Identification of single nucleotide polymorphisms associated with growth traits in rainbow trout by RNA Sequencing

by
Ashlin Powell Harris

A thesis presented to the Honors College of Middle Tennessee State University in partial fulfillment of the requirements for graduation from the University Honors College

Spring 2015

Identification of single nucleotide polymorphisms associated
with growth traits in rainbow trout by RNA Sequencing

by
Ashlin Powell Harris

APPROVED:

_____
Dr. Mohamed Salem
Assistant Professor, Biology Department

_____
Dr. Lynn Boyd
Chair, Biology Department

_____
Dr. Dennis M. Mullen
Professor, Biology Department
Honors Council Representative

_____
Dr. John Vile
Dean, University Honors College

ACKNOWLEDGMENTS

This thesis is a result of the work of many people over several years. I would like to acknowledge the following people who directly contributed to this thesis.

LAB WORK

DNA Extraction	Fatima Abdouni, Ashlin Harris, Rhett Layman, Bam Paneru
RNA Extraction	Fatima Abdouni, Ali Ali, Bam Paneru
PCR	Fatima Abdouni, Ali Ali, Bam Paneru
Quality Check	Ashlin Harris, Rhett Layman, Bam Paneru
Organization	Fatima Abdouni, Ashlin Harris, Rhett Layman, Bam Paneru
Data Analysis	Fatima Abdouni, Rafet Al-Tobasei, Bam Paneru, Mohamed Salem

PREVIOUS PUBLICATIONS

Scholars Week 2013	Ashlin Harris, Rhett Layman, Bam Paneru
Scholars Week 2014	Ashlin Harris, Rhett Layman, Bam Paneru
Scholars Week 2015	Ali Ali, Bam Paneru, Rafet Al-Tobasei
PLoS One	Fatima Abdouni, Rafet Al-Tobasei, Bam Paneru, Mohamed Salem

HONORS THESIS

Drafting	Ali Ali, Ashlin Harris
Proofing	Ali Ali, Ashlin Harris, Bam Paneru, Mohamed Salem

ABSTRACT

The rainbow trout (*Oncorhynchus mykiss*) are a species of great importance for aquaculture. The efficiency and sustainability of commercial farms depends on strains with traits such as high muscle yield. These characteristics could be improved by selective breeding based on genetic differences between subgroups of *O. mykiss*. Single nucleotide polymorphisms (SNPs) account for the majority of these variations and can be studied via whole transcriptome sequencing. The characteristics of muscle mass, fat content, and flesh whiteness were measured in about a 100 families from two generations produced at the USDA/NCCCWA rainbow trout breeding program. RNA extracted from the samples was sorted by its correlation with the traits, sequenced by RNA-Seq, and mapped to a reference transcriptome. SNPs with a very high likelihood of trait influence were confirmed by Sequenom genotyping in a small population. Seven SNPs were confirmed as genetic markers for traits of interest.

TABLE OF CONTENTS

List of Figures 6
Introduction 7
Rainbow Trout 7
Aquaculture 8
Marker-Assisted Selection 8
Genome and Transcriptome 9
Single Nucleotide Polymorphisms 10
Quality Traits 11
Materials and Methods 13
Ethics Statement 13
Tissue Samples 13
Preparation of Libraries and Sequencing 13
Isolation of DNA for Genotyping 14
SNP Marker Validation and Genotyping 15
RNA-Seq Analysis for SNP Detection 17
Results 19
Discussion 21
Voltage-Dependent Anion-Selective Channel Protein 21
Cytoplasmic FMR1-Interacting Proteins 22
Taxilin Beta Muscle-Derived 77-Like 23
Nucleolin Isoform X2 23
Eukaryotic Translation Initiation Factor 3, Subunit I 24
Conclusion 25
References 26
Appendices 31
1. DNA Extraction from Fin Clip 31
2. Scholars Week Abstract 2013 32
3. Scholars Week Abstract 2014 33
4. Scholars Week Abstract 2015 34

LIST OF FIGURES

FIGURE 1:

Workflow of SNP Identification 16

This workflow was used for identification of SNPs associated with muscle mass, fat content, and flesh whiteness in rainbow trout.

FIGURE 2:

Bar Graph of Mapping Statistics 18

This bar graph shows the mapping statistics of the reads of high and low-ranked families. Reads were mapped to the previously generated reference transcriptome. An average of 8,600,000 reads were mapped per group.

FIGURE 3:

Table of Biological Significance of SNP Genes 20

This table shows the analytical results of the association between 41 validated nuclear SNPs and three traits (percent muscle mass, percent fat content, and percent color absorbance). The analysis was completed with PLINK genetic toolset in a population of 871 fish. These seven markers were each associated with at least one trait. Low false discovery rates indicate probable associations and are shaded deep red in this table. For each polymorphism, the major and minor alleles are designated and assigned color for contrast: yellow for adenine, green for guanine, blue for cytosine, and red for thymine. Each minor allele frequency [MAF] is also shown, with darker shading corresponding to greater frequency. Genotyping data for each SNP is given (number of individuals homozygous for major allele / number of heterozygous individuals / number of individuals homozygous for minor allele).

INTRODUCTION

RAINBOW TROUT

The rainbow trout (*Oncorhynchus mykiss*) are an iconic species which have greatly impacted humans. Native to the eastern coast of Russia and the western coast of North America, this valuable food animal was transported eastward and has been introduced to all continents besides Antarctica. They have become the most widely cultivated cold freshwater fish in the world (Palti *et al*., 2011).

Traits such as large size and adaptability to artificial rearing and large size make rainbow trout ideal for such various purposes as food production, sport fishing, and research (Ma *et al*., 2012, Pulcini *et al*., 2012, Thorgaard *et al*., 2002). They are larger than most other model fish species such as zebrafish and pufferfish, so greater quantities of specific organ tissue can be obtained from an individual (Ali *et al*., 2014).

O. mykiss have been used as a model organism for such fields of research as ecology (Davidson, 2012), evolutionary biology (Taylor *et al*., 2011), immunology (Nya and Austin, 2011), nutrition (Wong *et al*., 2013), oncology (Tilton *et al*., 2005), toxicology (Köllner *et al*., 2002), and physiology (Palti *et al*., 2011). More is known about the physiology and biology of rainbow trout than any other fish species (Salem *et al*., 2015).

AQUACULTURE

Aquaculture remains the most sustainable solution to rising global demand for seafood products. Aquaculture is the fastest growing animal production sector and has expanded to accommodate demand that fisheries alone cannot meet (Lazzarotto *et al*., 2015). The quota for aquaculture production worldwide increased from 38.5% to 48% from 2005 to 2011. Currently, aquaculture is experiencing a growth rate of 8%, the most for any global food production industry (Safari and Mehraban Sang Atash, 2015).

The success of rainbow trout aquaculture is due to innovations in management, better understanding of nutrition, advanced disease control, and improved strain genetics. Fast and efficient growth highly influence the profitability of food animal production. Through improvements in feed quality and uptake, growth rate and feed conversion ratio must be maximized and mortality must be minimized in order for aquaculture to remain profitable (Ramzani *et al*., 2014). Furthermore, weight gain can also influenced by pro nutrients that may affect utilization of nutrients (Ramzani *et al*., 2014). There is also interest in using terrestrial plant resources to replace or supplement aquatically sourced components in this salmonid's carnivorous diet, both to reduce cost and to minimize dependence on marine fishery resources (Lazzarotto *et al*., 2015).

MARKER-ASSISTED SELECTION

Modifications to growth characteristics of commercial strains may help trout farmers bolster their production. While growth rate and adult size of *O. mykiss* can be increased by selection based on adult phenotype, this process leads to slow, modest gains

over several generations (Fishback, 2002, Salem *et al*., 2015). The efficiency of fish farms cannot improve, through traditional means alone, by the rate at which need increases (NCCCWA, 2005). Muscle traits are complex, polygenic and difficult to improve by conventional selection. As an additional impediment to traditional selection, some traits such as fatty acid content cannot easily be determined in living individuals; for such traits, once an individual's desirability for breeding is ascertained, it is no longer able to breed.

The information obtained by an analysis of genetic differences can lead to quicker, more cost-effective improvement in commercial populations. If a characteristic in an individual's genome serves as a marker for improved adult phenotype, the individual can be chosen for breeding at any stage of development. Genetic marker assisted selection for desired phenotypes can be accomplished in fewer generations with better results than traditional selection, allowing for quicker improvement in commercial populations (Salem *et al*., 2015).

Genome and Transcriptome

Genetic research of *O. mykiss* is complicated by the size and complexity of its genome. An ancestor of the ray-finned fishes underwent an autotetraploid event, which effectively doubled its genome. An additional autotetraploid event occurred in the ancestor of the salmonids (Genet *et al*., 2011). Due to crossovers between homologous chromosomes, modern rainbow trout still possess duplicate copies of most genes (Sakamoto and Danzmann, 2000, Palti *et al*., 2011, Christensen *et al*., 2013). In addition,

over half of all rainbow trout genes contains interspersed repeating sequences (Genet *et al*., 2011; Salem *et al*., 2015). The presence of many similar sequences hinders the development of a standardized genome reference sequence.

Alternatively, genes transcribed from different regions of the genome can be deduced from a transcriptome, the collection of RNA molecules synthesized in a population of cells. The muscle cell transcriptome of rainbow trout can reveal which variants in the DNA are correlated with important aquacultural traits by comparison with phenotypic data from adult individuals.

Single Nucleotide Polymorphisms

A single nucleotide polymorphism (SNP) is a variation in one particular base of the DNA in an individual in comparison either to a paired chromosome of the same individual or to a homologous chromosome in another individual from the same population or species. Though any of the 4 nucleotides may be present in a position, SNPs are typically bi-allelic. SNPs by common definition must be significant in a population, with the minor variation having a frequency of at least 1% (Vignal *et al*., 2002). In rainbow trout, SNPs can potentially exist as variation between alleles in a single chromosome or between two paralogous chromosomes (Christensen *et al*., 2013).

SNPs are a significant portion of the information that can be determined from comparison of individual genomes; potentially, hundreds of thousands of SNPs can be found in a population (Platzer 2013). Because they represent a large proportion of total genetic differences (as much as 90% among humans), they are ideal for genetic-assisted

selection of desired fish muscle traits (Pertea 2012). SNPs are also ubiquitous and readily scored, providing a relatively simple and inexpensive means to generate a vast amount of genotyping data (Vignal et al., 2002). SNPs which are proximal to or located within coding sequence are more likely to alter a completed protein and are the primary interest of this study.

Quality Traits

Traits that are beneficial for aquaculture include fast and efficient growth of individuals as well as pleasing appearance of fillets. The traits of percent muscle mass, percent fat mass, and flesh whiteness were selected for an operational definition of fish quality in this study.

As stated above, growth traits such as muscle yield and fat content are highly complex, as the adult size of an individual is determined by many genes, overall nutrition, and exposure and response to antigens. Diverse genes with many functions may potentially affect the adult phenotype.

Color is perceived as an indicator of overall quality and is a highly influential aspect of consumer approval (Safari and Mehraban Sang Atash, 2015). The characteristic pinkish coloration of rainbow trout fillets depends primarily on carotenoid pigments supplied in feed, as salmonids are unable to synthesize these pigments *de novo*. If carotenoids are not supplied in feed, the fillets will be a darker red (Hardy and Barrows, 2002). Carotenoids are absorbed enzymatically as lipophilic pigments and transmitted to the circulatory systems via carotenoid-carrying lipoprotein (Storebakken and No, 1992).

Carotenoids are deposited into the fillets, and only a negligible portion is remobilized in the blood of adult fish (Guillou and Choubert, 1992). However, color loss can occur in processed fillets during storage (Safari and Mehraban Sang Atash, 2015). Any gene that contributes to absorption, circulation, and retention of carotenoids could potentially affect this trait. Flesh whiteness, without the influence of dietary carotenoids, is heritable and was selected as a trait of interest for this study.

MATERIALS AND METHODS

ETHICS STATEMENT

Institutional Animal Care and Use Committee of The United States Department of Agriculture, National Center for Cool and Cold Water Aquaculture (Leetown, WV) specifically reviewed and approved all husbandry practices used in this study (IACUC approval #056).

TISSUE SAMPLES

Our samples were obtained from rainbow trout that were raised in Leetown, West Virginia at a USDA facility, the National Center for Cool and Cold Water Aquaculture. Extensive phenotypic data, including several size measurements and muscle and organ masses, were recorded for approximately 500 fish. Rainbow trout with phenotypic variations in traits were selected for this experiment from a hundred families and 2 generations (hatch years 2010 and 2012). Tissue samples were shipped on ice to MTSU and stored at -80°C. TRIzol® from Invitrogen (Carlsbad, CA) was utilized for total RNA isolation from each sample.

PREPARATION OF LIBRARIES AND SEQUENCING

Preparation of libraries and sequencing were performed at the University of Illinois. The TruSeq RNA Preparation Kit from Illumina (San Diego, CA) was used for construction of RNA-seq libraries. High-quality RNA (1 µg) was used as the source of mRNA. Strands were synthesized and 3'-end A-tailed double-stranded DNA was ligated

to indexed adaptors and amplified by PCR for 10 cycles. Libraries were Qubit-quantitated and pooled, followed by qPCR quantification to produce the maximum number of clusters in the flow cell. The libraries were sequenced from one end of the molecules on an Illumina HiSeq2000 to give reads with a length of 100 nucleotides.

Reads were mapped to the reference transcriptome assembled from 13 tissues of a single double-haploid fish using the default settings. Putative SNPs were called using the default parameters of CLC. Very strict parameters for SNPs detection were used (depth of coverage > 10; $P > 97\%$, quality score > 30) and followed by Fisher's exact test ($P < 0.05$).

Isolation of DNA for Genotyping

To isolate DNA, 500 µL extraction buffer, 5 µL DTT (1M) and 6 µL Pro-K (10 mg/mL) were added to each fin clip in a 1.5 mL microtube. The tubes were stored at 37°C overnight on a rotary shaker, then spun for 2 min at 14,000rpm to separate the pigments. The supernatant was transferred to a new 1.5mL microtube and mixed with 200 µL 5M NaCl. These tubes were spun for 10 min at 14,000 rpm, and the supernatant was transferred to a new 2 mL screw-cap tube containing cold 100% ethanol (1 mL). The tubes were cooled on ice for 10 min or in -20°C freezer then spun for 10 min at 14,000 rpm and 4°C. The resulting pellets were aspirated. For a total of 2 times, the pellet was rinsed in 1 mL of 70% EtOH, spun for 5 min at 14,000 rpm and 4°C, and aspirated. Afterwards, the pellets were allowed to air dry for 20 min. Finally, 100 µL Low TE

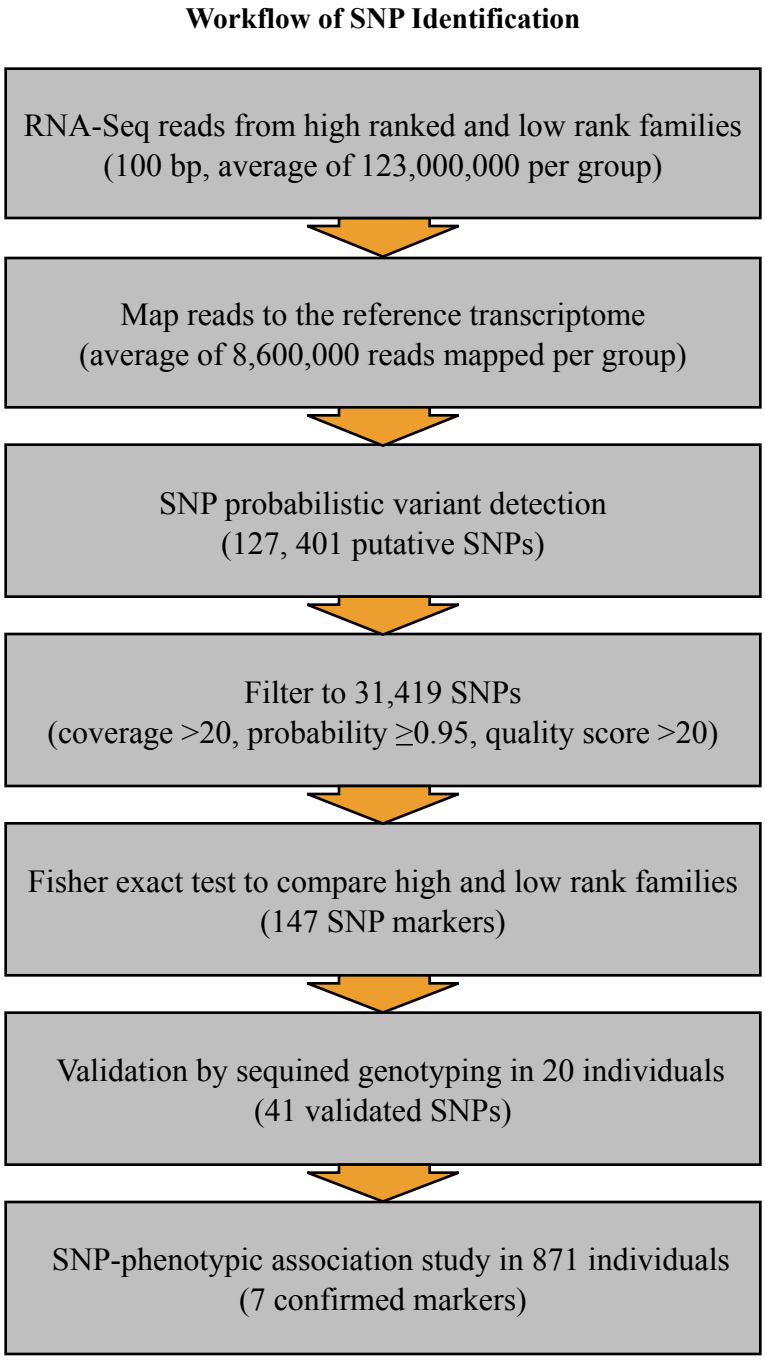
(10mM Tris, pH8.0; 0.1mM EDTA) was added, and each tube was left overnight at 37°C. DNA extracts were analyzed for concentration and purity by spectrometer.

SNP Marker Validation and Genotyping

Samples were sent to the University of Arizona Genetics Core for genotyping to validate SNPs. We expected the SNPs in the resulting genome to match many of the ones we had developed computationally from data obtained by RNA-Seq.

SNPs markers identified by RNA-Seq were validated by Sequenom genotyping at University of Arizona Genetics Core. Associations of markers with muscle mass, fat content, and flesh whiteness were assessed by genotyping 871 additional fish. SNPs associated with each phenotype were determined using PLINK, a whole genome association toolset. The detailed workflow for our pipeline to identify SNPs associated with muscle yield, fat content, and color is shown in Figure 1.

FIGURE 1



This workflow was used for identification of SNPs associated with muscle mass, fat content, and flesh whiteness in rainbow trout.

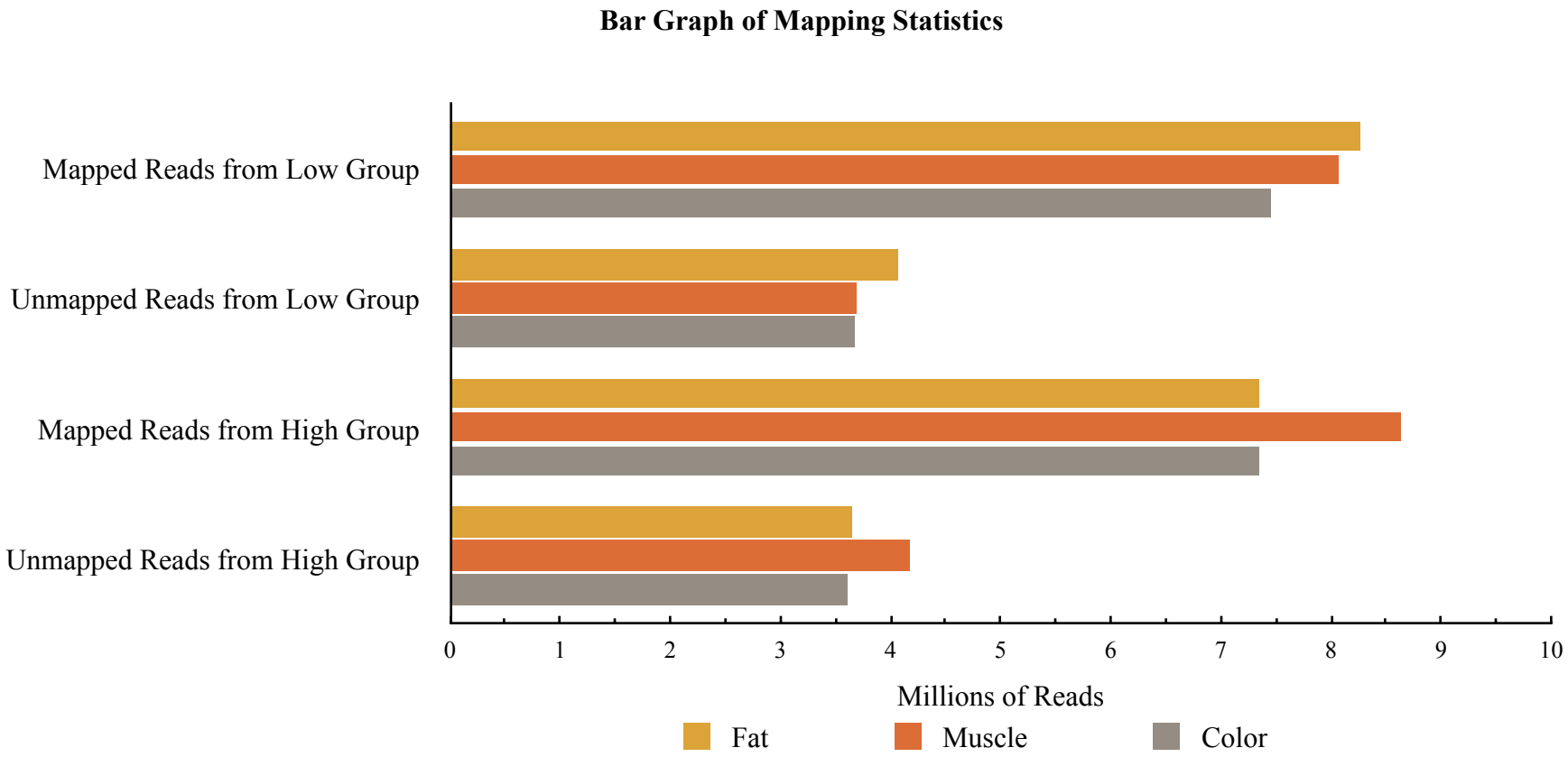
RNA-SEQ ANALYSIS FOR SNP DETECTION

SNPs discovery was achieved using tissues collected from high-ranked families versus low-ranked families. For muscle mass and fat content, cDNA libraries from 8 different full-sibling families of high fat content (5 families) versus low fat content (3 families) and high muscle yield (4 families) versus low muscle yield (4 families) were used for RNA-Seq, with two families shared between muscle yield and fat content. For flesh whiteness, cDNA libraries from 9 different full-sib families of high color (5 families) versus low color (4 families) were used for RNA-Seq, with two families shared between fat content and flesh whiteness. Bar-coded libraries were prepared using single-end multiplexed sequencing protocol. A total of 21 RNA-Seq libraries were produced for all traits. Libraries were pooled and sequenced on a single Illumina HiSeq2000 flow cell.

In this study, important putative SNPs in cDNAs from high versus low ranked families were identified using the reads from RNA-Seq high rank and low rank families (each read 100 base pairs in length, an average of 123,000,000 reads per group). Reads were mapped to the previously generated reference transcriptome (Salem *et al*., 2015). An average of 8,600,000 reads were mapped per group. Mapping statistics for the traits of muscle yield, fat content, and color are shown in Figure 2.

Putative SNPs identified by RNA-Seq were validated and genotyped by Sequenom genotyping platform. Associations of markers with muscle mass were further estimated by genotyping 871 individuals of the selected line at NCCCWA (hatch years 2010 and 2012).

FIGURE 2



This bar graph shows the mapping statistics of the reads of high and low-ranked families. Reads were mapped to the previously generated reference transcriptome. An average of 8,600,000 reads were mapped per group.

RESULTS

Probabilistic variant detection yielded 127,401 putative SNPs. Strict SNP filtration parameters (coverage > 20 reads; 4 reads to call variant; $P \geq 0.95$; forward/reverse balance < 0.05; quality score > 30) resulted with 31,419 putative SNPs. Fisher exact test on high rank and low rank families yielded 147 SNP ($P < 0.1$). These were submitted for genotyping and pared down to 41 validated SNPs considered for association study with phenotypes.

We measured muscle yield for individuals in 100 families from the USDA rainbow trout breeding program, and variations were compared with SNP allele frequencies. Out of 41 validated SNPs, the study identified 7 SNPs as potential genetic markers associated with quality traits. These results are given in Figure 3 and examined in the discussion.

FIGURE 3

Table of Biological Significance of SNP Genes

Gene with SNP	False Discovery Rate			Single Nucleotide Polymorphism			Genotyping Data
	Muscle Percentage	Fat Percentage	Color	Major Allele	Minor Allele	MAF	
VDAC2	4.39×10^{-5}	3.00×10^{-3}	3.00×10^{-3}	G	T	0.19	32/266/573
CYFIP1	0.008	0.776	0.775	A	C	0.22	30/315/526
CYFIP2	0.008	1	1	T	C	0.10	3/167/700
Taxilin beta muscle derived 77 like	0.038	0.46	0.46	A	G	0.31	78/388/405
EIF3I	0.16	0.008	0.14	A	G	0.27	66/325/459
EIF3I	0.18	0.016	0.016	A	G	0.30	87/350/425

This table shows the analytical results of the association between 41 validated nuclear SNPs and three traits (percent muscle mass, percent fat content, and percent color absorbance). The analysis was completed with PLINK genetic toolset in a population of 871 fish. These seven markers were each associated with at least one trait. Low false discovery rates indicate probable associations and are shaded deep red in this table. For each polymorphism, the major and minor alleles are designated and assigned color for contrast: yellow for adenine, green for guanine, blue for cytosine, and red for thymine. Each minor allele frequency [MAF] is also shown, with darker shading corresponding to greater frequency. Genotyping data for each SNP is given (number of individuals homozygous for major allele / number of heterozygous individuals / number of individuals homozygous for minor allele).

DISCUSSION

Each of the 7 SNPs identified in this study were synonymous mutations, meaning that each single base substitution does not directly alter the transcribed protein. Most of the SNPs were associated with housekeeping genes or genes with significant immune function. Housekeeping genes are always expressed and contribute to basic cell function, whereas tissue-specific genes are differentially expressed and contribute to cell specificity (Salem *et al*., 2015).

VOLTAGE-DEPENDENT ANION-SELECTIVE CHANNEL PROTEIN 2

A guanine-to-thymine SNP with a frequency of 0.19 was correlated with muscle mass, fat content, and flesh whiteness. This SNP is associated with a gene for voltage-dependent anion-selective channel protein 2 (VDAC2).

VDAC2 is a protein found abundantly in the outer membrane of mitochondria as a constituent of mitochondrial porins. These channels mediate a significant portion of all molecular traffic to and from the mitochondria and are involved in regulation of mitochondrial apoptosis (Hoogenboom *et al*., 2007). This gene could be a potential marker for immune response in rainbow trout (Long *et al*., 2015).

Since the gene for VDAC2 is important for cell housekeeping and immune response, its associated SNP could conceivably have implications for many diverse traits.

Cytoplasmic FMR1-Interacting Proteins

Two SNPs were found associated genes with cytoplasmic FMR1-interacting proteins (CYFIP). An adenine-to-cytosine SNP with a frequency of 0.22 was correlated with muscle mass. This SNP is associated with a gene for CYFIP1. A thymine-to-cytosine SNP with a frequency of 0.10 was also correlated with muscle mass. This SNP is associated with a gene for CYFIP2.

CYFIP are very highly conserved proteins. CYFIP1 and CYFIP2 in humans are, respectively, 92% and 98% identical to the homologous genes in zebrafish. This suggests structural of functional importance for these genes, though neither protein contains a recognized functional domain or motif. More likely, these proteins function by binding to other proteins (Mayne *et al*., 2004).

CYFIP is broadly expressed in the CNS (Schenck et al., 2001), where it has implications for axon pathfinding and retinal lamination (Kobayashi et al., 1998). It is also associated with both fibronectin-mediated primary CD4+ cell adhesion and Rac-1-mediated T cell adhesion (Mayne *et al*., 2004). CYFIP2 is highly abundant in CD4+ cells in patients with multiple sclerosis, an autoimmune disorder (Pittman *et al*., 2010).

Since CYFIP has an immune component, it is plausible that these 2 associated SNPs may affect immune response to parasites and other diseases, which in turn affect adult weight.

Taxilin Beta Muscle-Derived 77-Like

An adenine-to-guanine SNP with a frequency of 0.31 was correlated with muscle mass. This SNP is associated with a gene for taxilin btw muscle-derived 77-like.

The taxilin family interacts with syntaxin proteins in the process of intracellular vesicle traffic, especially calcium-dependent exocytosis. They are expressed ubiquitously as housekeeping proteins (Nogami *et al*., 2003; Nogami *et al*., 2004). They are also involved with recycling transferrin receptors which have been endocytosed. (Sakane *et al*., 2014).

Since this gene for a taxilin protein contributes to exocytosis and iron uptake, it is possible that an associated SNP could have implications for muscle development.

Nucleolin Isoform X2

A thymine-to-cytosine SNP with a frequency of 0.50 was correlated with fat content and flesh whiteness. This SNP is associated with a gene for nucleoli isoform X2.

Nucleolin is a eukaryotic nucleolar phosphoprotein, and the most abundant nucleolar protein, excluding ribosomal proteins (Lapeyre *et al*., 1997). It influences synthesis and maturation of ribosomes, from transcription of rDNA to assembly of ribosomal components (Ginisty *et al*., 2000). Primarily located in dense fibrillar regions of the nucleolus, it is essential for general cell activity but is expressed in greater quantities in metabolically active cells (Bayne *et al*., 2006).

For rainbow trout in particular, nucleolin production increases in cold seasons when fish are less metabolically active, serving to down-regulate production of both rDNA (Roger *et al*., 2002) and rRNA (Alvarez *et al*., 2003).

Since this gene for a nucleoli protein is involved in metabolic regulation, it appears likely that this associated SNP could affect fat storage in an adult rainbow trout. The association with color could be due to the altered composition of muscle cells as a result of metabolic down-regulation.

Eukaryotic Translation Initiation Factor 3, Subunit I

Two SNPs were found associated genes with Eukaryotic translation initiation factor 3, subunit I (EIF3I). An adenine-to-guanine SNP with a frequency of 0.27 was correlated with fat content. A adenine-to-cytosine SNP with a frequency of 0.30 was correlated with fat content and flesh whiteness.

EIF3I is one of 13 subunits of eukaryotic initiation factor (EIF) 3 (Graf *et al*., 2014). EIFs support the initiation of translation in eukaryotes (Rezende *et al*., 2014). EIF subunits may be overproduced in a carcinoma cell, resulting in larger size, increased proliferation, and enhanced cyclic progression (Cuchalová *et al*., 2010).

Since EIFs have implications for many cellular processes, it is conceivable that this associated SNP could affect fat content and flesh whiteness.

CONCLUSION

RNA Sequencing was used in this current study to identify single nucleotide polymorphisms (SNPs) to serve as markers associated with the traits of muscle mass, fat content, and flesh whiteness in rainbow trout. We validated many of these presumptive SNPs. We attempted to identify candidate genes that could be responsible for phenotypic variations. Additionally, we investigated why each SNP may elicit its effect and identified genes that could be responsible for variations occurring among phenotypes.

This study included SNP identification in only the muscle tissue of adult rainbow trout. In future studies, we intend to examine other tissues and life stages and tissues for SNPs associated with desirable traits for aquaculture in order to improve the productivity and yield of commercially farmed strains.

REFERENCES

Ali A, Rexroad C, Thorgaard G, Jianbo Y, Salem M. Characterization of the rainbow trout spleen transcriptome and identification of immune-related genes. Frontiers In Genetics [serial online]. October 2014;5:1-17.

Alvarez M, Quezada C, Navarro C, Molina A, Bouvet P, Krauskopf M, Vera M.I, An increased expression of nucleolin is associated with a physiological nucleolar segregation, Biochemical and Biophysical Research Communications, Volume 301, Issue 1, 31 January 2003, Pages 152-158, ISSN 0006-291X, http://dx.doi.org/10.1016/S0006-291X(02)02978-9.

Andreassen R., Lunner S., Hoyheim B. Characterization of full-length sequenced cDNA inserts (FLIcs) from Atlantic salmon (Salmo salar). BMC Genomics 10:502-502(2009).

Bayne, Christopher J., Lena Gerwick, Paul A. Wheeler, Gary H. Thorgaard, Transcriptome profiles of livers and kidneys from three rainbow trout (Oncorhynchus mykiss) clonal lines distinguish stocks from three allopatric populations, Comparative Biochemistry and Physiology Part D: Genomics and Proteomics, Volume 1, Issue 4, December 2006, Pages 396-403, ISSN 1744-117X, http://dx.doi.org/10.1016/j.cbd.2006.10.001.

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. The Annals of Statistics, 29(4), 1165–1188. Retrieved from http://projecteuclid.org/euclid.aos/1013699998.

Bullard, J. H., Purdom, E., Hansen, K. D., & Dudoit, S. (2010). Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. BMC Bioinformatics, 11, 94. http://doi.org/10.1186/1471-2105-11-94.

Christensen K, Brunelli J, Lambert M, DeKoning J, Phillips R, Thorgaard G. Identification of single nucleotide polymorphisms from the transcriptome of an organism with a whole genome duplication. BMC Bioinformatics [serial online]. December 2013;14(1):1-23.

Cuchalová, L., Kouba, T., Herrmannová, A., Dányi, I., Chiu, W.-L., & Valásek, L. (2010). The RNA recognition motif of eukaryotic translation initiation factor 3g (eIF3g) is required for resumption of scanning of posttermination ribosomes for reinitiation on GCN4 and together with eIF3i stimulates linear scanning. Molecular and Cellular Biology, 30(19), 4671–86. http://doi.org/10.1128/MCB.00430-10.

Davidson WS (2012) Adaptation genomics: next generation sequencing reveals a shared haplotype for rapid early development in geographically and genetically distant populations of rainbow trout. Mol Ecol 21: 219–222. pmid: 22329016 doi: 10.1111/j.1365-294x.2011.05387.x

Dekamin M, Veisi H, Safari E, Liaghati H, Khoshbakht K, Dekamin M. Life cycle assessment for rainbow trout (Oncorhynchus mykiss) production systems: a case study for Iran. Journal Of Cleaner Production [serial online]. March 15, 2015;91:43-55.

Edwards, J. S., Ibarra, R. U., & Palsson, B. O. (2001). In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. Nature Biotechnology, 19(2), 125–30. http://doi.org/10.1038/84379.

Ellington, A., & Szostak, J. (1990). In vitro selection of RNA molecules that bind specific ligands. Nature. Retrieved from http://www.researchgate.net/publication/21459179_In_vitro_selection_of_RNA_molecules_that_bind_specific_ligands.

Eya, J. C., Yossa, R., Ashame, M. F., Pomeroy, C. F. and Gannam, A. L. (2014), Effects of dietary lipid levels on mitochondrial gene expression in low and high-feed efficient families of rainbow trout Oncorhynchus mykiss. Journal of Fish Biology, 84: 1708–1720. doi: 10.1111/jfb.12391.

Fishback, A. G., Danzmann, R. G., Ferguson, M. M., & Gibson, J. P. (2002). Estimates of genetic parameters and genotype by environment interactions for growth traits of rainbow trout (Oncorhynchus mykiss) as inferred using molecular pedigrees. Aquaculture, 206(3-4), 137–150. doi: 10.1016/S0044-8486(01)00707-4.

General Guidelines for Working with RNA. (n.d.). From http://abe.leeward.hawaii.edu/Protocols/General Guidelines for Working with RNA.htm.

Genet C, Dehais P, Boussaha M, et al. Analysis of BAC-end sequences in rainbow trout: Content characterization and assessment of synteny between trout and other fish genomes. BMC Genomics [serial online]. January 2011;12(1): 314-321.

Ginisty, H. G. Serin, L. Ghisolfi-Nieto, B. Roger, V. Libante, F. Amalric, P. Bouvet. Interaction of nucleolin with an evolutionarily conserved pre-ribosomal RNA sequence is required for the assembly of the primary processing complex. J. Biol. Chem., 275 (2000), pp. 18845–18850

Graf, Alexander, Stefan Krebs, Mari Heininen-Brown, Valeri Zakhartchenko, Helmut Blum, Eckhard Wolf, Genome activation in bovine embryos: Review of the literature and new insights from RNA sequencing experiments, Animal Reproduction Science, Volume 149, Issues 1–2, September 2014, Pages 46-58, ISSN 0378-4320, http://dx.doi.org/10.1016/j.anireprosci.2014.05.016.

Grishin, N. V. (2001). Fold change in evolution of protein structures. Journal of Structural Biology, 134(2-3), 167–85. http://doi.org/10.1006/jsbi.2001.4335.

Guillou, A., G. Choubert. Absorption and blood clearance of labelled carotenoids ([14C] astaxanthin, [3H] canthaxanthin and [3H] zeaxanthin) in mature female rainbow trout (Oncorhynchus mykiss). Comp. Biochem. Physiol., 103A (1992), pp. 301–306.

Han M, Zheng Y. Comprehensive Analysis of Single Nucleotide Polymorphisms in Human MicroRNAs. Plos ONE [serial online]. November 2013;8(11):1-13.

Hand B, Hether T, Luikhart G et al. Genomics and introgression: Discovery and mapping of thousands of species-diagnostic SNPs using RAD sequencing. Current Zoology [serial online]. January 2015;61(1): 146-154.

Hardy, R.W., F. Barrows. Diet formulation and manufacture. J.E. Halver, R.W. Hardy (Eds.), Fish Nutrition, Academic Press (2002), pp. 505–600.

Hoogenboom, Bart W., Kitaru Suda, Andreas Engel, Dimitrios Fotiadis, The Supramolecular Assemblies of Voltage-dependent Anion Channels in the Native Membrane, Journal of Molecular Biology, Volume 370, Issue 2, 6 July 2007, Pages 246-255, ISSN 0022-2836, http://dx.doi.org/10.1016/j.jmb.2007.04.073. (http://www.sciencedirect.com/science/article/pii/S002228360700602X)

Identification and characterization of genes affecting cool and cold water aquaculture production ... annual report. (2005). Washington, D.C.: U.S. Dept. of Agriculture, Agricultural Research Service. Retrieved from http://www.lib.muohio.edu/multifacet/record/mu3ugb3993054.

Itoh, S., Fujimori, K. E., Uyeda, A., Matsuda, A., Kobayashi, H., Shinomiya, K., Tanaka, J. and Taguchi, T. (2005), Long-term effects of muscle-derived protein with molecular mass of 77 kDa (MDP77) on nerve regeneration. J. Neurosci. Res., 81: 730–738. doi: 10.1002/jnr.20582.

Kirchner S, McDaniel N, Ferraris R, et al. Salmonid microarrays identify intestinal genes that reliably monitor P deficiency in rainbow trout aquaculture. Animal Genetics [serial online]. August 2007;38(4):319-331.

Kobayashi, K. S. Kuroda, M. Fukata, T. Nakamura, T. Nagase, N. Nomura, Y. Matsuura, N. Yoshida-Kubomura, A. Iwamatsu, K. Kaibuchi. p140Sra-1 (specifically Rac1-associated protein) is a novel specific target for Rac1 small GTPase. J. Biol. Chem., 273 (1998), pp. 291–295

Kolditz CI. et al. BMC genomics: Changes induced by dietary energy intake and divergent selection for muscle fat content in rainbow trout (Oncorhynchus mykiss), assessed by transcriptome and proteome analysis of the liver. BMC Genomics [serial online]. January 3, 2008;9:506-521.

Köllner B, Wasserrab B, Kotterba G, Fischer U. Evaluation of immune functions of rainbow trout (Oncorhynchus mykiss)—how can environmental influences be detected?. Toxicology Letters [serial online]. May 10, 2002;131(1/2):83.

Lapeyre, B. H. Bourbon, F. Amalric. Nucleolin, the major nucleolar protein of growing eukaryotic cells: an unusual protein structure revealed by the nucleotide sequence. Proc. Natl. Acad. Sci. USA, 84 (1987), pp. 1472–1476.

Lazzarotto V, Corraze G, Leprevost A, Quillet E, Dupont-Nivet M, Médale F. Three-Year Breeding Cycle of Rainbow Trout (Oncorhynchus mykiss) Fed a Plant-Based Diet, Totally Free of Marine Resources: Consequences for Reproduction, Fatty Acid Composition and Progeny Survival. Plos ONE [serial online]. February 2015;10(2):1-17.

Librán-Pérez M, Velasco C, López-Patiño M, Míguez J, Soengas J. Counter-Regulatory Response to a Fall in Circulating Fatty Acid Levels in Rainbow Trout. Possible Involvement of the Hypothalamus-Pituitary-Interrenal Axis. Plos ONE [serial online]. November 2014;9(11):1-10.

Ma H, Hostuttler M, Hairong W, Rexroad I, Jianbo Y. Characterization of the Rainbow Trout Egg MicroRNA Transcriptome. Plos ONE [serial online]. June 2012;7(6):1-8.

Mayne, M., Moffatt, T., Kong, H., McLaren, Paul J., Fowke, Keith R., Becker, Kevin G., Namaka, M., Schenck, A., Bardoni, B., Bernstein, Charles N. and Melanson, M. (2004), CYFIP2 is highly abundant in CD4+ cells from multiple sclerosis patients and is involved in T cell adhesion. Eur. J. Immunol., 34: 1217–1227. doi: 10.1002/eji.200324726.

Meng Long, Juan Zhao, Tongtong Li, Carolina Tafalla, Qianqian Zhang, Xiehao Wang, Xiaoning Gong, Zhixin Shen, Aihua Li, Transcriptomic and proteomic analyses of splenic immune mechanisms of rainbow trout (Oncorhynchus mykiss) infected by Aeromonas salmonicida subsp. salmonicida, Journal of Proteomics, Volume 122, 3 June 2015, Pages 41-54, ISSN 1874-3919, http://dx.doi.org/10.1016/j.jprot.2015.03.031.

NCCCWA. Identification and characterization of genes affecting cool and cold water aquaculture production. (2005). Washington, D.C.: U.S. Dept. of Agriculture.

Nogami S, Satoh S, Shirataki H, et al. Taxilin; a novel syntaxin-binding protein that is involved in Ca2+ -dependent exocytosis in neuroendocrine cells. Genes To Cells [serial online]. January 2003;8(1):17-28.

Nogami S, Satoh S, Shirataki H, et al. Identification and characterization of taxilin isoforms. Biochemical & Biophysical Research Communications [serial online]. July 2, 2004;319(3): 936-943.

Nya E. J., Austin B. (2011). Development of immunity in rainbow trout (Oncorhynchus mykiss, walbaum) to aeromonas hydrophila after the dietary application of garlic. Fish Shellfish Immunol. 30, 845–850 10.1016/j.fsi.2011.01.008.

Palti, Y., Genet, C., Luo, M.-C., Charlet, A., Gao, G., Hu, Y., … Rexroad, C. E. (2011). A first generation integrated map of the rainbow trout genome. BMC Genomics, 12, 180. http://doi.org/10.1186/1471-2164-12-180.

Pertea, M. (2012). The human transcriptome: an unfinished story. Genes, 3(3), 344–60. http://doi.org/10.3390/genes3030344.

Pittman, Andrew J., John A. Gaynes, Chi-Bin Chien, nev (cyfip2) is required for retinal lamination and axon guidance in the zebrafish retinotectal system, Developmental Biology, Volume 344, Issue 2, 15 August 2010, Pages 784-794, ISSN 0012-1606, http://dx.doi.org/10.1016/j.ydbio.2010.05.512.

Platzer A. Visualization of SNPs with t-SNE. Plos ONE [serial online]. February 2013;8(2):1-6.

Pulcini D. Tommaso Russo1, Paola Reale, Alexia Massa-Gallucci, Grace Brennan Stefano Cataudella. Aquaculture research: Rainbow trout (Oncorhynchus mykiss, Walbaum 1792) develop a more robust body shape under organic rearing. Blackwell Publishing; 02/2014;45:397.

Ramzani S, Soltani M, Gholipourkanani H. Influence of Dietary Dinamune on Growth Performance and Lysozyme Activity in Rainbow Trout (Oncorhynchus mykiss) Fry. Walailak Journal Of Science & Technology [serial online]. December 20, 2014;11(9-12):865-869.

Rezende A, Assis L, de Melo Neto O, et al. The translation initiation complex eIF3 in trypanosomatids and other pathogenic excavates - identification of conserved and divergent features based on orthologue analysis. BMC Genomics [serial online]. December 15, 2014;15(1):3423-3458.

Roger, B. A. Moisand, F. Amalric, P. Bouvet. Repression of RNA polymerase I transcription by nucleolin is independent of the RNA sequence that is transcribed. J. Biol. Chem., 277 (2002), pp. 10209–10219

Safari O, Mehraban Sang Atash M. The effects of dietary supplement of annatto (Bixa orellana) seed meal on blood carotenoid content and fillet color stability in rainbow trout (Oncorhynchus mykiss). Aquaculture [serial online]. February 2015;437:275-281.

Sakamoto T, Danzmann R. A Microsatellite Linkage Map of Rainbow Trout (Oncorhynchus mykiss) Characterized by Large Sex.. Genetics [serial online]. July 2000;155(3):1331.

Sakane H, Horii Y, Nogami S, Kawano Y, Kaneko-Kawano T, Shirataki H. α-Taxilin Interacts with Sorting Nexin 4 and Participates in the Recycling Pathway of Transferrin Receptor. Plos ONE [serial online]. April 2014;9(4):1-11.

Salem, M., Rexroad, C. E., Wang, J., Thorgaard, G. H., & Yao, J. (2010). Characterization of the rainbow trout transcriptome using Sanger and 454-pyrosequencing approaches. BMC Genomics, 11(1), 564. http://doi.org/10.1186/1471-2164-11-564.

Salem, M., Vallejo, R. L., Leeds, T. D., Palti, Y., Liu, S., Sabbagh, A., … Yao, J. (2012). RNA-Seq identifies SNP markers for growth traits in rainbow trout. PloS One, 7(5), e36264. http://doi.org/10.1371/journal.pone.0036264.

Salem M, Paneru B, Al-Tobasei R, Abdouni F, Thorgaard GH, et al. (2015) Transcriptome Assembly, Gene Annotation and Tissue Gene Expression Atlas of the Rainbow Trout. PLoS ONE 10(3):e0121778. doi: 10.1371/journal.pone.0121778.

Sánchez O. I, Sanguino O. W, Gómez C. A, García C. R. Evaluation of a rainbow trout (Oncorhynchus mikyss) culture water recirculating system. Revista MVZ Córdoba [serial online]. September 2014;19(3):4226-4241.

Schenck, A., B. Bardoni, C. Langmann, N. Harden, J.L. Mandel, A. Giangrande. CYFIP/Sra-1 controls neuronal connectivity in Drosophila and links the Rac1 GTPase pathway to the fragile X protein. Neuron, 38 (2003), pp. 887–898

Storebakken, T., No, H.K., 1992. Pigmentation of rainbow trout. Aquaculture 100, 209-229.

Taylor E. B., Tamkee P., Keeley E. R., Parkinson E. A. (2011). Conservation prioritization in widespread species: the use of genetic and morphological data to assess population distinctiveness in rainbow trout (Oncorhynchus mykiss) from british columbia, canada. Evol. Appl. 4, 100–115 10.1111/j. 1752-4571.2010.00136.x.

Thorgaard G, Bailey G, Palti Y, et al. Status and opportunities for genomics research with rainbow trout. Comparative Biochemistry & Physiology Part B [serial online]. December 2002;133(4):609.

Tilton S. C., Gerwick L. G., Hendricks J. D., Rosato C. S., Corley-Smith G., Givan S. A., et al. (2005). Use of a rainbow trout oligonucleotide microarray to determine transcriptional patterns in aflatoxin b1-induced hepatocellular carcinoma compared to adjacent liver. Toxicol. Sci. 88, 319–330 10.1093/toxsci/kfi309.

Valdivia K, Jouanno E, Guiguen Y, et al. High Temperature Increases the Masculinization Rate of the All-Female (XX) Rainbow Trout "Mal" Population. Plos ONE [serial online]. December 2014;9(12):1-16.

Vignal A, Milan D, SanCristobal M, Eggen A. A review on SNP and other types of molecular markers and their use in animal genetics. Genetics, Selection, Evolution : GSE. 2002;34(3):275-305. doi: 10.1186/1297-9686-34-3-275.

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

Wong S., Waldrop T., Summerfelt S., Davidson J., Barrows F., Kenney P. B., et al. (2013). Aquacultured rainbow trout (Oncorhynchus mykiss) possess a large core intestinal microbiota that is resistant to variation in diet and rearing density. Appl. Environ. Microbiol. 79, 4974–4984 10.1128/AEM.00924-13.

Yoshida K, Nogami S, Shirataki H, et al. Interaction of the taxilin family with the nascent polypeptide-associated complex that is involved in the transcriptional and translational processes. Genes To Cells [serial online]. May 2005;10(5):465-476.

Zhang, Jing, Lingfei Hou , Eric Klann , David L. Nelson. Altered Hippocampal Synaptic Plasticity in the Fmr1 Gene Family Knockout Mouse Models Journal of Neurophysiology May 2009, 101 (5) 2572-2580; DOI: 10.1152/jn.90558.2008.

APPENDIX 1

DNA Extraction from Fin Clip

Extraction Buffer	Volume (mL)	Volume (mL)
1M Tris pH 8.0	0.5	5
0.5M EDTA pH 8.0	1	10
10% SDS	1	10
5M NaCl	1	10
Water	46.5	465
Total	**50**	**500**

1. **Cut** 1 cm^2 of fin tissue into small pieces. **Place** in a 1.5mL centrifuge microtube and **add** 500 µL extraction buffer, 5 µL DTT (1M), and 6 µL Pro-K (10mg/mL).
2. **Store** at 56 ˚C for a few hours or 37 ˚C overnight on a *rotary shaker*. If digestions are incomplete, add more Pro-K and **store** at 56 ˚C.
3. **Spin** in *centrifuge* for 2 min at 14,000rpm. **Transfer** supernatant to a new 1.5mL tube.
4. **Add** 200 µL 5M NaCl. **Mix** thoroughly, but do not vortex.
5. **Spin** in *centrifuge* for 10 min at 14,000 rpm.
6. **Transfer** clear supernatant to a new 2mL screw-cap tube containing 1 mL cold 100% EtOH.
7. **Store** on ice for 10 min or in -20 ˚C *freezer*.
8. **Spin** in *centrifuge* for 10 min at 14,000 rpm and 4 ˚C.
9. **Aspirate**, or pour out supernatant.
10. Perform steps A B, and C at least once. Repeat as needed.
    A. **Rinse** the pellet in 1 mL of 70% EtOH.
    B. **Spin** for 5 min. at 14,000 rpm and 4 ˚C.
    C. **Aspirate**, or pour out supernatant.
11. **Dry** in open air for 10-20 min.
12. **Add** 100 µL Low TE (10 mM Tris, pH8.0 and 0.1 mM EDTA). **Store** in 56 ˚C water bath for 1 hour or overnight at 37 ˚C.
13. **Quantify** DNA on *spectrophotometer*.
14. For PCR, **dilute** an aliquot to 12.5 ng/µL.

- *Or maximum rpm*
- *~40% of starting volume*
- *May store overnight in freezer at -20 ˚C*
- *Pellet should adhere to wall of tube.*
- *Dislodge pellet from wall of tube.*
- *Do not overly dry pellet.*
- *A260/A280 ratio should be around 1.75-2.0.*

APPENDIX 2

Scholars Week Abstract 2013

RNA-SEQ IDENTIFIES SNP GENETIC MARKERS AND DIFFERENTIAL GENE EXPRESSION ASSOCIATED WITH INCREASED MUSCLE YIELD IN RAINBOW TROUT

Ashlin Harris, Undergraduate, (Honors College); Rhett Layman, Graduate student; Mohamed Salem, Faculty; Mohamed Salem (Faculty Sponsor),

Background: Rainbow trout is an important food animal, making high muscle mass a desired trait. Muscle traits are complex, polygenic and difficult to improve by conventional selection. In addition, genetic marker assisted selection for desired phenotypes can be accomplished in fewer generations than traditional selection, allowing for quicker improvement in commercial populations. Single nucleotide polymorphisms (SNPs), which account for 90% of genetic differences among individuals, are ideal for marker assisted selection of desired phenotypic traits The objective of this project is utilization of state-of-the-art RNA-Seq (whole-transcriptome sequencing) analysis to identify SNP genetic markers and characterize genes controlling muscle yield.

Results: Phenotypic variations in muscle yield were measured in 100 families from the USDA rainbow trout breeding program. Variations in muscle yield were correlated to transcriptome-wide SNP allele frequencies and to global patterns of gene expression in families showing extreme phenotypes (4 high-muscle yield families [50.9% of BW ± 1.8] versus 4 low-muscle yield families [43.2% of BW ± 2.1]). We identified 127,401 presumptive SNPs; from these we identified 143 SNPs in the “high” group and 96 SNPs in the “low” group, all with a false discovery rate (FDR) ≤ 0.1. Currently, these putative SNP markers are being genotyped and evaluated for association with muscle yield in a ~500-fish panel. Differential gene expression between groups “high” and “low” identified 60 genes with a fold change ≥ ±2 and a FDR ≤ 0.1. Gene pathway analysis identified 239 SNPs and 60 differentially expressed genes responsible for amino acid and sugar metabolism as well as various metabolic pathways.

Conclusion: The study identified SNP markers and differentially expressed genes and gene pathways predictive of increased muscle yield in rainbow trout. These SNPs and differentially expressed genes can be used by aquaculture institutions, such as the USDA, to improve the efficiency of trait selection in rainbow trout.

APPENDIX 3

Scholars Week Abstract 2014

RNA SEQ IDENTIFIES SINGLE NUCLEOTIDE POLYMORPHISM (SNP) ASSOCIATED WITH MUSCLE YIELD AND SHEAR FORCE OF FLESH IN RAINBOW TROUT

Bam Paneru, Graduate student, Biology; Ashlin Harris, Undergraduate, Biology; Rhett Layman, Graduate student, Biology; Timothy D. Leeds, NCCCWA: Jianbo Yao, UWV: Brett Kenney, UWV; Mohamed Salem (Faculty Sponsor), Biology

Rainbow trout (*Oncorhynchus mykiss*) is one of the top five sport fish species in North America and the second most widely cultivated fish species for food. Muscle content and shear force are two important characteristics of fish flesh which determine consumer satisfaction. Single nucleotide polymorphism (SNP) determines 90% of genetic difference between individuals and is utilized in genetic evaluation and selective breeding. Our previous studies using the RNA-Seq technique were successful in finding SNPs markers associated with growth traits in rainbow trout. In this study, RNA-Seq was used to identify SNPs associated with fish muscle yield and shear force. RNA was isolated from fish belonging to 100 different families. These families were classified into two groups: high muscle yield families (50.9% of BW ± 1.8) and low muscle yield families (43.2% of BW ± 2.1). RNA sample from all individuals belonging to each muscle yield group was pooled and sequenced by illumina platform. Sequencing reads from high muscle and low muscle families were mapped to a transcriptome reference to predict the SNPs associated with muscle content and shear force. Phenotypic variations in muscle yield were measured in these 100 families from the USDA rainbow trout breeding program. Predicted SNPs were validated by the sequenom genotyping technology. Out of validated SNPs, 42 SNPs were evaluated for association with muscle traits in 1000 fish from two generations. The study identified SNP (potential genetic markers) associated with shear force.

APPENDIX 4

Scholars Week Abstract 2015

IDENTIFICATION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNP) MARKERS FOR MUSCLE YIELD AND QUALITY IN RAINBOW TROUT USING RNA SEQ

Ali Ali, Graduate student, Biology; Bam Paneru, Graduate student, Biology; Rafet Al Tobasei, Graduate student, Biology; Mohamed Salem, Staff, Biology; Mohamed Salem (Faculty sponsor), Biology

Background: Rainbow trout (*Oncorhynchus mykiss*) is one of the favorite sport fish species in North America and the second most important fish species for aquaculture in the U.S. Muscle yield and quality traits are important factors of profitability for the food fish aquaculture industry. Most of the genetic variations are caused by Single nucleotide polymorphism (SNP). Our previous studies were successful in finding SNP markers associated with different traits in rainbow trout.

Methods: In the current study, whole-transcriptome sequencing (RNA-Seq) was used to detect SNPs associated with fish muscle yield, fat content, flesh whiteness and softness (shear force). Variations in these muscle quality traits were measured in about a 100 families from two generations produced at the USDA/NCCCWA rainbow trout breeding program. RNAs isolated from four full-sib families (5 fish per family), sorted as high or low for each of these traits, were separately sequenced using an illumine RNA-Seq technique. Sequences from different families were mapped to the reference genome/ transcriptome to predict SNPs that are associated with each quality traits. Predicted SNPs were confirmed on a small set of fish using the Sequenom genotyping technology. Then, the validated SNPs were assessed for association with the aforementioned traits in approximately 1000 fish from two generations. Results: The study identified seven SNPs (potential genetic markers) associated with these quality traits. In addition, digital gene expression profiling identified a small number of differentially regulated protein coding genes and long non-coding RNAs (LncRNAs) that are associated with each quality attribute, perhaps, indicating that selective breeding have had a relatively low impact on modifying gene expression.