</sub>/Pathogen/Worms) shows several groupings of points. It appears that the majority of samples from EFG are grouped together rather tightly (circled in black). Additionally, there is a grouping that migrates far to the right relative to all of the other points (circled in green). These points are almost exclusively part of the QCPOST and BL sites. When compared to the chart including only Mem/Sec/V<sub>H</sub> data (Figure 17B) this migration to the right does not occur, indicating that the movement to the right is caused primarily by the pathogen data. Knowing the rates of infection at QCPOST and BL this may be expected, but even without the pathogen data the QCPOST and BL points are still relatively grouped together. This in itself lends support for the idea of an antibody fingerprint. Even without the pathogen data included the effect is such on the V<sub>H</sub> usage that the two sites that demonstrated the greatest rates of infection are more like each other than all other sampled sites

An important point to consider here is that if juveniles are exposed to pathogens earlier in life they should theoretically have resistance to those pathogens later in life. The question then becomes, if there is an antibody fingerprint in response to the pathogen fingerprint, why are so many of the fish found to be infected? It is possible that given none of the fish are going to survive spawning it may be advantageous to develop an antibody fingerprint that, while not capable of effective resistance to pathogens, is able to increase the tolerance of the fish to the pathogens so that they can survive long enough to spawn. This would explain why at certain sites a large percentage of fish are infected while still presenting an antibody fingerprint.

#### Grouping According to V<sub>H</sub> Usage

When pathogen data are not included several groupings of points can be observed. First, near to the center of the chart (circled in orange) the points representing

EFG are grouped together, while to the right the points of ML and MoK are grouped together (circled in black). These groupings indicate that these fish share characteristics; making them more like fish from the same location than from other locations. This in turn means that spawning location is a critical factor affecting the variables being measured. If spawning location were not an important factor we would expect to observe an essentially homogeneous mixture of points, with all of the different colors and shapes being evenly dispersed.

Interestingly, the points representing QCPRE appear to be broadly distributed across the plot. This could potentially indicate that the variables measured at QCPRE demonstrated a wide degree of variance that prevented fish from being seen as similar to each other. Additionally, the grouping of ML and MoK is quite puzzling. While we do expect the points to group with other points from the same site, we do not necessarily expect the points of two different sites to group so strongly with each other. MoK and ML should theoretically be quite different; MoK is the start of a spawning run, while ML is a spawning point from an entirely different spawning run. However, it is possible that this comigration may occur as a result of the large amount of variance inherent in MoK samples. On the other hand, MoK and ML fish may display similar  $V_H$  usage and MemHCmu/SecHCmu patterns despite their differences in location. The cause of such a similarity is unknown, but perhaps an additional pathogen not analyzed in this study could be present at both sites and influence these fish to display similar characteristics.

### Visible Correlations

The significant correlations from the correlation table offer an interesting insight into the interaction of the variables measured in this study. The  $V_H$  families (with the exception of  $V_H10$ ) are all correlated with each other to varying degrees ( $r=0.46$ - $r=0.77$ ). This may mean that if the fish is still able to produce antibodies with  $V_H$  gene segments

from one  $V_H$  family, they are likely able to produce antibodies using gene segments from other  $V_H$  families. Perhaps high levels of  $V_H$  family usage indicates enhanced survival of the cells that are producing the antibodies. However, there is a negative correlation between all of the  $V_H$  families (except  $V_{H10}$ ) and secreted and membrane expression. If the positive correlation observed between  $V_H$  families is a result of the population of antibody producing cells, we would expect to observe a positive correlation between at least SecHCmu and the  $V_H$  families, however the opposite is true. The negative correlation between each of the  $V_H$  families and SecHCmu expression could suggest that production of each of the individual  $V_H$  families results in a decreased *need* for secreted IgM. In other words, increases in “effective”  $V_H$  family usage reduces the total amount of SecHCmu expression necessary, because the individual  $V_H$  families are better able to successfully deal with the immune challenge.

MemHCmu transcripts are mostly generated by resting B cells. The slight negative correlation observed between MemHCmu expression and the individual  $V_H$  families could indicate that when individual  $V_H$  family usage is high for production of SecHCmu by (long-lived) plasma cells, relative levels of MemHCmu, expressed on non-stimulated B cells, is low.

It is interesting to see that there is a slight, though significant, negative correlation between  $V_{H9}$  and IHNV, and between  $V_{H2}$  and IHNV and BCWD. In the case of  $V_{H9}$  this indicates that a higher usage level of  $V_{H9}$  results in a lower level of IHNV measured. For  $V_{H2}$  it appears that a higher level of  $V_{H2}$  usage results in lower IHNV and lower BCWD. These correlations could represent the basis for understanding a proper immune response towards two of the pathogens analyzed.

There was a reasonably strong correlation between SecHCmu and MemHCmu expression ( $r=0.68$ ), suggesting that expression of these two is linked. This could



potentially be explained in the general health of the fish being analyzed. Fish that are capable of producing MemHCmu are also able to produce SecHCmu and vice versa. On the other hand, those fish that are potentially less healthy at the time of sampling might lack the cells needed to produce either MemHCmu or SecHCmu. As we cannot distinguish between mature resting B cells (mem+), plasmablasts (mem+/sec+), and plasma cells (sec+) using this method of analysis, additional flow cytometric data are needed. Finally, there appears to be a relatively strong correlation ( $r=0.56$ ) between IHNV and Fp indicating that there is a tendency for coinfection with these two pathogens in the wild. Interestingly, this correlation has been observed by others (Greg Wiens, personal communication, ("Research - Wargo Lab" 2016))

### Conclusions and Future Directions

Taken together my research suggests answers for some questions while simultaneously posing numerous additional ones. The antibody fingerprints of many of the spawning locations sampled were indeed different – as predicted by the Immunological Imprinting Hypothesis. These differences appear to exist both at the antibody level, as well as at the pathogen level. While the differences in  $V_H$  usage patterns observed might be expected, this study does not directly assess whether these patterns are a result of production of LLPC's in juvenile fish. Additionally, I have not ascertained whether fish that stray to different spawning sites experience higher rates of mortality than those fish that successfully return to their natal site. While certainly interesting questions, the logistical hurdles associated with monitoring anadromous fish throughout their life cycles are considerable.

In the future, information gathered through long-term analysis following fish through their lives (perhaps through the use of radio tags or other means) would be invaluable. Fish could be analyzed from fry to spawning and clear conclusions could be

drawn with regards to the specificity of  $V_H$  usage patterns throughout the lives of individual fish. Additionally, continued collection of samples to further define the antibody fingerprints associated with different spawning sites would be quite useful. More data could allow for continued evaluation of potential changes in antibody fingerprints over time; and perhaps in response to potential environmental perturbations – events that are likely only to increase in frequency. While the current study only analyzed four distinct pathogens, there are likely a wide variety of other significant pathogens affecting wild fish. Increasing the number of pathogens analyzed and increasing the number of fish samples would enhance our ability to investigate correlations that might exist between pathogens and the  $V_H$  families used to combat them.

The Immunological Imprinting Hypothesis suggests an answer to a question that has been around since humans first observed anadromy in nature. While this research does not conclusively prove the hypothesis true, my data support the hypothesis in that different spawning sites demonstrate unique patterns of antibody usage, suggesting that the fish are responding to pathogens unique to those sites.

Appendix.

Supplementary Table I.

| Sample | Site  | Gender | VH5.1    | VH12     | VH9      | VH10     | VH2      | VH1.1    | Worms | Secreted | Membrane | IHNV -<br>SPL<br>Copy<br>Number | Rs -<br>SPL<br>Copy<br>Number | Fp - SPL<br>Copy<br>Number | IN -<br>Copy<br>Number | Year |
|--------|-------|--------|----------|----------|----------|----------|----------|----------|-------|----------|----------|---------------------------------|-------------------------------|----------------------------|------------------------|------|
| 361    | BL    | Female | 0.395934 | 0.190342 | 0.463294 | 0.073302 | 0.072796 | 0.188156 | 0     | 3.024437 | 0.331405 | 132.6                           | 3.05                          | 0                          | 0                      | 2013 |
| 362    | BL    | Female | 0.276752 | 0.401461 | 2.089754 | 0.001931 | 0.123564 | 0.225313 | 0     | 0.381565 | 0.061925 | 7583.333                        | 11.81                         | 4.696                      | 0                      | 2013 |
| 363    | BL    | Female | 0.40239  | 0.383332 | 0.899171 | 0.066523 | 0.269807 | 0.332171 | 0     | 0.664343 | 1.295342 | 14390                           | 5845                          | 1.36                       | 0                      | 2013 |
| 364    | BL    | Female | 0.814131 | 0.888843 | 2.17347  | 0.003808 | 0.139984 | 0.933033 | 0     | 0.240927 | 0.218141 | 63.32                           | 7757                          | 0                          | 0                      | 2013 |
| 365    | BL    | Male   | 0.586417 | 0.695762 | 1.681793 | 0.112396 | 0.152477 | 0.451668 | 0     | 0.597358 | 1.130269 | 199.4                           | 127.7                         | 0                          | 0                      | 2013 |
| 366    | BL    | Male   | 0.672062 | 0.460094 | 0.954842 | 0.014445 | 0.075887 | 0.350301 | 0     | 0.680657 | 1.855318 | 301.62                          | 29.8                          | 0                          | 0                      | 2013 |
| 550    | BL    | Female | 0.143919 | 0.120185 | 0.526072 | 0.035814 | 0.127922 | 0.150726 | 0     | 2.1386   | 0.366868 | 1063                            | 0                             | 4.77                       | 0                      | 2015 |
| 551    | BL    | Female | 0.697372 | 0.113702 | 0.681444 | 0.006479 | 0.261824 | 0.740549 | 0     | 0.842842 | 0.097171 | 41.31                           | 0                             | 0                          | 0                      | 2015 |
| 552    | BL    | Female | 0.371989 | 0.178006 | 0.82932  | 0.054409 | 0.163044 | 0.378929 | 0     | 1.292353 | 0.17234  | 2.63                            | 0                             | 0                          | 0                      | 2015 |
| 553    | BL    | Female | 0.254664 | 0.094296 | 0.582367 | 0.007494 | 0.173941 | 0.271057 | 0     | 1.217004 | 0.289841 | 785.7                           | 0                             | 0                          | 0                      | 2015 |
| 554    | BL    | Female | 0.260616 | 0.199344 | 0.350301 | 0.022457 | 0.181747 | 0.312083 | 0     | 1.60956  | 0.422396 | 54060                           | 0                             | 0                          | 0                      | 2015 |
| 555    | BL    | Male   | 0.420448 | 0.230047 | 0.054788 | 0.002022 | 0.12385  | 0.021246 | 0     | 1.178267 | 0.171744 | 5.49                            | 0                             | 0                          | 0                      | 2015 |
| 556    | BL    | Male   | 0.138696 | 0.120463 | 0.257028 | 0.037508 | 0.074497 | 0.206374 | 0     | 0.632878 | 0.106088 | 1569                            | 279                           | 0                          | 0                      | 2015 |
| 557    | BL    | Male   | 0.100134 | 0.106579 | 0.381565 | 0.002031 | 0.048361 | 0.240371 | 0     | 0.579682 | 0.094078 | 88.21                           | 0                             | 0                          | 0                      | 2015 |
| 558    | BL    | Male   | 0.288505 | 0.222211 | 0.517632 | 0.007529 | 0.092355 | 0.469761 | 1     | 0.345079 | 0.115957 | 131.1                           | 0                             | 0                          | 0                      | 2015 |
| 332    | MLPRE | Male   | 2.027919 | 2.07053  | 1.24545  | 0.018841 | 0.543367 | 0.63728  | 0     | 0.013634 | 0.014478 | 0                               | 0                             | 0                          | 0                      | 2013 |

|     |       |        |          |          |          |          |          |          |   |          |          |       |      |   |   |      |
|-----|-------|--------|----------|----------|----------|----------|----------|----------|---|----------|----------|-------|------|---|---|------|
| 333 | MLPRE | Male   | 2.271009 | 1.990779 | 1.183724 | 0.020333 | 0.53465  | 1.401204 | 0 | 0.037163 | 0.014612 | 316.7 | 0    | 0 | 0 | 2013 |
| 334 | MLPRE | Female | 1.981601 | 1.658639 | 1.569168 | 0.036991 | 0.293887 | 1.154019 | 0 | 0.016064 | 0.009291 | 7.19  | 0    | 0 | 0 | 2013 |
| 490 | MLPRE | Male   | 1.233992 | 0.370274 | 0.97041  | 0.012633 | 0.329877 | 2.123828 | 0 | 0.000115 | 6.65E-05 | 0     | 0    | 0 | 0 | 2014 |
| 491 | MLPRE | Male   | 1.22264  | 0.950439 | 5.566099 | 0.036906 | 1.453973 | 1.617015 | 0 | 2.98E-05 | 2.87E-05 | 0     | 6    | 0 | 0 | 2014 |
| 492 | MLPRE | Female | 1.257013 | 0.743979 | 7.638733 | 0.002444 | 2.378414 | 2.1386   | 0 | 0.716978 | 0.553504 | 0     | 0    | 0 | 0 | 2014 |
| 493 | MLPRE | Male   | 2.056228 | 1.292353 | 1.794191 | 0.00087  | 0.643197 | 0.862542 | 0 | 0.7457   | 0.309212 | 0     | 0    | 0 | 0 | 2014 |
| 494 | MLPRE | Female | 3.538979 | 1.140764 | 4.510644 | 0.040667 | 0.831238 | 1.180993 | 0 | 0.289841 | 0.159689 | 0     | 0    | 0 | 0 | 2014 |
| 495 | MLPRE | Female | 2.841527 | 1.725084 | 2.560928 | 0.003529 | 1.081725 | 1.353474 | 0 | 0.380684 | 0.204948 | 0     | 0    | 0 | 0 | 2014 |
| 496 | MLPRE | Female | 3.182146 | 1.310393 | 2.394957 | 0.025149 | 0.89296  | 1.164734 | 0 | 0.441351 | 0.302848 | 0     | 0    | 0 | 0 | 2014 |
| 498 | MLPRE | Female | 2.719485 | 0.939523 | 5.302478 | 0.002687 | 1.154019 | 1.453973 | 0 | 0.160058 | 0.120185 | 0     | 0    | 0 | 0 | 2014 |
| 525 | MLPRE | Female | 1.251218 | 0.484085 | 1.420764 | 0.080959 | 0.903335 | 0.988514 | 0 | 0.103905 | 0.136156 | 78.24 | 0    | 0 | 0 | 2015 |
| 526 | MLPRE | Male   | 0.4954   | 0.529732 | 0.289172 | 0.043485 | 0.340722 | 0.225313 | 0 | 0.710382 | 0.189903 | 10.69 | 0    | 0 | 0 | 2015 |
| 527 | MLPRE | Female | 1.286395 | 0.326088 | 0.355191 | 0.050299 | 0.175962 | 0.300756 | 0 | 0.003065 | 0.00135  | 139.4 | 0    | 0 | 0 | 2015 |
| 528 | MLPRE | Male   | 0.351111 | 0.274206 | 8.49531  | 0.020665 | 0.175962 | 0.298679 | 0 | 0.297302 | 0.0625   | 0     | 0    | 0 | 0 | 2015 |
| 225 | QCPRE | Male   | 0.907121 | 0.133296 | 0.166013 | 0.00242  | 0.041599 | 0.085735 | 1 | 0.5042   | 0.029792 | 0     | 0    | 0 | 0 | 2012 |
| 226 | QCPRE | Female | 1.780028 | 0.156247 | 4.403244 | 0.000477 | 0.151274 | 0.17098  | 0 | 0.40084  | 0.021487 | 0     | 1.88 | 0 | 0 | 2012 |
| 227 | QCPRE | Female | 0.865297 | 0.177751 | 0.853384 | 0.018512 | 0.130724 | 0.145383 | 1 | 0.021391 | 0.001109 | 0     | 0    | 0 | 0 | 2012 |
| 228 | QCPRE | Male   | 0.628666 | 0.273643 | 0.415725 | 0.000447 | 0.03683  | 0.162334 | 1 | 0.33948  | 0.02717  | 14.61 | 0    | 0 | 0 | 2012 |
| 229 | QCPRE | Male   | 0.642351 | 0.09817  | 0.259073 | 3.89E-05 | 0.016155 | 0.020974 | 0 | 1.799588 | 0.067697 | 0     | 0    | 0 | 0 | 2012 |
| 455 | QCPRE | Female | 0.692555 | 0.121582 | 0.965936 | 0.000541 | 0.087575 | 0.279968 | 0 | 1.725084 | 0.279968 | 181.1 | 0    | 0 | 0 | 2014 |
| 461 | QCPRE | Female | 1.594753 | 0.838956 | 1.94082  | 0.002333 | 0.149685 | 1.079228 | 1 | 0.48971  | 0.243164 | 0     | 0    | 0 | 0 | 2014 |

|     |        |        |          |          |          |          |          |          |   |          |          |        |       |          |   |      |
|-----|--------|--------|----------|----------|----------|----------|----------|----------|---|----------|----------|--------|-------|----------|---|------|
| 463 | QCPRE  | Female | 1.654811 | 0.421421 | 0.780967 | 0.005732 | 0.293209 | 0.125579 | 0 | 0.707107 | 0.733736 | 30.47  | 0     | 0        | 0 | 2014 |
| 571 | QCPRE  | Female | 0.366021 | 0.117984 | 0.607097 | 0.00456  | 0.133046 | 0.231647 | 1 | 1.084227 | 0.243726 | 0      | 0     | 0        | 0 | 2015 |
| 572 | QCPRE  | Male   | 0.156764 | 0.181747 | 5.278032 | 0.165702 | 0.08362  | 0.170755 | 1 | 0.469761 | 0.103905 | 0      | 0     | 0        | 0 | 2015 |
| 574 | QCPRE  | Male   | 0.922316 | 0.181747 | 0.449585 | 0.050883 | 0.066986 | 0.282567 | 1 | 1.689582 | 0.245421 | 0      | 0     | 0        | 0 | 2015 |
| 575 | QCPRE  | Female | 0.779165 | 0.158952 | 0.260616 | 0.030678 | 0.083235 | 0.202127 | 1 | 0.006244 | 0.000686 | 106.8  | 0     | 0        | 0 | 2015 |
| 576 | QCPRE  | Male   | 0.609909 | 0.225313 | 0.289172 | 0.039282 | 0.120463 | 0.206851 | 1 | 0.702222 | 0.280616 | 0.6167 | 0     | 0        | 0 | 2015 |
| 578 | QCPRE  | Female | 0.385108 | 0.261219 | 4.913213 | 0.004007 | 0.103665 | 0.128812 | 1 | 2.084932 | 0.339151 | 0      | 0     | 0        | 0 | 2015 |
| 579 | QCPRE  | Male   | 0.605696 | 0.278677 | 4.901874 | 0.00499  | 0.177186 | 0.205898 | 1 | 0.496546 | 0.052073 | 9434   | 0     | 0        | 0 | 2015 |
| 371 | QCPOST | Female | 0.366021 | 0.366021 | 0.174343 | 0.147283 | 0.066986 | 0.136787 | 1 | 1.844632 | 0.246558 | 36060  | 79.76 | 4.42     | 0 | 2013 |
| 372 | QCPOST | Male   | 0.533416 | 0.582367 | 0.293887 | 0.012062 | 0.055424 | 0.227405 | 1 | 0.248847 | 0.043485 | 246.9  | 206.1 | 59.126   | 0 | 2013 |
| 373 | QCPOST | Male   | 0.515246 | 0.463294 | 0.370274 | 0.044915 | 0.272312 | 0.34151  | 1 | 1.372367 | 1.725084 | 1073   | 15.56 | 0        | 0 | 2013 |
| 374 | QCPOST | Female | 0.672062 | 0.493116 | 0.795536 | 0.048585 | 0.419478 | 0.388683 | 1 | 1.162046 | 0.242043 | 29810  | 28.56 | 1.213    | 0 | 2013 |
| 456 | QCPOST | Female | 4.702192 | 1.091768 | 0.650671 | 0.002794 | 0.6846   | 0.790041 | 1 | 0.036736 | 0.003988 | 373.3  | 0     | 199.033  | 0 | 2014 |
| 457 | QCPOST | Male   | 3.301984 | 0.590496 | 1.936341 | 0.002449 | 0.129109 | 0.756109 | 1 | 2.042024 | 0.390483 | 143.4  | 0     | 0        | 0 | 2014 |
| 458 | QCPOST | Female | 0.002668 | 0.000941 | 2.07053  | 0.00123  | 0.000538 | 0.390483 | 1 | 0.009866 | 0.0014   | 1752   | 0     | 3.926667 | 0 | 2014 |
| 459 | QCPOST | Male   | 0.096055 | 0.087171 | 0.074842 | 9.89E-05 | 0.069509 | 0.036567 | 0 | 1.280464 | 0.411796 | 0      | 0     | 0        | 0 | 2014 |
| 460 | QCPOST | Female | 1.628263 | 0.866537 | 0.858565 | 0.013292 | 0.153893 | 0.41851  | 1 | 1.186463 | 0.197967 | 48.8   | 0     | 1.296667 | 0 | 2014 |
| 462 | QCPOST | Female | 0.025442 | 0.006479 | 0.016326 | 1.36E-05 | 0.004196 | 0.005025 | 1 | 118.0565 | 0.259415 | 537    | 0     | 2.09     | 0 | 2014 |
| 464 | QCPOST | Male   | 1.914101 | 0.355191 | 0.366021 | 0.014082 | 0.168404 | 0.223756 | 1 | 0.899171 | 0.224274 | 1516   | 0     | 0.196    | 0 | 2014 |
| 161 | MoK    | ?      | 3.697799 | 0.41466  | 2.657372 | 0.206851 | 0.433269 | 0.935191 | 0 | 0.390483 | 0.257623 | 0      | 0     | 0        | 0 | 2011 |
| 162 | MoK    | Male   | 1.074253 | 2.40605  | 17.18804 | 12.32344 | 1.125058 | 5.869889 | 0 | 0.000699 | 0.003645 | 0      | 0     | 0        | 0 | 2011 |

|     |     |        |          |          |          |          |          |          |   |          |          |        |        |   |   |      |
|-----|-----|--------|----------|----------|----------|----------|----------|----------|---|----------|----------|--------|--------|---|---|------|
| 163 | MoK | Male   | 0.002285 | 0.00135  | 22.62742 | 0.056983 | 4.208579 | 8.876556 | 1 | 0.371131 | 0.496546 | 0      | 0      | 0 | 0 | 2011 |
| 164 | MoK | Male   | 0.381565 | 0.222211 | 0.324585 | 0.048585 | 0.192999 | 0.334482 | 1 | 0.02024  | 0.009099 | 0      | 0      | 0 | 0 | 2011 |
| 165 | MoK | Male   | 0.310644 | 0.102238 | 0.548412 | 0.043285 | 0.090454 | 0.324585 | 1 | 0.006346 | 0.002027 | 0      | 0      | 0 | 0 | 2011 |
| 340 | MoK | Female | 0.045227 | 0.00525  | 0.228458 | 0.002238 | 0.073983 | 0.066064 | 0 | 21.95788 | 10.31496 | 0      | 0      | 0 | 0 | 2013 |
| 341 | MoK | Female | 0.512871 | 0.508152 | 2.854689 | 0.003464 | 1.474269 | 1.057018 | 0 | 0.00822  | 0.005002 | 0      | 0      | 0 | 0 | 2013 |
| 342 | MoK | Female | 1.248331 | 0.961483 | 2.894538 | 0.004624 | 1.159364 | 1.909683 | 1 | 1.265757 | 0.304955 | 0      | 0      | 0 | 0 | 2013 |
| 343 | MoK | Male   | 0.029701 | 0.001962 | 0.130007 | 0.004425 | 0.080959 | 0.051237 | 0 | 3.045474 | 0.279968 | 0      | 0      | 0 | 0 | 2013 |
| 344 | MoK | Male   | 1.219819 | 0.463294 | 2.234574 | 0.003424 | 0.761368 | 1.681793 | 1 | 1.681793 | 7.22E-05 | 0      | 0      | 0 | 0 | 2013 |
| 345 | MoK | Male   | 1.697408 | 0.868541 | 2.060984 | 0.002197 | 0.97716  | 1.023374 | 1 | 0.858565 | 0.000375 | 0      | 0      | 0 | 0 | 2013 |
| 346 | MoK | Female | 0.582367 | 0.502316 | 1.874709 | 0.002846 | 0.624165 | 0.844791 | 1 | 0.68302  | 3.34E-05 | 0      | 0      | 0 | 0 | 2013 |
| 347 | MoK | Male   | 1.36604  | 0.565135 | 4.913213 | 0.05954  | 0.911722 | 2.763826 | 1 | 0.409897 | 1.047294 | 0.8368 | 0      | 0 | 0 | 2013 |
| 348 | MoK | Female | 1.725084 | 1.242575 | 2.688249 | 2.537369 | 0.76313  | 1.914101 | 1 | 0.284533 | 0.607097 | 0      | 0      | 0 | 0 | 2013 |
| 349 | MoK | Female | 1        | 1        | 1        | 1        | 1        | 1        | 1 | 1        | 1        | 0      | 1.68   | 0 | 0 | 2013 |
| 327 | EFG | Male   | 0.52003  | 0.424352 | 2.467984 | 0.252321 | 0.455335 | 0.552227 | 1 | 0.621288 | 0.226356 | 0.4469 | 0      | 0 | 0 | 2013 |
| 329 | EFG | Female | 0.300062 | 0.230579 | 2.123828 | 0.216634 | 0.268874 | 0.217638 | 0 | 6.932296 | 1.785919 | 0      | 0      | 0 | 0 | 2013 |
| 331 | EFG | Male   | 1.375542 | 0.7457   | 23.42537 | 0.020428 | 0.770215 | 0.95705  | 0 | 1.42405  | 0.174343 | 0      | 0      | 0 | 0 | 2013 |
| 466 | EFG | Female | 1.54043  | 0.404254 | 1.617015 | 0.014816 | 0.459032 | 0.614152 | 0 | 0.049721 | 0.030046 | 0.9195 | 0      | 0 | 0 | 2014 |
| 467 | EFG | Male   | 2.318728 | 0.573024 | 1.666321 | 0.01278  | 0.720298 | 0.771997 | 0 | 0.023848 | 0.018841 | 0      | 0      | 0 | 0 | 2014 |
| 469 | EFG | Male   | 0.868541 | 0.476319 | 0.827406 | 0.008239 | 0.259415 | 0.205423 | 0 | 1.861759 | 1.487958 | 0      | 0      | 0 | 0 | 2014 |
| 470 | EFG | Male   | 1.3692   | 0.664343 | 1.091768 | 0.016213 | 0.509328 | 0.331405 | 0 | 0.544624 | 0.275476 | 0      | 0.1414 | 0 | 0 | 2014 |
| 515 | EFG | Female | 0.204476 | 0.065759 | 0.30566  | 0.282567 | 0.117169 | 0.15822  | 0 | 8.092956 | 0.972655 | 0      | 0      | 0 | 0 | 2015 |

|     |        |        |          |          |          |          |          |          |   |          |          |          |      |      |   |      |
|-----|--------|--------|----------|----------|----------|----------|----------|----------|---|----------|----------|----------|------|------|---|------|
| 516 | EFG    | Male   | 0.218141 | 0.069028 | 0.219658 | 0.167241 | 0.113178 | 0.158952 | 0 | 10.22007 | 0.565135 | 0        | 0    | 0    | 0 | 2015 |
| 517 | EFG    | Female | 0.430276 | 0.115824 | 0.469761 | 0.399611 | 0.263036 | 0.356013 | 0 | 9.917662 | 0.808508 | 0        | 0    | 0    | 0 | 2015 |
| 519 | EFG    | Male   | 0.392292 | 0.127038 | 0.695762 | 0.028889 | 0.249423 | 0.274841 | 0 | 3.087987 | 0.271057 | 0        | 0    | 0    | 0 | 2015 |
| 521 | EFG    | Female | 0.23871  | 0.150378 | 0.281265 | 0.030395 | 0.112396 | 0.248273 | 0 | 5.253698 | 0.943874 | 0        | 0    | 0    | 0 | 2015 |
| 522 | EFG    | Female | 0.245989 | 0.106826 | 0.383332 | 0.001508 | 0.123279 | 0.22688  | 0 | 2.292095 | 0.276752 | 0        | 7.58 | 0    | 0 | 2015 |
| 523 | EFG    | Female | 0.216134 | 0.180075 | 0.410845 | 0.001319 | 0.195242 | 0.217638 | 0 | 2.802407 | 0.579682 | 0        | 0    | 0    | 0 | 2015 |
| 524 | EFG    | Male   | 0.258219 | 0.160799 | 0.297302 | 0.001295 | 0.240371 | 0.232183 | 0 | 2.051482 | 0.448548 | 0        | 0    | 0    | 0 | 2015 |
| 580 | QCPOST | Female | 0.56188  | 0.339543 | 0.701412 | 0.02238  | 0.153007 | 0.416099 | 1 | 0.904379 | 0.125289 | 15866.67 | 0    | 1.15 | 0 | 2015 |
| 581 | QCPOST | Male   | 0.480742 | 0.590496 | 0.708742 | 0.046071 | 0.148995 | 0.529732 | 1 | 0.260015 | 0.042004 | 557000   | 0    | 0    | 0 | 2015 |
| 582 | QCPOST | Male   | 0.348686 | 0.259415 | 0.583714 | 0.021394 | 0.089622 | 0.356836 | 1 | 0.270431 | 0.094078 | 5796.667 | 0    | 0    | 0 | 2015 |
| 583 | QCPOST | Female | 0.743979 | 0.740549 | 1.594753 | 0.039646 | 0.257623 | 0.556068 | 0 | 0.726986 | 0.123279 | 30066.67 | 0    | 2.08 | 0 | 2015 |

*Supplementary Table I. Summary of the data collected for this research. Each fish is identified by a unique number (listed under the “Sample” column). For each sample the location and year are listed, in addition to other variables measured. Gender was recorded, as well as the fold change for each  $V_H$  family ( $V_H1.1$ ,  $V_H2$ ,  $V_H5.1$ ,  $V_H9$ ,  $V_H10$ , and  $V_H12$ ) and SecHCmu/MemHCmu relative to the reference sample (#349). Presence or absence of worms in the body cavity is indicated by a “1” or a “0” in the “Worms” column. For listed pathogens the number of copies found are listed under the columns for each respective pathogen. Fold Changes for  $V_H$  families and SecHCmu/MemHCmu were calculated from anterior kidney samples, while pathogen copy numbers were calculated from spleen samples. Due to limited quantities of spleen tissue pathogen detection for certain samples used anterior kidney instead of spleen; these values are marked in red on the table.*

## Bibliography

- Altschul, SF, W Gish, W Miller, EW Myers, and DJ Lipman. 1990. "Basic Local Alignment Search Tool." *Journal of Molecular Biology* 215 (3): 403–10.
- Amanna, Ian J., Nichole E. Carlson, and Mark K. Slifka. 2007. "Duration of Humoral Immunity to Common Viral and Vaccine Antigens." *New England Journal of Medicine* 357 (19): 1903–1915.
- "An Introduction to Corrplot Package." 2016. Accessed July 11. <https://cran.r-project.org/web/packages/corrplot/vignettes/corrplot-intro.html>.
- Andersson, Elisabet, and Takeshi Matsunaga. 1998. "Evolutionary Stability of the Immunoglobulin Heavy Chain Variable Region Gene Families in Teleost." *Immunogenetics* 47 (3): 272–277.
- Baker, M.R., and C.H. Vynne. 2014. "Cortisol Profiles in Sockeye Salmon: Sample Bias and Baseline Values at Migration, Maturation, Spawning, and Senescence." *Fisheries Research* 154 (June): 38–43. doi:10.1016/j.fishres.2014.01.015.
- Bernardet, J. F., and B. Kerouault. 1989. "Phenotypic and Genomic Studies of 'Cytophaga Psychrophila' isolated from Diseased Rainbow Trout (*Oncorhynchus Mykiss*) in France." *Applied and Environmental Microbiology* 55 (7): 1796–1800.
- Blanchet, S., O. Rey, and G. Loot. 2010. "Evidence for Host Variation in Parasite Tolerance in a Wild Fish Population." *Evolutionary Ecology* 24 (5): 1129–39. doi:10.1007/s10682-010-9353-x.
- Bowers, Robert M., Scott E. Lapatra, and Arun K. Dhar. 2008. "Detection and Quantitation of Infectious Pancreatic Necrosis Virus by Real-Time Reverse Transcriptase-Polymerase Chain Reaction Using Lethal and Non-Lethal Tissue Sampling." *Journal of Virological Methods* 147 (2): 226–34. doi:10.1016/j.jviromet.2007.09.003.
- Brodeur, Peter H., and Roy Riblet. 1984. "The Immunoglobulin Heavy Chain Variable Region (Igh-V) Locus in the Mouse. I. One Hundred Igh-V Genes Comprise Seven Families of Homologous Genes." *European Journal of Immunology* 14 (10): 922–30.
- Bromage, E. 2004. "Use of Staphylococcal Protein A in the Analysis of Teleost Immunoglobulin Structural Diversity." *Developmental & Comparative Immunology* 28 (7–8): 803–14. doi:10.1016/j.dci.2003.12.001.
- Bromage, Erin S., Ilsa M. Kaattari, Patty Zwollo, and Stephen L. Kaattari. 2004. "Plasmablast and Plasma Cell Production and Distribution in Trout Immune Tissues." *The Journal of Immunology* 173 (12): 7317–7323.
- Brown, Gwynne D., Ilsa M. Kaattari, and Stephen L. Kaattari. 2006. "Two New Ig VH Gene Families in *Oncorhynchus Mykiss*." *Immunogenetics* 58 (11): 933–36. doi:10.1007/s00251-006-0149-7.
- Carlson, R.E., D.P. Anderson, and J.E. Bodammer. 1993. "In Vivo Cortisol Administration Suppresses the in Vitro Primary Immune Response of Winter Flounder Lymphocytes." *Fish & Shellfish Immunology* 3 (4): 299–312.
- Castro, Rosario, Luc Jouneau, Hang-Phuong Pham, Olivier Bouchez, Véronique Giudicelli, Marie-Paule Lefranc, Edwige Quillet, et al. 2013. "Teleost Fish Mount Complex Clonal IgM and IgT Responses in Spleen upon Systemic Viral Infection." Edited by Chris Allen. *PLoS Pathogens* 9 (1): e1003098. doi:10.1371/journal.ppat.1003098.
- "Creating Standard Curves with Genomic DNA or Plasmid Templates for Use in Quantitative PCR - Quant\_pcr.pdf." 2016. Accessed March 4. [http://www6.appliedbiosystems.com/support/tutorials/pdf/quant\\_pcr.pdf](http://www6.appliedbiosystems.com/support/tutorials/pdf/quant_pcr.pdf).



- Curran, James M. 2013. *Hotelling: Hotelling's T-Squared Test and Variants*. (version 1.0-2). <http://CRAN.R-project.org/package=Hotelling>.
- Dhabhar, Firdaus S. 2014. "Effects of Stress on Immune Function: The Good, the Bad, and the Beautiful." *Immunologic Research* 58 (2–3): 193–210. doi:10.1007/s12026-014-8517-0.
- Dhabhar, Firdaus S., and Bruce S. McEwen. 1997. "Acute Stress Enhances While Chronic Stress Suppresses Cell-Mediated Immunity in Vivo: A Potential Role for Leukocyte Trafficking." *Brain, Behavior, and Immunity* 11 (4): 286–306.
- Dittman, A., and T. Quinn. 1996. "Homing in Pacific Salmon: Mechanisms and Ecological Basis." *The Journal of Experimental Biology* 199 (1): 83–91.
- Domanico, Michael, and Ruth Phillips. 1995. "Phylogenetic Analysis of Pacific Salmon (Genus *Oncorhynchus*) Based on Mitochondrial DNA Sequence Data." May 2. [http://ac.els-cdn.com/S1055790385710342/1-s2.0-S1055790385710342-main.pdf?\\_tid=62979b46-863f-11e4-8b5a-00000aab0f6b&acdnat=1418856932\\_5d875eba240ada52bf4260ac06120b8b](http://ac.els-cdn.com/S1055790385710342/1-s2.0-S1055790385710342-main.pdf?_tid=62979b46-863f-11e4-8b5a-00000aab0f6b&acdnat=1418856932_5d875eba240ada52bf4260ac06120b8b).
- Eiras, Jorge C., Helmut Segner, Thomas Wahli, and B.G. Kapoor. 2008. *Fish Diseases*. Vol. 1. 2 vols. Enfield, NH, USA: Science Publishers.
- Fenton, Andy. 2013. "Dances with Worms: The Ecological and Evolutionary Impacts of Deworming on Coinfecting Pathogens." *Parasitology* 140 (9): 1119–32. doi:10.1017/S0031182013000590.
- Flores, A.-M., J. M. Shrimpton, D. A. Patterson, J. A. Hills, S. J. Cooke, T. Yada, S. Moriyama, S. G. Hinch, and A. P. Farrell. 2012. "Physiological and Molecular Endocrine Changes in Maturing Wild Sockeye Salmon, *Oncorhynchus Nerka*, during Ocean and River Migration." *Journal of Comparative Physiology B* 182 (1): 77–90. doi:10.1007/s00360-011-0600-4.
- French, R., H. Bilton, M. Osako, and A. Hartt. 1976. "Distribution and Origin of Sockeye Salmon (*Oncorhynchus Nerka*) in Offshore Waters of the North Pacific Ocean." *International North Pacific Fisheries Commission* 34. <http://www.npafc.org/new/inpfc/INPFC%20Bulletin/Bull%20No.34/Bulletin%2034.pdf>.
- Gadan, K., I. Singh Marjara, H. Sundh, K. Sundell, and Ø. Evensen. 2012. "Slow Release Cortisol Implants Result in Impaired Innate Immune Responses and Higher Infection Prevalence Following Experimental Challenge with Infectious Pancreatic Necrosis Virus in Atlantic Salmon (*Salmo Salar*) Parr." *Fish & Shellfish Immunology* 32 (5): 637–44. doi:10.1016/j.fsi.2012.01.004.
- "ggplot2 Stripchart (Jitter) : Quick Start Guide - R Software and Data Visualization - Documentation - STHDA." 2016. Accessed April 4. <http://www.sthda.com/english/wiki/ggplot2-stripchart-jitter-quick-start-guide-r-software-and-data-visualization>.
- Hammer, Øyvind, David Harper, and Paul D. Ryan. 2001. "PAST: Paleontological Statistics Software Package for Education and Data Analysis." *Palaeontologia Electronica* 4 (1): 9.
- Hansen, John D., Eric D. Landis, and Ruth B. Phillips. 2005. "Discovery of a Unique Ig Heavy-Chain Isotype (IgT) in Rainbow Trout: Implications for a Distinctive B Cell Developmental Pathway in Teleost Fish." *Proceedings of the National Academy of Sciences of the United States of America* 102 (19): 6919–6924.
- Hansen, John D., and Agustin G. Zapata. 1998. "Lymphocyte Development in Fish and Amphibians." *Immunological Reviews* 166 (1): 199–220.
- Hartman, Wilbur L., Theodore R. Merrell, and Richard Painter. 1964. "Mass Spawning Behavior of Sockeye Salmon in Brooks River, Alaska." *Copeia* 1964 (2): 362. doi:10.2307/1441028.

- Hothorn, Torsten, Kurt Hornik, Mark A. van de Wiel, Henric Winell, and Achim Zeileis. 2016. "General Information on the Coin Package." Accessed May 15. <http://finzi.psych.upenn.edu/library/coin/html/coin-package.html>.
- Hothorn, Torsten, Kurt Hornik, Mark A. van de Wiel, Henric Winell, and Achim Zeileis. 2015. *Coin: Conditional Inference Procedures in a Permutation Test Framework* (version 1.1-2). <https://cran.r-project.org/web/packages/coin/index.html>.
- "How to Generate a Random Number in R | inside-R | A Community Site for R." 2016. Accessed April 6. <http://www.inside-r.org/howto/how-generate-random-number-r>.
- Hoyer, B F, I M Mumtaz, T Yoshida, F Hiepe, and A Radbruch. 2008. "How to Cope with Pathogenic Long-Lived Plasma Cells in Autoimmune Diseases." *Annals of the Rheumatic Diseases* 67 (Suppl 3): iii87-iii89. doi:10.1136/ard.2008.098418.
- "IHN Virus Database." 2016. Accessed March 4. <http://gis.nacse.org/ihnv/>.
- Isaksen, Trond E., Egil Karlsbakk, Oddvar Repstad, and Are Nylund. 2012. "Molecular Tools for the Detection and Identification of Ichthyobodo Spp. (Kinetoplastida), Important Fish Parasites." *Parasitology International* 61 (4): 675–83. doi:10.1016/j.parint.2012.07.006.
- Kaattari, Stephen, Erin Bromage, and Ilsa Kaattari. 2005. "Analysis of Long-Lived Plasma Cell Production and Regulation: Implications for Vaccine Design for Aquaculture." *Aquaculture* 246 (1–4): 1–9. doi:10.1016/j.aquaculture.2004.12.024.
- Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, et al. 2012. "Geneious Basic: An Integrated and Extendable Desktop Software Platform for the Organization and Analysis of Sequence Data." *Bioinformatics* 28 (12): 1647–49. doi:10.1093/bioinformatics/bts199.
- Keeling, S E, C L Brosnahan, C Johnston, R Wallis, N Gudkovs, and W L McDonald. 2013. "Development and Validation of a Real-Time PCR Assay for the Detection of *Aeromonas Salmonicida*." *Journal of Fish Diseases* 36 (5): 495–503. doi:10.1111/jfd.12014.
- Kibenge, Frederick S.B., Marcos G. Godoy, Mark Fast, Samuel Workenhe, and Molly J.T. Kibenge. 2012. "Countermeasures against Viral Diseases of Farmed Fish." *Antiviral Research* 95 (3): 257–81. doi:10.1016/j.antiviral.2012.06.003.
- Kurath, G. 2003. "Phylogeography of Infectious Haematopoietic Necrosis Virus in North America." *Journal of General Virology* 84 (4): 803–14. doi:10.1099/vir.0.18771-0.
- Kusnecov, Alexander W., and Alba Rossi-George. 2002. "Stressor-Induced Modulation of Immune Function: A Review of Acute, Chronic Effects in Animals." *Acta Neuropsychiatrica* 14 (6): 279–291.
- Lau, Matthew K. 2013. *DTK: Dunnett-Tukey-Kramer Pairwise Multiple Comparison Test Adjusted for Unequal Variances and Unequal Sample Sizes* (version 3.5). <http://CRAN.R-project.org/package=DTK>.
- Lefranc, M.-P., V. Giudicelli, C. Ginestoux, J. Jabado-Michaloud, G. Folch, F. Bellahcene, Y. Wu, et al. 2009. "IMGT(R), the International ImMunoGeneTics Information system(R)." *Nucleic Acids Research* 37 (Database): D1006–12. doi:10.1093/nar/gkn838.
- Li, Jian, Linda G. Cowden, Janice D. King, David A. Briles, Harry W. Schroeder, Alan B. Stevens, Rodney T. Perry, et al. 2007. "Effects of Chronic Stress and Interleukin-10 Gene Polymorphisms on Antibody Response to Tetanus Vaccine in Family Caregivers of Patients With Alzheimer's Disease." *Psychosomatic Medicine* 69 (6): 551–59. doi:10.1097/PSY.0b013e3180cc2c61.
- Livak, Kenneth J., and Thomas D. Schmittgen. 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.

- Ma, Cuiyan, Jianmin Ye, and Stephen L. Kaattari. 2013. "Differential Compartmentalization of Memory B Cells versus Plasma Cells in Salmonid Fish: Cellular Immune Response." *European Journal of Immunology* 43 (2): 360–70. doi:10.1002/eji.201242570.
- Marancik, David P., and Gregory D. Wiens. 2013. "A Real-Time Polymerase Chain Reaction Assay for Identification and Quantification of *Flavobacterium Psychrophilum* and Application to Disease Resistance Studies in Selectively Bred Rainbow Trout *Oncorhynchus Mykiss*." *FEMS Microbiology Letters* 339 (2): 122–29. doi:10.1111/1574-6968.12061.
- Meyers, Theodore, Tamara Burton, Collette Bentz, and Norman Starkley. 2008. *Common Diseases of Wild and Cultured Fishes in Alaska*. 2nd ed. Alaska Department of Fish and Game, Division of Sport Fish, Research and Technical Services.
- Multiple Primer Analyzer. 2016. Thermo Fisher Scientific, Inc. Accessed March 4. <https://www.thermofisher.com/us/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/multiple-primer-analyzer.html>.
- Peters, Maureen B., and Thomas F. Turner. 2008. "Genetic Variation of the Major Histocompatibility Complex (MHC Class II  $\beta$  Gene) in the Threatened Gila Trout, *Oncorhynchus Gilae Gilae*." *Conservation Genetics* 9 (2): 257–70. doi:10.1007/s10592-007-9336-9.
- Petney, Trevor N., and Ross H. Andrews. 1998. "Multiparasite Communities in Animals and Humans: Frequency, Structure and Pathogenic Significance." *International Journal for Parasitology* 28 (3): 377–93.
- "Plotting Means and Error Bars (ggplot2)." 2016. Accessed April 4. [http://www.cookbook-r.com/Graphs/Plotting\\_means\\_and\\_error\\_bars\\_\(ggplot2\)/](http://www.cookbook-r.com/Graphs/Plotting_means_and_error_bars_(ggplot2)/).
- PrimerQuest® Program (version 3.1). 2012. Coralville, USA: IDT. <http://www.idtdna.com/Scitools>.
- Purcell, Maureen K, Gael Kurath, Kyle A Garver, Russell P Herwig, and James R Winton. 2004. "Quantitative Expression Profiling of Immune Response Genes in Rainbow Trout Following Infectious Haematopoietic Necrosis Virus (IHNV) Infection or DNA Vaccination." *Fish & Shellfish Immunology* 17 (5): 447–62. doi:10.1016/j.fsi.2004.04.017.
- Quinn, Thomas P., and Katherine W. Myers. 2004. "Anadromy and the Marine Migrations of Pacific Salmon and Trout: Rounsefell Revisited." *Reviews in Fish Biology and Fisheries* 14 (4): 421–42. doi:10.1007/s11160-005-0802-5.
- "R: A Visualization of a Correlation Matrix." 2016. Accessed April 11. <http://127.0.0.1:18512/library/corrplot/html/corrplot.html>.
- R Core Team. 2015. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- "Real-Time PCR: Understanding Ct Application Note." 2016. Accessed March 6. <https://www.thermofisher.com/us/en/home/life-science/pcr/real-time-pcr/qpcr-education/pcr-understanding-ct-application-note.html>.
- "Research - Wargo Lab." 2016. Accessed June 9. <http://wmpeople.wm.edu/site/page/arwargo/research>.
- Roman, T., J. Charlemagne, T. Matsunaga, E. Andersson, E. Bengtén, L. Pilström, J. Hansen, and S. Kaattari. 1996. "Unified Nomenclature of IGHV Genes in Rainbow Trout (*Oncorhynchus Mykiss*): Definition of eleven VH Families." *Immunogenetics* 43 (5): 325–26. doi:10.1007/BF02441003.
- Rounsefell, George A. 1958. "Anadromy in North American Salmonidae." *Fishery Bulletin of the Fish and Wildlife Service* 58: 169–85.

- Salinas, Irene, Yong-An Zhang, and J. Oriol Sunyer. 2011. "Mucosal Immunoglobulins and B Cells of Teleost Fish." *Developmental & Comparative Immunology* 35 (12): 1346–65. doi:10.1016/j.dci.2011.11.009.
- Sandell, Ta, and Kc Jacobson. 2011. "Comparison and Evaluation of Renibacterium Salmoninarum Quantitative PCR Diagnostic Assays Using Field Samples of Chinook and Coho Salmon." *Diseases of Aquatic Organisms* 93 (2): 129–39. doi:10.3354/dao02289.
- Schouten, Jonathan, Terri Clister, Amber Bruce, Lidia Epp, and Patty Zwollo. 2013. "Sockeye Salmon Retain Immunoglobulin-Secreting Plasma Cells throughout Their Spawning Journey and Post-Spawning." *Developmental & Comparative Immunology* 40 (2): 202–9. doi:10.1016/j.dci.2013.01.013.
- Soto, E, K Bowles, D Fernandez, and Jp Hawke. 2010. "Development of a Real-Time PCR Assay for Identification and Quantification of the Fish Pathogen Francisella Noatunensis Subsp. Orientalis." *Diseases of Aquatic Organisms* 89 (April): 199–207. doi:10.3354/dao02204.
- Strepparava, Nicole, Thomas Wahli, Helmut Segner, and Orlando Petrini. 2014. "Detection and Quantification of Flavobacterium Psychrophilum in Water and Fish Tissue Samples by Quantitative Real Time PCR." *BMC Microbiology* 14 (1): 105. doi:10.1186/1471-2180-14-105.
- Sugahara, Kazuhiro, Erina Fujiwara-Nagata, Ayami Fukuda, and Mitsuru Eguchi. 2010. "Viable but Non-Culturable State of Bacterial Cold-Water Disease Pathogen Flavobacterium Psychrophilum at Various Temperatures." *Fish Pathology* 45 (4): 158–63.
- "Summarizing Data." 2016. Accessed April 4. [http://www.cookbook-r.com/Manipulating\\_data/Summarizing\\_data/](http://www.cookbook-r.com/Manipulating_data/Summarizing_data/).
- Traxler, G. S., J. R. Roome, M. E. Kent, and others. 1993. "Transmission of Infectious Hematopoietic Necrosis Virus in Seawater." *Diseases of Aquatic Organisms* 16: 111–111.
- Traxler, G. S., J. R. Roome, K.A. Lauda, and S LaPatra. 1997. "Appearance of Infectious Hematopoietic Necrosis Virus (IHNV) and Neutralizing Antibodies in Sockeye Salmon Onchorynchus Nerka during Their Migration and Maturation Period." *Diseases of Aquatic Organisms* 28 (1): 31–38.
- Ueda, Hiroshi. 2011. "Physiological Mechanism of Homing Migration in Pacific Salmon from Behavioral to Molecular Biological Approaches." *General and Comparative Endocrinology* 170 (2): 222–32. doi:10.1016/j.ygcen.2010.02.003.
- Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm, and S. G. Rozen. 2012. "Primer3--New Capabilities and Interfaces." *Nucleic Acids Research* 40 (15): e115–e115. doi:10.1093/nar/gks596.
- "VDJ Recombination | Laika's MedLibLog." 2014. Accessed December 17. <https://laikaspoetnik.wordpress.com/tag/vdj-recombination/>.
- Wang, W.B., A.H. Li, T. Z. Cai, and J.G. Wang. 2005. "Effects of Intraperitoneal Injection of Cortisol on Non-Specific Immune Functions of Ctenopharyngodon Idella." *Journal of Fish Biology* 67 (3): 779–93.
- Wei, Taiyun, and Viliam Simko. 2016. *Corrplot: Visualization of a Correlation Matrix* (version 0.77). <https://cran.r-project.org/web/packages/corrplot/index.html>.
- Wickham, Hadley, Winston Chang, and RStudio. 2016. *ggplot2: An Implementation of the Grammar of Graphics* (version 2.1.0). <https://cran.r-project.org/web/packages/ggplot2/index.html>.
- Williams, Alan F., and A. Neil Barclay. 1988. "The Immunoglobulin Superfamily-Domains for Cell Surface Recognition." *Annual Review of Immunology* 6 (1): 381–405.
- Williamson, Kurt E., Jamie V. Harris, Jasmin C. Green, Faraz Rahman, and Randolph M. Chambers. 2014. "Stormwater Runoff Drives Viral Community Composition Changes in

- Inland Freshwaters." *Frontiers in Microbiology* 5 (March).  
doi:10.3389/fmicb.2014.00105.
- Wilson, Melanie, Eva Bengtén, Norman W. Miller, L. William Clem, Louis Du Pasquier, and Gregory W. Warr. 1997. "A Novel Chimeric Ig Heavy Chain from a Teleost Fish Shares Similarities to IgD." *Proceedings of the National Academy of Sciences* 94 (9): 4593–4597.
- Wood, CHRIS C., DAVID W. Welch, LYSE Godbout, and JAMES Cameron. 2012. "Marine Migratory Behavior of Hatchery-Reared Anadromous and Wild Non-Anadromous Sockeye Salmon Revealed by Acoustic Tags." In *Am Fish Soc Symp*, 76:289–311.  
[https://www.researchgate.net/profile/Chris\\_Wood3/publication/235702457\\_Marine\\_Migratory\\_Behavior\\_of\\_Hatchery-Reared\\_Anadromous\\_and\\_Wild\\_Non-Anadromous\\_Sockeye\\_Salmon\\_Revealed\\_by\\_Acoustic\\_Tags/links/09e41512c1ae398970000000.pdf](https://www.researchgate.net/profile/Chris_Wood3/publication/235702457_Marine_Migratory_Behavior_of_Hatchery-Reared_Anadromous_and_Wild_Non-Anadromous_Sockeye_Salmon_Revealed_by_Acoustic_Tags/links/09e41512c1ae398970000000.pdf).
- Ye, Jian, George Coulouris, Irena Zaretskaya, Ioana Cutcutache, Steve Rozen, and Thomas L. Madden. 2012. "Primer-BLAST: A Tool to Design Target-Specific Primers for Polymerase Chain Reaction." *BMC Bioinformatics* 13 (1): 1.
- Zeileis, Achim, Mark A. Wiel, Kurt Hornik, and Torsten Hothorn. 2008. "Implementing a Class of Permutation Tests: The Coin Package." *Journal of Statistical Software* 28 (8): 1–23.
- Zhang, Yong-An, Irene Salinas, Jun Li, David Parra, Sarah Bjork, Zhen Xu, Scott E LaPatra, Jerri Bartholomew, and J Oriol Sunyer. 2010. "IgT, a Primitive Immunoglobulin Class Specialized in Mucosal Immunity." *Nature Immunology* 11 (9): 827–35.  
doi:10.1038/ni.1913.
- Zwollo, Patty. 2012. "Why Spawning Salmon Return to Their Natal Stream: The Immunological Imprinting Hypothesis." *Developmental & Comparative Immunology* 38 (1): 27–29.