

ABSTRACT

RAJAB, SARAH, A. S.. An Integrated Metabolomic and Transcriptomic Approach for Understanding White Muscle Growth Regulation in Hybrid Striped Bass Aquaculture. (Under the direction of Dr. Harry V. Daniels and Dr. Benjamin J. Reading).

Farming of hybrid striped bass (HSB) (striped bass x white bass) is the fourth largest finfish aquaculture industry in the U.S., behind catfish, salmon and tilapia, however several knowledge gaps remain, particularly in regards to the cellular processes that underlie growth performance of these fish. Insight into these processes is critical for improvement of breeding and rearing methodologies geared towards foodfish production. Here, we describe a novel combinatorial machine learning analysis of transcriptomic and metabolomic data generated from white muscle and liver tissues to identify discrete molecular processes that underlie the regulation of enhanced growth and muscle accretion of HSB. HSB were reared in tanks and ponds according to standard two-phase culture procedures until harvest at market size. Individuals of the top and bottom 10% body size in weight and total length relative to the cohort were sampled (N = 10 fish per group) as representatives of fish that grow good and poorly, respectively. These cutoffs were chosen as they represent a body size that is typically larger and smaller than that desired at market and therefore represent extremes of the distribution. The data were linked to corresponding changes in metrics of growth including body weight, total length, hepatosomatic index, and gonadosomatic index. Transverse sections of white muscle tissue were taken to evaluate differences in muscle growth patterns and revealed that good-growth group fish trended towards hyperplastic growth and that poor-growth group fish had switched to hypertrophic growth. Additionally, muscle and liver samples were subjected to metabolomics analysis and a total of 469 and 464 metabolites, in muscle and

liver, respectively were identified and analyzed by decision tree, where 30 metabolites were classified as the top most important growth predictors. As metabolites are the end products of gene and protein activities within the cell, they therefore underlie the physiological response of the animal. A determination of differential gene expression between groups was completed with Illumina-based RNA-Seq data generated from muscle tissue and a total of 72,893 genes were identified and ranked by importance using statistical inference, based on adjusted *p*-value, and machine learning (Support Vector Machines). Two reduced gene lists, 143 from inferential statistics and 150 from SVMAttributeEval machine learning model, were selected as a cutoff for pathway analysis and only eleven genes were found to be shared between the two lists. To ascertain the biological processes that determine size differences in the two groups of fish, Ingenuity Pathway Analysis (IPA) was conducted on the two differentially expressed gene lists and the top 30 metabolites. The IPA results, identified potential candidates and pathways that influence growth performance for guiding future experimental testing and feed design. In particular, pathways such as cell proliferation and cell differentiation, apoptosis, and inflammation may possibly underly the growth performance of the fish. Furthermore, combined with application of genetic improvement technology, manipulation of these pathways may enhance animal growth and body size. These studies combined, for the first time, two powerful approaches (metabolomics and transcriptomics), along with a novel machine learning-based model analysis tool using Support Vector Machines (SVMAttributeEval), to generate an informative list of the most important metabolite and gene biomarkers and their likely integrated pathways associated with HSB growth performance. These approaches, which ten years ago were neither available nor accessible,

provide the most practical predictors of growth performance in fish and identify key metabolites and gene transcripts that may be indicators of growth performance. The integrated pathways of these components provide a great insight into the mechanisms of muscle development, which will inform future breeding and nutrition studies to help improve HSB aquaculture production and assure the successful continuation of the industry.

© Copyright 2020 Sarah A. S. Rajab

All Rights Reserved

An Integrated Metabolomic and Transcriptomic Approach for Understanding White Muscle
Growth Regulation in Hybrid Striped Bass Aquaculture

By
Sarah A. S. Rajab

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Zoology

Raleigh, North Carolina
2020

APPROVED BY:

Dr. Harry V. Daniels
Co-Chair of Advisory Committee

Dr. Benjamin J. Reading
Co-Chair of Advisory Committee

Dr. Peter Ferket

Dr. Christopher Ashwell

DEDICATION

This thesis work is wholeheartedly dedicated to my beloved parents, Ahmad and Batoul, for encouraging my pursuit in sciences from a young age. For loving me unconditionally and whose good examples have taught me to work hard for the things that I aspire to achieve. They have been my source of inspiration and gave me strength when I thought of giving up, and have continuously provided their moral, spiritual, emotional, and financial support.

This work is also dedicated to my beloved husband, Yousef. Yousef is also my best friend for being my constant source of support and encouragement during the challenges of graduate school and life. I am thankful that you have always believed in me and I am genuinely grateful for having you in my life.

BIOGRAPHY

Sarah Ahmad Sayed Rajab was born during the year 1984 to Ahmad Rajab and Batoul Haji. She has been passionate for science from a very young age and was mostly fascinated in nature and explicitly in animals. This led her to obtain a zoology degree with a minor in marine sciences from Kuwait University. She was awarded an honor student certificate for two consecutive years 2005-2006 during her undergraduate studies. Throughout her undergraduate career she had the opportunity to be involved in variety of different research projects all which were integrated with Kuwait marine environment. From that point onward, Sarah's passion for the marine environment grow more and she contained to peruse her love for the sea by working during the summer at Lothan Youth Achievement Centre (LOYAC).

Shortly after her graduation in 2006, she joined Kuwait Institute for Scientific Research (KISR) as a Research Assistant specializing in fish age determination and stock assessment. During that time, she participated in collaborated projects with The Gulf Cooperation Council (GCC). While there, she mastered several laboratorial techniques such a tissue and bone histology for fish age determination, Principles of Polymerase Chain Reaction (PCR), fish anatomy and classification , image analysis using highly advanced motorized compound, stereo, and scanning electron microscopy. Additionally, she was given the opportunity to attend a stock assessment modeling training course arranged by the International Council for the Exploration of the Sea (ICES) in Copenhagen, Denmark. Also she participated in the world 5th fisheries congress in Yokohama, Japan.

In KISR, Sarah was granted a scholarship to continue her path in the field. She was awarded a Master Degree in “Applied Marine and Fishers Ecology an Implementation to Marine Protected Areas” at Aberdeen University in Scotland, the United Kingdom. Throughout her time in Scotland, Sarah also worked in Marine Scotland Science (MSS) where she refined her skills in the fish age determination laboratory, until her graduation in 2010.

After graduation, Sarah continued working for KISR until August 2014 where she was matriculated to Kuwait University and granted a scholarship to pursue a doctorate degree at the North Carolina State University (NCSU). During her time in NSCU, she was able to work with a number of undergraduate and graduate students and present her research at several national scientific symposia.

As of the writing of this dissertation, Sarah currently lives in Raleigh, NC with her Husband Yousef, her Mother Batoul, and her newly born son Mohammed. She has five siblings, two older brothers, Hashem and Mohammed, two older sisters, Saja and Sakeena, and one younger brother, Ibraheem. Sarah’s hobbies include watercolor painting, photography, reading National Geographic, and exploring the world.

ACKNOWLEDGMENTS

First and foremost, praises and thanks to God, the Almighty, for His showers of blessings throughout my research work to complete the research successfully.

I offer my sincerest gratitude to my committee advisors, Dr. Benjamin Reading and Dr. Harry Daniels, my mentors, and trusted guides for their unending support and encouragement throughout my Northern Carolina State University experience. Dr. Benjamin Reading nurtured and refined my scientific research talents by encouraging me to be bold and creative, not afraid to “think outside the box.” I will forever be indebted to him for how he enriched my understanding of science, the scientific research process, and the ability to approach problem-solving with precision and clarity. It was a great privilege and honor to work and study under his guidance. I am incredibly grateful for what he has offered me. I would also like to thank him for his friendship, empathy, and a great sense of humor. I am extending my heartfelt thanks to his wife, family, for their acceptance and patience during the discussion I had with him on research work and thesis preparation. My most profound appreciation to Dr. Harry Daniels, Senior Associate Dean of the College of Agriculture and Life Sciences, for giving me the opportunity to proceed with my graduate studies here at North Carolina State University. His dynamism, vision, sincerity, and motivation have deeply inspired me. He has taught me the methodology to carry out the research and present the research works as clearly as possible and he provided me with an invaluable guidance throughout my graduate school life.

Also, my committee members, Dr. Peter Ferket and Dr. Christopher Ashwell, who helped me understand and deal effectively with countless challenges and taught me how to

solve problems from various angles. Their lessons have been instrumental in my intellectual development and will guide me throughout my life and career.

I want to say thank you to my friends and research colleagues, Linnea Andersen, and Erin Ducharme. Linnea, who was there whenever needed, always positive, supportive, and eager to sacrifice her time and energies to assist and encourage me. She also provided badly needed “Funny MEMES” at times during the long hours working in the office. Erin, who went through the trials of graduate school with me and helped me through her work to finish writing up the dissertation on time. Being a graduate student has been an amazing experience and I would never have been able to reach to this accomplishment without the guidance and tremendous academic and emotional support of all the amazing graduate student in office 208, Elsit Kiekebusch, Eugene Cheung, Jonathan Giacomini, and Laura Hamon.

This research would not have been possible without the help from the various field labs and their staff who greatly improved my knowledge and understanding of aquaculture systems. They included Dr. Andrew McGinty, John Davis, and Joseph Bursey.

I offer my appreciations to my family; if not for their love, understanding, encouragement, wise counsel, and endless support, my walk would have been far more difficult. For them and their love, I am truly blessed far more than I will ever deserve. Also, I lovingly thank my husband, Yousef, who has sacrificed so much to help me succeed. He has always been patient and reliable throughout this process and his compassionate and understanding words of encouragement were my mainstay.

Lastly, but by no means least, I owe my deepest appreciativeness to Kuwait University, who believed in me and gave me the opportunity to obtain the Doctor of Philosophy at the North Carolina State University.

Sarah A. S. Rajab

TABLE OF CONTENTS

LIST OF TABLES	xiii
LIST OF FIGURES	xxi
CHAPTER I: Literature Review	1
Hybrid striped bass background.....	1
An overview of muscle growth in fishes	3
An overview of reduced frequency feeding strategies in aquaculture	7
Objectives.....	10
Study Design	10
Organization of the dissertation.....	12
References	16
Chapter II: White Muscle and Liver Metabolites Involved In Enhanced Growth	
Response of Hybrid Striped Bass	28
Abstract	28
Introduction	29
An overview of metabolites and muscle growth	29
An overview of technology platform for metabolomics	30
Specific literature review on metabolomics and proteomics of fish growth.....	31
Study design.....	33
Materials & Methods	34
Experimental animals and tissue sampling.....	34
Evaluations of growth parameters	36
Muscle histology analysis	37

Metabolomics analysis.....	38
Data Analysis	39
Growth parameters	39
Muscle histology	40
Metabolomics.....	40
Results	41
Experimental animals, tissue sampling, and evaluations of growth parameters	41
Muscle histology analysis	42
Metabolomics analysis.....	42
Discussion	44
References	58
 Chapter III: White Muscle Genes Involved In Enhanced Growth Response of	
Hybrid Striped Bass	
	109
Abstract	109
Introduction	110
An overview of transcriptomics.....	110
An overview of technology platform for RNA sequencing (RNA-Seq).....	111
Specific overview on myogenesis of fish growth	112
Study design.....	113
Materials & Methods	114
Experimental animals and tissue sampling.....	114
Evaluations of growth parameters	114
Muscle histology analysis	115

Data Analysis	115
Growth parameters	115
Muscle histology	115
Gene expression analysis.....	115
Quantitative RNA-Seq analysis.....	116
Machine learning evaluation of RNA-Seq expression.....	118
DAVID pathway analysis.....	122
Results	122
Experimental animals, tissue sampling, and evaluations of growth parameters...	122
Muscle histology analysis	122
Gene expression analysis.....	123
Quantitative RNS-Seq analysis	123
Machine learning evaluation of RNA-Seq expression.....	124
DAVID pathway analysis.....	129
Discussion	130
References	135
Chapter IV: A Correlation Between Gene Expression And Metabolites	169
Abstract	169
Introduction	170
Integrated pathways of metabolomics and transcriptomics data.....	170
Specific overview on integrated metabolomics and transcriptomics.....	171
Material & Methods	174
Comparison datasets.....	175

Ingenuity Pathway Analysis (IPA) upload	175
Ingenuity Pathway Analysis (IPA) Core Analysis.....	176
IPA Core Analysis settings.....	177
IPA Core Analysis statistics.....	178
IPA Comparison Analysis.....	180
IPA BioProfiler Analysis	181
Results & Discussion.....	181
Canonical Pathways	181
Upstream Regulators.....	183
Diseases and Functions	184
Comparison Analysis.....	186
BioProfiler Analysis	187
References.....	192
Chapter V: The Effects of Frequent and Infrequent Feeding Regimes on the Growth of Sunshine Hybrid Striped Bass (Male Striped Bass <i>Morone saxatilis</i> x Female White Bass <i>Morone chrysops</i>)	215
Abstract	215
Introduction.....	216
Materials & Methods.....	218
Results.....	221
Discussion & Conclusion.....	223
References.....	228
Chapter VI: Summary and Future Directions	241

References..... 250

LIST OF TABLES

CHAPTER I: Literature Review

Table 1.1	Growth of white axial muscle (removed from a site dorso-lateral and slightly posterior to the dorsal fin) fibers (measured under the microscope for the number and “diameter” of each fiber) of ten species of freshwater teleosts from five families (Cyprinidae, Centrarchidae, Percidae, Salmonidae, Esocidae) possessing widely different growth rates and ultimate sizes have been studied. The dynamics of muscle increase (i.e., increase in fiber numbers and/or diameter) appears to determine the ability for rapid somatic growth and large ultimate size in teleosts. The largest and fastest growing species (smallmouth bass, lake whitefish, rainbow trout, muskellunge) show evidence of sustained recruitment of muscle fibers to a large size, in contrast to the smaller and slower growing species (bluntnose minnow, longnose dace, grass pickerel). Pumpkinseed, bluegill, common bully, and yellow perch are all intermediate in fiber growth dynamics, growth, and ultimate size between the smaller and larger species (Weatherley et al., 1988)24
-----------	---

Chapter II: White Muscle and Liver Metabolites Involved In Enhanced Growth

Response of Hybrid Striped Bass

Table 2.1	Classification of metabolomic approaches and the optimal technology used. The terminologies are still evolving and there can be overlaps in their definition, however, the following classification highlights the
-----------	--

	current options available for monitoring the metabolome (Goodacre et al., 2004)	67
Table 2.2	Proteins associated with fast and slow growing yellow perch. Nine proteins/peptides were identified as predominantly metabolic enzymes, including: Alpha actinin, Glycogen phosphorylase, Phosphoglucose isomerase, ATP synthase, Enolase, Creatine kinase, Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), lactate dehydrogenase, and adenylate kinase. All of these proteins with the exception for sarcoendoplasmic reticulum calcium ATPase (SERCA) were found in bands that showed increased expression in fast-growing fish compared to slow-growing fish suggesting possible association of these proteins with faster growth of yellow perch when expressed at higher levels (Kwasek et al. 2012; Pocket, 2006).....	69
Table 2.3	List of the 10 most important proteins that are directly related to muscle metabolic activities out of the 49 proteins identified, which were extracted from frozen muscle tissues of European sea bass obtained by 2-D electrophoresis (2-DE) and tandem mass spectrometry (MS/MS) analysis. The protein identities are listed according to their accession number obtained from the UniProt database (Terova et al., 2014)	71
Table 2.4	Hybrid striped bass growth parameters data for the poor-growth and good-growth groups. Values are provided for individual fish (Fish No.) along with mean values and SDE (standard error of the mean).....	73

- Table 2.5 Raw area counts of all metabolites identified within hybrid striped bass liver tissues. Supplemental file available at North Carolina State University library online repository.
- Table 2.6 Raw area counts of all metabolites identified within hybrid striped bass muscle tissues. Supplemental file available at North Carolina State University library online repository.
- Table 2.7 Statistical comparisons between the poor- and good-growth groups identifying the number of significant metabolites found in the hybrid striped bass liver tissue using Welch's Two-sample t-tests. Two thresholds were used: 1) for statistical significance, a $p \leq 0.05$ was used as a statistical cutoff, and 2) $0.05 < p < 0.10$ was used to determine a trend towards significance. The chart describes the total number of significantly different metabolites (109 in the good-growth group compared to the poor-growth group) with an indication of those metabolites that were elevated (30) and decreased (79) in the good-growth group. A gender effect was noted in the principle component analysis (Figure 2.7), so a statistical contrast was also provided for the female cohort. (F): denotes only female samples were compared75
- Table 2.8 Statistical comparisons between the poor- and good-growth groups identifying the number of significant metabolites found in the hybrid striped bass muscle tissue using Welch's Two-sample t-tests. Two thresholds were used: 1) for statistical significance, a $p \leq 0.05$ was used as a statistical cutoff, and 2) $0.05 < p < 0.10$ was used to determine a

trend towards significance. The chart describes the total number of significantly different metabolites (50 in the good-growth group compared to the poor-growth group) with an indication of those metabolites that were elevated (23) and decreased (27) in the good-growth group. A gender effect was noted in the principle component analysis (Figure 2.8), so a statistical contrast was also provided for the female cohort. (F): denotes only female samples were compared76

Table 2.9 Confusion matrix showing the predictive accuracy of the Random Forest analysis (Figure 2.7), conducted on hybrid striped bass liver tissues (N=9). Predictive accuracy was 89% and calculated in the shown table by comparing the predicted group versus actual group (Class Error).....77

Table 2.10 Confusion matrix showing the predictive accuracy of the Random Forest analysis (Figure 2.8), conducted on hybrid striped bass muscle tissues (N=9). Predictive accuracy was 72% and calculated in the shown table by comparing the predicted group versus actual group (Class Error).....78

Chapter III: White Muscle Genes Involved In Enhanced Growth Response of Hybrid Striped Bass

Table 3.1 List of selected rainbow trout skeletal muscle genes grouped by functional categories. An RNA-seq of the transcriptome of skeletal muscle identified 1,085 gene sequences in red muscle (R) and 1,228 gene sequences in white muscle (W) (for more information about the other genes see Palstra et al., 2013) 143

Table 3.2	List of selected paralogous genes identified in the Gilthead sea bream skeletal muscle transcriptome that are mainly related to fast skeletal muscle activities (Estévez et al., 2012).....	145
Table 3.3	Summary of hybrid striped bass gene expression data. The number of expressed genes that had an average read count of ≥ 5 , informative genes that passed DESeq2 filtering, and differentially expressed genes (DEGs) are provided. There were 143 up-regulated and 9 down-regulated genes in fish from poor-growth group. These DEGs were annotated via the top 5 hits of BLASTN, against non-redundant nucleotide of NCBI, with a e-value cutoff of 1×10^{-10}	146
Table 3.4	List of the top 143 most important genes in the white skeletal muscle of hybrid striped bass (143 is the number of genes that significantly differed in expression between fish from the poor- and good- growth groups). The genes were ranked based on their significant <i>p</i> -value level ($q \leq 0.05$, FDR = 0.05). The difference in the elevation level between the two growth groups (i.e., poor- and good-growth) is expressed as the average FPKM value (Fragments per Kilobase of Transcript per Million Mapped Reads). The black box in the table represents the upregulated genes between the two growth groups. The variant designation indicates genes that were identified more than once, such that different <i>de novo</i> transcript assemblies corresponding to the same gene were identified (e.g., spliced variants).....	147

Table 3.5 List of the top 150 most important genes in the white skeletal muscle of hybrid striped bass (150 is the number of genes that were arbitrarily chosen for the pathway analysis). The genes were ranked based on the Machine learning SVMAttributeEval model. The difference in the elevation level between the two growth groups (i.e., poor- and good-growth) is expressed as the Average FPKM value (Fragments per Kilobase of transcript per Million mapped reads). The black box in the table represents the upregulated genes between the two growth groups (i.e., poor- and good growth). The variant designation indicates genes that were identified more than once, such that different *de novo* transcript assemblies corresponding to the same gene were identified (e.g., spliced variants)..... 154

Table 3.6 List of alphabetically arranged hybrid striped bass muscle genes that were shared between the two lists of the top 143 and 150 (143 is the number of genes that significantly differed in expression between fish from the poor- and good- growth groups by *p*-value ($q \leq 0.05$, FDR = 0.05) and 150 is the number of genes that were chosen for the pathway analysis based on machine learning rankings by SVMAttributeEval). The table shows the ranking of each gene in each classifying approach and the number of variants for each of the genes also is provided; more than one ranking number corresponds to the different variant designations. The arrows correspond to the gene elevation level which was based on

the value of the average FPKM (Fragments per Kilobase of transcript per Million mapped reads) between the two growth groups..... 161

Table 3.7 DAVID Functional Annotation Clustering report based on the top 150 most important expressed gene transcripts to hybrid striped bass growth as ranked using Support Vector Machine (SVMAttributeEval) model. Each Annotation cluster shares gene members with similar biological function. Gene Ontology (GO) class provides the common biological characteristics for all the Gene Members in the corresponding cluster. The Enrichment Score and the *p*-value of each cluster also are provided..... 162

Chapter IV: A Correlation Between Gene Expression And Metabolites

Table 4.1 Qiagen IPA Diseases and Function Analysis of muscle GeneListA (ranked by *p*-value) and GeneMetaboliteListA (ranked by *p*-value) identified some molecules that were predicted to either activate or inhibit key pathways that were related to the growth performance of hybrid striped bass muscle from the good-growth group based on the gene and metabolite signatures provided. These genes and metabolites were represented in the network map presented in Figure 4.9. The table shows the category of each molecule, their elevation level, and their direct effect on the growth performance 202

Chapter V: The Effects of Frequent and Infrequent Feeding Regimes on the Growth of Sunshine Hybrid Striped Bass (Male Striped Bass *Morone saxatilis* x Female White Bass *Morone chrysops*)

Table 5.1 Water temperature and salinity data during the hybrid striped bass feeding trials 233

Table 5.2 Summary data of the sunshine hybrid striped bass growth trials including number of fish per replicate tank, initial weight, initial feeding rate (FR % body weight), final weight, final FR, weight fed, and feed conversion ratio (FCR) for fish either were fed by hand or automated belt feeder at either three or five days per week (Hand x 3, Belt x 3, Hand x 5, and Belt x 5) 234

Table 5.3 Summary of the *p*-values for the sunshine hybrid striped bass growth trials including initial weight, initial feeding rate (FR % body weight), final weight, final FR, weight fed, and feed conversion ratio (FCR) for fish either were fed by hand or automated belt feeder at either three or five days per week (Hand x 3, Belt x 3, Hand x 5, and Belt x 5) 235

LIST OF FIGURES

CHAPTER I: Literature Review

Figure 1.1 A. Striped bass (*Morone saxatilis*). B. White bass (*Morone chrysops*). C. Sunshine cross (♂ Striped bass (*M. saxatilis*) x ♀ White bass (*M. chrysops*)), it can be distinguished from the striped bass by broken rather than solid horizontal stripes on the body (from AquaSol, Inc, 2016).....26

Figure 1.2 Satellite cells occupy a sub-laminar position in adult skeletal muscle. Elongated, spindle-shaped, quiescent myoblasts lying in close contact with adult muscle. The satellite cells can be distinguished from the myonuclei by a surrounding basal lamina and more abundant myofibril. The satellite cells are thought to play a role in muscle repair and regeneration (from Hawke et al., 2001)27

Chapter II: White Muscle and Liver Metabolites Involved In Enhanced Growth

Response of Hybrid Striped Bass

Figure 2.1 Pie chart distribution of protein identities directly related to muscle metabolic activities of 49 proteins identified related to muscle growth in European sea bass based on: A) Function of the protein, B) Location of the proteins (Terova et al., 2014)79

Figure 2.2 Bar charts showing the weight class (g) frequency in numbers of the hybrid striped bass. The white bars represent the top and bottom 10% in terms of body size (N=10 fish each group, 10 x 2 fish total) that were selected as representatives of fish that grow good (good-growth) or poor

(poor-growth), respectively. These cutoffs were chosen as they represent a body size that is typically larger and smaller than that desired at market and therefore represent the extremes of the distribution.....80

Figure 2.3 Bar charts showing the (a) average total length (mm), (b) average wet weight (g), (c) average Gonadosomatic index (GSI), (d) average Hepatosomatic index (HSI), (e) average Condition Factor (K), and (f) Average Specific Growth Rate (SGR) of hybrid striped bass in each growth group (poor-growth or good-growth). Average values are represented by each bar and the standard error of the mean is shown by the brackets. There were ten fish in each group. There were significant differences between groups (One-way Student’s t-test at an alpha level of 0.05; *** $p = 0.0001$; ** $p = 0.0037$; * $p = 0.0147$)81

Figure 2.4 Representative histological cross-sections of hybrid striped bass (HSB) white muscle from fish in the poor- (a) and good- (b) growth groups. TL: Fish total length, WW: Fish Wet weight, Avg. FN: Average muscle fiber number, Avg. FD: Average muscle fiber diameter (μm). Comparing the two cross-sections, HSB in the poor-growth group have noticeably fewer, but larger muscle fibers, indicating hypertrophic growth status, whereas HSB in the good-growth group have more numerous but smaller muscle fibers, indicating hyperplastic growth. The bar corresponds to 100 μm 83

Figure 2.5 Bar charts showing the (a) average total muscle fiber number, (b) average muscle fiber number ($\leq 20 \mu\text{m}$), and (c) average muscle fiber

diameter (μm) of sunshine hybrid striped bass in each growth group. The fish were divided into the poor-growth or good-growth groups. Average values are represented by each bar and the standard error of the mean is shown by the brackets. There were ten fish in each group. There were significant differences between groups (One-way Student's t-test at an alpha level of 0.05; * $p = 0.0274$).....84

Figure 2.6 Linear regression graph between the: a) Average total muscle fiber number and hybrid striped bass wet weight (g) of the poor- and good-growth groups, b) Average muscle fiber diameter (μm) and the fish wet weight (g) of the poor- and good- growth groups. For each growth group, the linear regression equation is given as $y = mx + b$, where the slopes (m) of the lines represent the correlation between the variables in each growth groups, and correlation coefficient (R^2). A linear correlation between the Average total muscle fiber number and fish Wet Weight of the good- and poor-growth groups showed a weak correlation ($R^2=0.0858$ and 0.0185 , respectively) with little relationship. A linear correlation between the Average muscle fiber diameter and fish Wet Weight of the good- and poor-growth groups showed a weak correlation ($R^2=0.1813$ and 0.2165 , respectively)85

Figure 2.7 Principle Component Analysis (PCA) of hybrid striped bass liver tissue. Component 1 appeared to separate the growth groups based on gender. A statistical test is provided in Table 2.5 for female cohort to help discriminate gender-related effects in the dataset86

Figure 2.8 Principle Component Analysis (PCA) of hybrid striped bass muscle tissue. Component 1 appeared to separate the growth groups based on gender. A statistical test is provided in Table 2.6 for female cohort to help discriminate gender-related effects in the dataset87

Figure 2.9 Random Forest analysis identifying the top thirty most important metabolites when comparing liver tissue of hybrid striped bass from the poor- and good-growth groups. Metabolites were ranked by increasing importance and mean decrease accuracy (sensitivity analysis) during Random Forest decision tree analysis. The legend describes the metabolic super-pathway to which each listed metabolite belongs. An asterisk designates metabolites with identifications that could not be confirmed based on a control standard, but Metabolon was confident in its identification.....88

Figure 2.10 Random Forest analysis identifying the top thirty most important metabolites when comparing muscle tissue of hybrid striped bass from the poor- and good-growth groups. Metabolites were ranked by increasing importance and mean decrease accuracy (sensitivity analysis) during Random Forest decision tree analysis. The legend describes the metabolic super-pathway to which each listed metabolite belongs. An asterisk designates metabolites with identifications that could not be confirmed based on a control standard, but Metabolon was confident in its identification.....89

Figure 2.11 Boxplots showing (a) glucose, (b) pyruvate, (c) UDP-N-acetylglucosamine, and (d) UDP-N-acetylgalactosamine of scaled intensity range for poor- and good-growth groups of hybrid striped bass white muscle tissue. Welch’s Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass muscle tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the “+” symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.....90

Figure 2.12 Schematic diagram of the tricarboxylic acid (TCA) cycle where the grey sub-pathways shown correspond to the metabolites identified as important for growth in the hybrid striped bass white muscle tissue.....91

Figure 2.13 Boxplots showing (a) butyrylcarnitine, (b) propionylcarnitine, and (c) succinylcarnitine of scaled intensity range for poor- and good- growth groups of hybrid striped bass white muscle tissue. Welch’s Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass muscle tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the “+”

symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.....92

Figure 2.14 Boxplots showing (a) alpha-ketoglutarate, (b) acetylphosphate, and (c) citrate of scaled intensity range for poor- and good- growth groups of hybrid striped bass white muscle tissue. Welch’s Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass muscle tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the “+” symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.....93

Figure 2.15 Boxplots showing (a) cortisol, (b) 12-HETE, (c) 12-HEPE, (d) 15-HETE, (e) cholate, (f) taurocholate, and (g) 7-hydroxycholesterol of scaled intensity range for poor- and good- growth groups of hybrid striped bass white muscle tissue. Welch’s Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass muscle tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in

each group. Median value is shown as the line through each box. Mean value is represented as the “+” symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.....94

Figure 2.16 Boxplots showing (a) cholate, (b) taurocholate, and (c) cholesterol of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch’s Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (there were no significant differences between groups, alpha = 0.05). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the “+” symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers95

Figure 2.17 Schematic diagram of the cholesterol metabolic pathway. The schematic representation of the cholesterol biosynthetic pathway includes a number of cholesterol by-products including bile acids and cortisol. The pathway shown corresponds to the metabolites identified as important to growth in the hybrid striped bass liver tissue.....96

Figure 2.18 Boxplots showing (a) choline phosphate, (b) glycerophosphorylcholine (GPC), and (c) glycerophosphoethanolamine (GPE) of scaled intensity

range for poor- and good- growth groups of hybrid striped bass white muscle tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass muscle tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.....97

Figure 2.19 Boxplots showing (a) glucose, (b) glucose-6-phosphate, (c) fructose-6-phosphate, (d) dihydroxyacetone phosphate (DHAP), (e) phosphoenolpyruvate (PEP), and (f) 3-phosphoglycerate of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers98

Figure 2.20 Boxplots showing (a) pyruvate, (b) sedoheptulose-7-phosphate, (c) AICA ribonucleotide (AICAR), (d) hypoxanthine, (e) xanthine, (f) allantoin, (g) adenosine 3'-monophosphate (3'-AMP), (h) adenosine 3',5'-diphosphate (ADP), and (i) guanosine 5'- monophosphate (5'-GMP) of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers 99

Figure 2.21 Boxplots showing (a) xanthine, (b) adenosine 2'-monophosphate (2'-AMP), (c) guanine, (d) inosine, and (e) allantoin of scaled intensity range for poor- and good- growth groups of hybrid striped bass white muscle tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass muscle tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a

small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers 100

Figure 2.22 Boxplots showing (a) maltohexaose, (b) maltopentaose, and (c) maltotetraose of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch’s Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (there were no significant differences between groups, alpha = 0.05). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the “+” symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers..... 101

Figure 2.23 Boxplots showing (a) citrate, (b) acetyl-CoA, (c) acetylphosphate, (d) alpha-ketoglutarate, and (e) glutamine of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch’s Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented

as the “+” symbol; potential outliers are indicated with a small circle.

Upper and lower quartile range are the top and bottom of each box.

Maximum and minimum distribution for each group is designated by the

top and bottom whiskers 102

Figure 2.24 Boxplots showing a number of amino acids (a) aspartate, (b) glutamine, (c) histidine, (d) isoleucine, (e) leucine, (f) methionine, (g) proline, (h) serine, (i) threonine, (j) tryptophan, (k) tyrosine, and (l) valine of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch’s Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the “+” symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers 103

Figure 2.25 Boxplots showing (a) urea, (b) ornithine, (c) citrulline, (d) creatine, and (e) creatinine of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch’s Two-sample t-tests comparison between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (* indicates significant difference between groups, $p \leq 0.05$). There

were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the “+” symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers 104

Figure 2.26 Boxplots showing (a) 4-methyl-2-oxopentanoate, (b) 3-methyl-2-oxovalerate, (c) 3-methyl-2-oxobutyrate, (d) succinate, and (e) succinylcarnitine of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch’s Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the “+” symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers 105

Figure 2.27 Schematic diagram for some of the branched chain amino acids (BCAA) (valine, isoleucine, leucine) catabolism pathways. Catabolized BCAA were assimilated for energy production through the tricarboxylic acid (TCA) cycle. Consistent with catabolic use, the keto-acids (4-methyl-2-oxopentanoate, 3-methyl-2-oxovalerate, and 3-methyl-2-oxobutyrate) were elevated. The pathway shown corresponds to metabolites

identified as important to growth in the good-growth group hybrid striped bass liver tissue 106

Figure 2.28 Boxplots showing (a) carnitine-dicarboxylate, (b) adipoylcarnitine, and (c) pimeloylcarnitine/3-methyladipoylcarnitine of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers 107

Figure 2.29 Boxplot showing adipoylcarnitine of scaled intensity range for poor- and good- growth groups of hybrid striped bass white muscle tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify significant metabolites found in the hybrid striped bass muscle tissue (there was no significant difference between groups, $\alpha = 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and

minimum distribution for each group is designated by the top and bottom whiskers..... 108

Chapter III: White Muscle Genes Involved In Enhanced Growth Response of Hybrid Striped Bass

Figure 3.1 Muscle specification from the mesoderm depends on myogenic transcription factors (PAX3, PAX7, MyoD, Myf5, myogenin and Myf6). Postnatal muscle maintenance and regeneration invokes quiescent muscle precursor cells within the satellite cell pool that proliferate, terminally differentiate and fuse to generate multinucleated myotubes (From Hattmer et al., 2010) 163

Figure 3.2 Principal Component Analysis using top 500 highest variance genes in the white muscle samples of hybrid striped bass. The data were clustered together based on growth performance of the white muscle in the hybrid striped bass, where the light grey color points show the poor-growth group, and the dark grey points show the good-growth group. PC1 and PC2, are the parameters that comprise these 500 genes and together they explained 91% of the variation in growth performance 164

Figure 3.3 Performance of machine learning models to evaluate all informative gene transcripts (72,893) in hybrid striped bass white muscle (i.e., related to the good- and poor-growth groups) that were ranked by importance during classification using the statistical inference tests ($p \leq 0.05$). Hybrid striped bass growth status was predicted (classified into good-growth and poor-growth groups) based on expression of

muscle gene transcripts. (a) The Multilayer Perceptron Artificial Neural Networks (MLP) model, (b) The J48 Decision Tree (J48), (c) The Sequential Minimal Optimization Support Vector Machines (SMO) model, and (d) Random Forest model were used. Data points were fit with a polynomial line of the second order. The graphs show model robustness performance of the cross-validation as percentage of correctly classified instances when using the top highly ranked genes (10, 25, 50, 75, 100, 150, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, and 10000). Two cross-validation strategies were used to evaluate each model: (1) a percentage split whereby 66% of the data were randomly selected and used to train the models and the remaining 33% of the data were input as a cross-validation and (2) a 8-fold stratified hold out with n = 8 folds where one fold was used for cross-validation and n – 1 folds of the randomly reordered data set were used for training. Both classes (good- and poor-growth groups) were properly represented in the model training and cross-validation data sets 165

Figure 3.4 Performance of machine learning models to evaluate all informative gene transcripts (72,893) in hybrid striped bass white muscle (i.e., related to the good- and poor-growth groups) that were ranked by importance during classification using the Machine Learning SVMAttributeEval model. Hybrid striped bass muscle growth status was predicted (classified into good-growth and poor-growth groups) based

on expression of muscle gene transcripts. (a) The Sequential Minimal Optimization Support Vector Machines (SMO) model and (b) The Multilayer Perceptron Artificial Neural Networks (MLP) model were used. Data points were fit with a polynomial line of the third order. The graphs show model robustness performance of the cross-validation as percentage of correctly classified instances when using the top highly ranked genes (10, 25, 50, 75, 100, 150, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, and 70000). The MLP model could not be performed with 1100 or more genes due to limitations in computational capacity. Two cross-validation strategies were used to evaluate the model learning: (1) a percentage split whereby 66% of the data were randomly selected and used to train the models and the remaining 33% of the data were input as a cross-validation and (2) a 8-fold stratified hold out with $n = 8$ folds where one fold was used for cross-validation and $n - 1$ folds of the randomly reordered data set were used for training. Both classes (good- and poor-growth groups) were properly represented in the model training and cross-validation data sets..... 167

Figure 3.5 Performance of machine learning models to evaluate all informative gene transcripts (72,893) in hybrid striped bass white muscle (i.e., related to the good- and poor-growth groups) that were ranked by importance during classification using the Machine Learning

SVMAttributeEval model. Hybrid striped bass growth status (classified into good-growth and poor-growth groups) was predicted based on expression of muscle gene transcripts using the Sequential Minimal Optimization Support Vector Machines (SMO) model and the Multilayer Perceptron Artificial Neural Networks (MLP) model. Data points were fit with a polynomial line of the third order. The graph shows the top 10,000 genes that were used in the SMO and MLP, if applicable, model and subsequently the top 10, 25, 50, 75, 100, 150, 250, and 500 genes that were eliminated from the list. The curves are the SMO and MLP model performance evaluation in what is called "reduction of data dimensionality". The MLP model could not be performed with 1100 or more genes due to limitations in computational capacity. Two cross-validation strategies were used to evaluate the model learning: (1) a percentage split whereby 66% of the data were randomly selected and used to train the models and the remaining 33% of the data were input as a cross-validation and (2) a 8-fold stratified hold out with $n = 8$ folds where one fold was used for cross-validation and $n - 1$ folds of the randomly reordered data set were used for training. Both classes (good- and poor-growth groups) were properly represented in the model training and cross-validation data sets 168

Chapter IV: A Correlation Between Gene Expression And Metabolites

Figure 4.1 Bar charts of significantly-enriched Qiagen IPA Canonical Pathways ($p \leq 0.05$) from Core Analysis of GeneMetaboliteListA (ranked by p -value) for hybrid striped bass muscle from the good-growth group. Each bar represents a canonical pathway's p -value on a negative logarithmic scale, such that the taller bars are more significant than the shorter bars. Bars are shaded according to their z-score activity predictions. Orange: pathways with positive z-scores (upregulated), blue: pathways with negative z-scores (down regulated), white: pathways that have z-score of 0, indicating that the differential gene expression data did not allow for a clear determination of the activity prediction; in other words, the weight of the evidence for a prediction of activation is equal to that of inhibition, and grey: pathways for which no activity predictions can currently be made due to a lack of information in the Ingenuity Knowledge Base 203

Figure 4.2 Bar charts of significantly-enriched Qiagen IPA Canonical Pathways ($p \leq 0.05$) from core analysis of GeneMetaboliteListB (ranked by the Machine Learning-based SVMAttributeEval) for hybrid striped bass muscle from the good-growth group. Each bar represents a canonical pathway's p -value on a negative logarithmic scale, such that the taller bars are more significant than the shorter bars. Bars are shaded according to their z-score activity predictions. Orange: pathways with positive z-scores (upregulated), blue: pathways with negative z-scores (downregulated), white: pathways that have z-score of 0, indicating that

the differential gene expression data did not allow for a clear determination of the activity prediction; in other words, the weight of the evidence for a prediction of activation is equal to that of inhibition, and grey: pathways for which no activity predictions can currently be made due to a lack of information in the Ingenuity Knowledge Base 204

Figure 4.3 Cellular senescence pathway enriched in GeneMetaboliteListA (ranked by *p*-value) for hybrid striped bass muscle from the different growth groups. The overlapping molecules from the gene list (green for downregulated and red for upregulated) were used to predict the activity pattern of the pathway using Molecule Activity Predictor (MAP) in IPA. It predicted gene expression which probably leads to inhibition of cellular senescence (blue) and activation of proliferation of cells (orange). Downregulation (green) refers to lower expression in fish muscle from the good-growth group relative to fish muscle from the poor-growth group, while upregulation (red) refers to increased expression in fish muscle from the good-growth group relative to fish muscle from the poor-growth group 205

Figure 4.4 B Cell receptor signaling pathway enriched in GeneMetaboliteListB (ranked by the Machine Learning-based SVMAttributeEval) for hybrid striped bass muscle from the different growth groups. The shared molecules from the gene list (green for downregulated and red for upregulated) were used to predict the activity pattern of the pathway using Molecule Activity Predictor (MAP) in IPA. It predicted gene

expression which probably leads to inhibition of apoptosis (blue) and activation of transcription (orange). Downregulation (green) refers to lower expression in fish muscle from the good-growth group relative to fish muscle from the poor-growth group, while upregulation (red) refers to increased expression in fish muscle from the good-growth group relative to fish muscle from the poor-growth group..... 206

Figure 4.5 Upstream Regulator analysis of GeneMetaboliteListA (ranked by *p*-value) identified a number of factors (*il6*, *il1b*, *stat3*, and *pparg*) that were predicted to be inhibited based on gene expression in muscle of hybrid striped bass. The network indicates direct relationships from five inhibited factors (colored blue due to predicted inhibition in fish from the good-growth group). Connections are drawn between molecules that have been found to have relationships in the literature. Arrows indicate activation and perpendicular lines indicate inhibition. Genes and metabolites from the input dataset with a relationship to the Upstream Regulators are colored by fold change (red: upregulated in fish muscle from the good-growth group; green; downregulated in fish muscle from the good-growth group). The pathway also includes curated functions (mass of organism, apoptosis, and inflammatory response) that are associated with the molecules. Based on the signature of input genes, apoptosis and inflammatory response were predicted to be inhibited (blue) in fish muscle from the good-growth group relative to fish from

the poor-growth group. Conversely, mass of organism was predicted to be activated (orange)..... 207

Figure 4.6 Upstream Regulator Analysis of GeneMetaboliteListB (ranked by the Machine Learning-based SVMAttributeEval) identified a number of developmental regulators (*myc*, *sirt1*, *egfr*, *notch1*) that were predicted to be activated based on the gene signature provided from hybrid striped bass muscle. Network map indicates the direct relationships of these factors (colored orange due to predicted activation in fish from the good growth group). Genes and metabolites from the input dataset with a relationship to the Upstream Regulators are colored by fold change (red: upregulated in fish muscle from the good-growth group; green; downregulated in fish muscle from the good-growth group). The pathway includes curated functions (hyperplasia of cell lines; proliferation of fibroblasts, size of body) and that are associated with the molecules. Based on the signature of input genes and the relationships with the upstream regulators, each of these functions were predicted to be activated (orange)..... 208

Figure 4.7 Comparison analysis of IPA Upstream regulators of muscle from hybrid striped bass in the good-growth group. Heatmaps were generated to show the key Upstream Regulators that were identified in Core Analyses from each gene list (order of columns from left to right: Upstream Regulators, GeneMetaboliteListA, GeneListA, GeneMetaboliteListB, GeneListB, MetaboliteList). Blue squares indicate predicted inhibition of

the upstream regulator or function, while orange indicates the component is predicted to be activated based on the dataset provided. Intensity of color is scaled by z-score, with legend provided above heatmap 209

Figure 4.8 Comparison Analysis of IPA Diseases and Biological Functions of muscle from hybrid stiped bass in the good-growth group. Heatmaps were generated to show the key Diseases and Functions that were identified in Core Analyses from each gene list (order of columns from left to right: Disease and Biological Functions, GeneMetaboliteListA, GeneListA, GeneMetaboliteListB, GeneListB, MetaboliteList). Blue squares indicate predicted inhibition of the upstream regulator or function, while orange indicates the component is predicted to be activated based on the dataset provided. Intensity of color is scaled by z-score, with legend provided above heatmap..... 210

Figure 4.9 One metabolite and 12 genes shown to most influence the growth performance hybrid striped bass (mass of organism) in GeneMetaboliteListA (ranked by *p*-value). Table 4.1 lists all molecules that had a direct effect on the growth performance in fish muscle from the good-growth group. Network map indicates the direct relationships of these factors (colored orange due to predicted activation in fish from the good growth group). Genes and metabolites from the input dataset with a relationship to the Upstream Regulators are colored by fold change (red: upregulated in fish muscle from the good-growth group;

green; downregulated in fish muscle from the good-growth group). The filtered functions based on the signature of input genes and metabolites were predicted to be inhibited (blue) or activated (orange) in fish muscle from the good-growth group relative to fish muscle from the poor-growth group. Connections are drawn between molecules that have been found to have relationships in the literature. Arrows indicate activation, and perpendicular lines indicate inhibition. Orange and blue lines indicate agreement and yellow line indicates disagreement with previously published literature 211

Figure 4.10 BioProfiler analysis of GeneMetaboliteListA (ranked by *p*-value) of hybrid striped bass. Network was generated from filtered results for muscle and liver interactions. Genes and metabolites from the input dataset are colored by fold change (fish muscle from the good-growth group/fish muscle from the poor-growth group; red: upregulated in fish muscle from the good-growth group; green: downregulated in fish muscle from the good-growth group). The filtered functions based on the signature of input genes were predicted to be inhibited (blue) or activated (orange) in fish muscle from the good-growth group relative to fish muscle from the poor-growth group. Connections are drawn between molecules that have been found to have relationships in the literature. Arrows indicate activation, and perpendicular lines indicate inhibition. Legend is provided indicating what each color represents.

Orange and blue lines indicate agreement and yellow line indicates disagreement 212

Figure 4.11 BioProfiler analysis of GeneMetaboliteListB (ranked by the Machine Learning-based SVMAttributeEval) from hybrid striped bass. Network was generated from filtered results for muscle and liver interactions. Genes and metabolites from the input dataset are colored by fold change (fish muscle from the good-growth group/fish muscle from the poor-growth group; red: upregulated in fish muscle from the good-growth group; green: downregulated in fish muscle from the good-growth group). The filtered functions based on the signature of input genes were predicted to be inhibited (blue) or activated (orange) in fish muscle from the good-growth group relative to the fish muscle from the poor-growth group. Connections are drawn between molecules that have been found to have relationships in the literature. Arrows indicate activation, and perpendicular lines indicate inhibition. Legend is provided indicating what each color represents. Orange and blue lines indicate agreement and yellow line indicates disagreement..... 213

Figure 4.12 Similar to Figure 4.11, but represented as Organic View in IPA. BioProfiler analysis of GeneMetaboliteListB (ranked by the Machine Learning-based SVMAttributeEval) from hybrid striped bass. Network was generated from filtered results for muscle and liver interactions. Genes and metabolites from the input dataset are colored by fold change (fish muscle from the good-growth group/fish muscle from the poor-

growth group; red: upregulated in fish muscle from the good-growth group; green: downregulated in fish muscle from the good-growth group). The filtered functions based on the signature of input genes were predicted to be inhibited (blue) or activated (orange) in fish muscle from the good-growth group relative to the fish muscle from the poor-growth group. Connections are drawn between molecules that have been found to have relationships in the literature. Arrows indicate activation, and perpendicular lines indicate inhibition. Legend is provided indicating what each color represents. Orange and blue lines indicate agreement and yellow line indicates disagreement in the published literature..... 214

Chapter V: The Effects of Frequent and Infrequent Feeding Regimes on the Growth of Sunshine Hybrid Striped Bass (Male Striped Bass *Morone saxatilis* x Female White Bass *Morone chrysops*)

Figure 5.1 Bar charts showing the average (a) initial and (b) final weights (g/fish) of sunshine bass in the feeding trials. The fish were fed either by hand or automated belt feeder three or five days per week (Hand x 3, Hand x 5, Belt x 3, and Belt x 5). Average values are represented by each bar and the SE of the mean is shown by the brackets. There were three tank replicates per experimental group. There were no significant differences between groups (ANOVA, LSMeans Student’s t-test or LSMeans Difference Tukey HSD, alpha = 0.05) 236

Figure 5.2 Bar charts showing the average (a) weight gain per day (g/fish), (b) total weight fed per fish, and (c) feed conversion ratio (FCR) of sunshine bass

hybrid striped bass in the feeding trials. The fish were fed either by hand or automated belt feeder three or five days per week (Hand x 3, Hand x 5, Belt x 3, and Belt x 5). Average values are represented by each bar and the *SE* of the mean is shown by the brackets. There were three tank replicates per experimental group. Data points with the same letter superscript indicate mean values that are not significantly different at $\alpha = 0.05$ (ANOVA, LSMeans Student's t-test or LSMeans Difference Tukey HSD). There were no significant differences between groups in average weight gain per day (g/fish) or average feed conversion ratio (FCR) (ANOVA, LSMeans Student's t-test, $\alpha = 0.05$). There was a significant difference between Hand x 3 and Belt x 5 in average total weight fed per fish

(LSMeans Difference Tukey HSD, $p \leq 0.05$) 237

Figure 5.3 Line graph showing the average Initial feeding rates (Initial FR) and Final feeding rates (Final FR) of sunshine hybrid striped bass expressed as percent body weight (%BW) for all treatment groups. The fish were fed either by hand or automated belt feeder three or five days per week (Hand x 3, Belt x 3, Hand x 5, and Belt x 5). The values represent the mean of three replicate tanks. Data points with the same letter superscript indicate mean values that are not significantly different (ANOVA, LSMeans Student's t-test or LSMeans Difference Tukey HSD, $p \leq 0.05$). The slopes of the lines representing change in feeding rate also are provided 239

Figure 5.4 Line graph showing the average Initial and Final body weights (g/fish) of sunshine hybrid striped bass for all treatment groups. The values

represent the mean of three replicate tanks. The fish were fed either by hand or automated belt feeder three or five days per week (Hand x 3, Belt x 3, Hand x 5, and Belt x 5). There were no significant differences between groups in Initial or Final body weight (ANOVA, LSMeans Student's t-test or LSMeans Difference Tukey HSD, alpha = 0.05). The slopes of the lines representing changes in body weight (i.e., growth) also are provided 240

Chapter VI: Future Directions

Figure 6.1 Genomic selection within an aquaculture breeding program. Full siblings from a number of families are split into selection candidates and animals for phenotypic evaluation. These full siblings of the selection candidates can be grown in different environmental conditions and phenotyped for different traits, for example, to measure growth performance traits in ponds and tanks. The selection candidates and their phenotyped full siblings are all genotyped, and a genomic relationship matrix reflecting the genetic similarity between each pair of animals is built. This relationship matrix and the collected phenotypes enable the estimation of breeding values for the selection candidates through the use of genomic selection models such as GBLUP (genomic best linear unbiased prediction) or Bayesian models (Georges et al., 2019). gEBV, genomic estimated breeding value. (from Houston et al., 2020)..... 251

Chapter I: Literature Review

Hybrid striped bass background

The striped bass (*Morone saxatilis*) is an anadromous species that typically lives most of its life in the Atlantic Ocean or estuarine waters and migrates into freshwater rivers to spawn and have long been utilized as food fish and as an excellent sport fish (Harrell and Webster, 1997; Stickney, 1993; Harrell et al., 1990). Taking advantage of market opportunities created by the decline of the wild striped bass fishery in the mid-1980s, a foodfish farming industry began to develop based largely on a hybrid produced by crossing the striped bass with the freshwater white bass, *Morone chrysops* (Harrell and Webster, 1997). Hybridization of striped bass was first accomplished in 1965 in South Carolina by Robert Stevens (Bishop, 1968). The main reason of creating such a hybrid was to produce a new fish for recreational fisheries that had the size, longevity, food habits, and angling qualities of the striped bass and the hardiness of the white bass in different environments, particularly warm freshwater (Bishop, 1968; D'Abramo and Frinsko, 2008; Dasgupta and Thompson, 2013; Harrell and Webster, 1997; Hodson, 1989b; Kelly and Kohler, 1996; NOAA, 2016; Rudacille and Kohler, 2000). Similarly, hybrid striped bass exhibit superior commercial aquaculture production traits compared to that of the white bass and striped bass used to create them (Harrell et al., 1990; Hodson, 1989b). Hybrid striped bass have a greater tolerance to extremes in temperature (from 4 to 33°C) compared to either of its parents and is thus better suited for aquaculture (D'Abramo and Frinsko, 2008), however optimum growth occurs within a temperature range of 25 to 27°C (Hodson, 1989b; Woiwode and Adelman, 1983). The name "hybrid striped bass" is the common name that refers to either the original cross between a male white bass (Figure

1.1B) and a female striped bass (Figure 1.1A), or the reciprocal cross between a male striped bass and a female white bass. The original cross, is also known as the “palmetto bass”, which was first produced in South Carolina in the mid-1960s (Hodson, 1989b). The reciprocal cross, also known as the “sunshine bass” (Figure 1.1C) are the predominant market product, because white bass females are more available and are easy to spawn in captivity (Lochmann, 2015). Sunshine bass grow faster and have higher survival rates than striped bass, white bass or palmetto bass when reared in indoor water recirculation systems (Kelly and Kohler, 1996; Rudacille and Kohler, 2000). Also, sunshine bass were found in many studies to grow more quickly and convert feed more efficiently than striped bass, white bass and palmetto bass during part of the production cycle (D’Abramo and Frinsko, 2008; Dasgupta and Thompson, 2013; Lochmann, 2015; Rudacille and Kohler, 2000). Despite the performance differences, both reciprocal hybrid crosses exhibit hybrid vigor (heterosis), manifested as improved survival and growth during the first two years, and resistance to handling stress and diseases (Kerby, 1993; Kerby and Harrell, 1990; Noga et al., 1994). Thus, they gained widespread acceptance either as a sportfish for stock enhancement or as an aquaculture species for food (Dasgupta and Thompson, 2013; Kerby, 1993).

Henceforth, hybrid striped bass production has dramatically grown across the United States (U.S.) (Hodson, 1989b; Rudacille and Kohler, 2000; Whitehurst et al., 1990). Farming of hybrid striped bass is currently the fourth largest form of finfish aquaculture in the U.S., behind only catfish, salmon and tilapia. The annual national production of hybrid striped bass was estimated around 12 million pounds with a farm-gate value of

approximately \$50 million USD (USDA 2019). The industry has been expanding at a rate of 10-15% per year for the past 20 years (Dasgupta and Thompson, 2013).

Despite the size and importance of hybrid striped bass aquaculture in the U.S., growth of the industry is hampered in part by increasing production costs. As with other forms of finfish aquaculture, the rising costs of feeds is a primary concern for industry expansion and feed typically represents 40–50% of the total production costs in aquaculture (Lochmann, 2015). The Striped Bass Growers Association polls their members annually and feed cost is consistently their number one concern among producers nationwide (Gatlin et al., 2007). There are several different strategies for addressing this problem, including: 1) reducing fishmeal or animal-derived protein components of the feeds, which tend to be expensive, 2) improving the feed conversion efficiency of the animals such that less feed is required to grow them to market size, 3) improving the growth rate of the animals using selective breeding, biotechnology, or modification to husbandry practices. In this study we will focus on points 2) and 3) above, in particular to better understand the molecular basis of growth differences in hybrid striped bass in order to identify important gene and metabolic pathways that could be targeted for future improvement by selective breeding and to explore the use of reduced frequency feeding strategies that may improve feed conversion efficiency in the fish.

An overview of muscle growth in fishes

The primary consumable product of fish is the fillet, composed of white, fast-twitch muscle fibers (Johnston, 1999). The white muscle accounts for more than half of the body mass of fishes (Weatherley 1990; Weatherley et al., 1988) and therefore understanding the

regulatory mechanisms underlying the growth of this tissue is critical for improving aquaculture yields (Johnston, 1999). Muscle growth in fishes also is different from that of other vertebrates in that it is continuous (Johnston et al., 2011, Stickland, 1988; Valente et al., 2013). Many studies have suggested (Fauconneau et al., 1993; Hatae et al., 1990; Hurling et al., 1996) that variation in quality of the fish fillet is largely due to the variation in the size distribution of muscle fibers. In aquaculture, the number and size distribution of fibers, therefore, is thought to be an important determinant of not only growth performance, but also of the textural characteristics of the flesh (Fauconneau et al., 1993; Hurling et al., 1996).

Morphometric procedures, including average cross-sectional area or diameter of muscle fibers, enables the direct visualization and analysis of muscle fiber distribution patterns, such as, flesh texture and morphology (Dunajski, 1979). No previous literature was found, which has associated the difference in growth performance of the white muscle fibers of hybrid striped bass based on a single morphometric measure. However, substantial progress has taken place in the aquaculture industry in an attempt to understand the relationship between white muscle fiber growth and the whole body growth capacity of various fish species (Mommsen, 2001; Suresh and Sheehan, 1998; Velez et al., 2016; Weatherley et al., 1988). Such as the study by Weatherley et al. (1988), which was conducted on ten different freshwater fish species (Table 1.1) with the intent of understanding the ability of different teleost fishes to grow rapidly and reach a comparatively large ultimate size only by recruiting new white muscle fibers. The results showed two major findings that later came into use for the aquaculture industry. First, while the majority of species used in the study revealed rapid somatic growth rates (i.e.,

fork length) synchronized with dynamic increase of white muscle fiber number (i.e., hyperplasia), a few revealed less progressive somatic growth rate synchronized with an increase in white muscle fiber diameter (i.e., hypertrophy). This suggests that hyperplasia, in particular, plays a major role in fish muscle growth at early ages, wherein the longer a species is capable of recruiting new white muscle fibers, the faster is its growth rate and the larger is its ultimate size (Johnston, 1999; Rowleson and Veggetti, 2001; Suresh and Sheehan, 1998). The second major finding of Weatherley et al. (1988) was that, the proportional distribution of muscle fibers of different sizes was strongly correlated with body size, regardless of different rates of organismic growth. Gill et al. (1989) postulated that this relationship between muscle fiber size frequency and body size is a species-specific physiological acclimating effect to ensure muscle tissues meeting the structural and functional requirements imposed by a particular body size. On the account of these two findings, similar muscle growth dynamics may be found in hybrid striped bass white muscle samples, where a progressive hypertrophic muscle growth may be perceived in both larger and smaller groups, while the larger sized group (i.e., good growth) may generally show greater hyperplasia than that of the smaller sized group (i.e., poor growth).

Fish muscle morphometrics have important implications for the end-product fillet quality as they influence textural characteristics of the consumable product (Johnston et al., 2000a; Listrat, et al., 2016). The majority of studies on fish muscle growth have measured the average cross-sectional areas or diameters of fibers (Gill et al., 1989; Johnston, 1999; Suresh and Sheehan, 1998; Weatherley et al., 1988), as average fiber size is a relatively unreliable indicator of hypertrophic growth, because of the recruitment of new muscle fibers (Johnston, 1999). Stickland (1988) conducted a study on the rainbow trout

(*Oncorhynchus mykiss*) and noticed that the average fiber diameter in the white muscle remained within the range 90-95 μm between 34 and 52 cm body length due to the addition of new fibers, but increased to 135-140 μm at 62 cm body length once hyperplasia had ceased and hypertrophy of the muscle had begun. Another consideration of using fish muscle morphometrics, is that the presence of small fibers does not necessarily indicate fast growth because they are typical of fish size rather than growth rate (Weatherley and Gill, 1982, 1987a; Weatherley et al., 1988). For example, when the grass pickerel (*Esox americanus*) in Weatherley and Gill (1987b) was categorized as a slow-growing species in terms of the ability for rapid large and ultimate somatic size growth, the white muscle primarily contained small diameter fibers, which is a character of some fast-growing fish species (Weatherly and Gill, 1987b). Therefore, muscle fiber number and diameter must be considered when using muscle morphometrics as an indicator of growth status.

Additional factors, such as small molecules (i.e., metabolites) and expressed gene transcripts also may serve as good indicators of growth status in fishes. From the molecular level, the main expansion of muscle fibers with growth occurs by a process called mosaic hyperplasia, in which myogenic progenitor cells (i.e., satellite cells) fuse to form new myotubes on the surface of existing muscle fibers giving rise to a mosaic of fiber diameters as the fish matures (Johnston et al., 2009; Rowleson et al., 1995). Satellite cells are elongated cells with limited cytoplasm located between the plasma membrane of the muscle fiber and the basement membrane (Figure 1.2) (Mauro, 1961). Their main function, after generating muscle fibers, is to maintain the homeostasis of muscle cell growth through giving rise to new satellite cells or by differentiating into muscle fibers as necessary for growth (Brack, 2012; Kadi et al., 2004; Kadi et al., 2005; Pannérec. 2013).

Satellite cells also contribute additional nuclei to the muscle fiber as it expands in length and diameter (Johnston et al., 2003). In the majority of teleosts, mosaic hyperplasia in muscle continues until the fish reaches around 40% of its maximum body length (Johnston et al., 2009; Rowleson et al., 1995). This is a complex process that involves many different regulatory genes, enzymes involved in metabolic processes such as energy production, and structural components of the muscle itself. Presently, the molecular mechanism(s) of white muscle growth in hybrid striped bass (i.e., hyperplasia *versus* hypertrophy) and the molecular dynamics of muscle growth, including the importance of small molecules (i.e. metabolites) and gene transcripts involved in the muscling process, are largely unknown. Therefore, further characterization of the processes that underly hybrid striped bass muscle growth is needed to expand the application of scientific findings to aquaculture production of these fish.

An overview of reduced frequency feeding strategies in aquaculture

There are two general strategies used to reduce feed costs without interfering with growth rates of fish in aquaculture: 1) formulate feeds that are comprised of less costly ingredients which provide an equivalent nutritional value and 2) optimize feeding and husbandry practices to maximize the feed conversion efficiency (FCE) of the animals. FCE is the kg or pounds of food required to increase the biomass in the fish by one kg or pound, respectively and it is a primary economic concern of fish farmers ($FCE = (\text{weight gained} / \text{weight of feed applied}) \times 100$) (D'Abramo and Frinsko, 2008). Many studies have been conducted to improve feed formulation to correspond with seasonal water temperature fluctuations without interference in the growth performance of the cultured hybrid striped

bass or the water quality (D'Abramo and Frinsko, 2008; Ludwig, 2004; Rawles et al., 2012). Since water temperatures influence feed intake and growth rates, these feeds have been formulated based on cost, ingredients, and most importantly nutrient composition (protein: energy ratio, fat). For example, some producers feed nutrient-dense (high-protein, high-fat) diets during mid-winter and mid-summer to reduce feed waste and nutrient stacking in ponds (D'Abramo and Frinsko, 2008; Rawles et al., 2012). These studies aimed to optimize and evaluate less expensive feed formulations that provide necessary energy and essential nutrients whilst reducing husbandry practice costs to maximize the FCE of hybrid striped bass.

Fewer studies have investigated the response of hybrid striped bass growth performance to changes in feeding frequency (Liu and Liao, 1999; Webster et al., 2001; Rawles et al., 2012). Rawles et al. (2012) recommended the best way to reduce the feed cost without affecting the growth performance of the hybrid striped bass seems to be in reducing feeding when water temperatures were not within the optimum growth range. Other studies took advantage of compensatory growth or "catch-up-growth", which is a physiological process whereby the organism exhibits accelerated growth after a period of restricted feed intake, to reduce feed cost whilst maintaining the growth rate efficiency and at the same time minimizing environmental impacts. Turano et al. (2007) conducted a 16 week pond study to evaluate the effectiveness of cyclic feeding regimens to elicit a compensatory growth response and mitigate water quality problems in fingerling hybrid striped bass. To evaluate the potential for compensatory growth, hybrid striped bass were subjected to one of four cyclic feeding regimens of control (0 week, fed twice daily), cycles of one week, two weeks, and 4 weeks deprivation, followed by one, two or four weeks

feeding. The results revealed a higher FCE (i.e. the inverse of FCR or feed conversion ratio) for fish subjected the two weeks regimen compared to the other regimens, although the fish were smaller in size at the end of the trials. These results put forward that compensatory growth can be effectively induced in pond-raised hybrid striped bass as a way of minimizing the feed and the husbandry costs at the same time as maintaining the water quality. An alternative to these procedures would be to feed fish on alternate days as opposed to every day in a reduced frequency feeding schedule, which may reflect the natural feeding behavior of the animals and help to mitigate the overall lower growth rate due to feed restriction as observed for compensatory growth. As hybrid striped bass are piscivores, they naturally feed infrequently as the gut is filled with large prey items which are digested and then waste is eliminated. These fish do not typically feed several times a day in nature, as is common practice in aquaculture. Thus, reduced frequency feeding schedules may offer an alternative for improving the FCR and reducing labor costs on hybrid striped bass farms.

Mechanisms underlying feed utilization and growth in hybrid striped bass are still poorly understood and no one has yet to evaluate the underlying metabolomics and transcriptomics processes that may be affecting hybrid striped bass growth performance. Based on these previous studies, we hypothesize that 1) differences in growth rates of hybrid striped bass will be associated with changes in the muscle and liver metabolites that are related to growth metabolic pathways; 2) differences in growth rates of hybrid striped bass will be associated with changes in the muscle transcriptome of genes that are related to growth pathways; 3) important metabolites and expressed genes may serve as indicators of current or future growth status of hybrid striped bass; and 4) it is possible to

improve the FCR of hybrid striped bass using reduced frequency feeding strategies, which in turn will reduce the overall husbandry costs.

Objectives

In order to facilitate a better understanding of the metabolome, transcriptome, and their pathways, particularly with respect to their effects on the white muscle growth performance of hybrid striped bass, and to improve the FCR of farmed fish, I proposed the following specific research objectives:

- 1) Determine which muscle and liver metabolites that are involved in enhanced growth responses of hybrid striped bass.
- 2) Determine the muscle genes that are involved in enhanced growth responses of hybrid striped bass.
- 3) Determine if there is a correlation between gene expression and metabolites.
- 4) Determine the effects of frequent and infrequent feeding regimes on the growth of hybrid striped bass.

Study design

We conducted a series of studies on hybrid striped bass that grow well and those that grow poorly in a common garden experimental design. The fish were reared communally at semi-commercial density in tanks and earthen ponds until the average size of the cohort was market-sized (approximately 1.25-1.5 lbs.). Fish were then harvested and a weight-frequency histogram generated to identify those fish with good and poor growth performance (i.e., top and bottom 10% of the distribution, respectively). To provide a

visual indicator for the underlying growth dynamics of the two most important processes of muscle growth (hyperplasia and hypertrophy), morphometric differences between muscle samples taken from each hybrid striped bass sampled from the good and poor growth groups were evaluated by histology. Tissues from these animals also were sampled and analyzed to identify the important metabolites and gene transcripts related to differences in growth performance. As metabolites represent the end products of gene and protein activities within the cell (and hence the organism), gene expression also will be measured and will be complementary to the metabolite data. For example, understanding metabolic pathways may be enhanced when complementary gene transcript data are available to support a physiological response (Gatlin, 2007; Goodacre et al., 2004), although gene expression data does not account for post-transcriptional regulation or post-translational modifications (i.e., identifying a particular gene transcript does not necessarily indicate that it is translated into a functional protein). Shared pathways identified between the gene expression and metabolite datasets were compiled and compared to better understand growth performance in hybrid striped bass. Moreover, the application of genomic technologies promises to revolutionize our understanding of the genetic and molecular basis of muscle growth and plasticity in farmed fish species, thereby increasing the efficiency and sustainability of aquaculture production (Estévez, 2012; Vélez et al, 2016).

Lastly, an experiment was conducted whereby the growth performance of hybrid striped bass was measured and compared between groups fed at a different frequency (three or five times per week for three months) in order to understand how feeding frequency may influence the FCR (FCR = weight of feed applied/ weight of fish gained) of

hybrid striped bass. Our approach to finding the optimum FCR under these feeding conditions may not only improve the growth rate of the hybrid striped but also may improve the husbandry practices by reducing the labor and feed costs in farms and also may mimic a more natural feeding style for the animals.

Organization of the dissertation:

The dissertation is comprised of six chapters. A literature review, four data chapters that are parallel to the four represented research objectives, and a summary and future directions chapter.

The first chapter, (**Chapter I: Literature Review**), serves as a literature review. The literature review covers the background of hybrid striped bass culture, a general overview of muscle growth in fish and some of the underlying metabolome and transcriptome processes, an overview of reduced frequency of feeding practices in aquaculture in general and specific to hybrid striped bass, as well as an outline of the research objectives and the research design.

The second chapter, (**Chapter II: White Muscle and Liver Metabolites Involved In Enhanced Growth Response of Hybrid Striped Bass**), explains how I achieved the first research objective (Determine which muscle and liver metabolites that are involved in enhanced growth responses of hybrid striped bass). This chapter is comprised of five sections: an abstract, an introduction, methods and methods, results, and discussion. In this chapter, I described the use of a novel metabolite profiling machine learning approach that incorporated data generated from white muscle and liver tissue samples from two distinctive hybrid striped bass growth groups (poor- and good-growth) to evaluate those

small molecule compounds (i.e., metabolites) that are related to growth performance. The major finding of this study indicated that metabolite pathways related to inflammation were elevated in fish from the poor-growth group performance.

The third chapter, (**Chapter III: White Muscle Genes Involved In Enhanced Growth Response of Hybrid Striped Bass**), explains my approach to the second research objective (Determine the muscle genes that are involved in enhanced growth responses of hybrid striped bass). This chapter consists of the following sections: an abstract, an introduction, material and methods, results, and a very brief discussion. In this chapter, I describe how a machine learning platform was used to analyze gene expression data generated from fish in both growth performance groups described above. The resulting list of genes determined to be relevant to growth provided compelling evidence of an immune response possibly generated in response to a stressor, which may have affected growth performance in the poor-growth fish group.

The lists of metabolites provided in chapter II and genes provided in chapter III however, do not necessarily provide comprehensive information as far as understanding the physiological phenomenon that underlies the growth performance differences, therefore, I performed a novel combinatorial analysis described in chapter IV. The fourth chapter, (**Chapter IV: A Correlation Between Gene Expression And Metabolites**), elucidated the third research objective (Determine if there is a correlation between gene expression and metabolites). This chapter contains: an abstract, an introduction, material and methods, and a collective results and discussion section. In this chapter, I looked at how the combined transcriptome and metabolome data determined in chapters II and III regulate the growth performance in hybrid striped bass. I used the innovative Canonical

Pathway Analysis provided by Qiagen IPA pathways. The chief verdict of the combined analyses provide potential candidates and pathways such as cell proliferation and differentiation, cell death, and inflammation that may possibly underly the growth performance of the fish in general.

The fifth chapter, (**Chapter V: The Effects of Frequent and Infrequent Feeding Regimes on the Growth of Sunshine Hybrid Striped Bass (Male Striped Bass *Morone saxatilis* x Female White Bass *Morone chrysops*)**), fulfills the final research objective (Determine the effects of frequent and infrequent feeding regimes on the growth of hybrid striped bass). The chapter is comprised of: an abstract, an introduction, material and methods, results, discussion and conclusion. To improve the growth performance of hybrid striped bass, different feeding frequency experiments were conducted in an attempt to improve the FCR value. The key inference is that FCR was approximately 20% higher in fish fed more frequently compared to those fed less frequently, and this may have an economic impact in commercial aquaculture operations.

The sixth and the final chapter of the dissertation, (**Chapter VI: Summary and Future Directions**), summarizes and discusses the understanding of the cellular processes that underlie growth performance of hybrid striped bass based on the studies outlined in Chapters II through V. It reviews the novel combinatorial machine learning analysis of metabolomic data (from **Objective 1, Chapter II**), that were generated from white muscle tissue and liver samples, and transcriptomic data, which were expressed in muscle tissue samples (from **Objective 2, Chapter III**) to identify discrete molecular processes that underlie the regulation of enhanced growth rate and accretion of hybrid striped bass. Together with the feeding study which is conducted to establish the effects of feeding

frequency on growth performance, measured by FCR (from **Objective 4, Chapter V**). This chapter, as a whole, provides the main findings of each research objective and the recommendations on how to extend the current work for future research benefit.

References

AQUASOL fish farming consultants. (2016). Retrieved December 17, 2019, from

<https://www.fishfarming.com/projects/hybrid-striped-bass.html>

Bishop, R. D. (1968). Evaluation of the striped bass (*Roccus saxatilis*) and white bass (*R. chrysops*) hybrids after two years. *Proceedings of the Annual Conference of the Southeastern Association of Game and Fish Commissioners*. 2(1), 245-254.

Brack AS and Rando TA. (2012) Tissue-specific stem cells: lessons from the skeletal muscle satellite cell. *Cell Stem Cell*. 10: 504-514.

D'Abramo, L.R. and M.O. Frinsko. (2008). Hybrid Striped Bass: Pond Production of Food Fish. Southern Regional Aquaculture Center. Stoneville, Mississippi. Publication Number 303.

Dasgupta, Siddhartha, and Kenneth R. Thompson. (2013). Comparison of coast of different Hybrid Striped Bass production systems in ponds. *Southern Regional Aquaculture Center. Publication Number 3000*.

Dunajski, E. (1979). Texture of fish muscle. *Journal of Texture Studies*, 10(4), 301-318.

Estévez, A., Andree, K., & Johnston, I. A. (2012). Fast skeletal muscle transcriptome of the Gilthead sea bream (*Sparus aurata*) determined by next generation sequencing. *BMC genomics*, 13(1), 1.

Fauconneau, B., Chmaitily, J., Andre, S., Cardinal, M., Cornet, J., Vallet, J.L., Dumont, J.P. and Laroche, M. (1993). Characteristics of rainbow trout flesh. 1. Chemical composition and cellularity of muscle and adipose tissues. *Sciences des Aliments (France)*.

- Gatlin, Delbert M., Frederic T. Barrows, Paul Brown, Konrad Dabrowski, T. Gibson Gaylord, Ronald W. Hardy, Eliot Herman, Gongshe Hu, Åshild Krogdahl, Richard Nelson, Kenneth Overturf, Michael Rust, Wendy Sealey, Denise Skonberg, Edward J Souza, David Stone, Rich Wilson, and Eve Wurtele.** (2007). Expanding the utilization of sustainable plant products in aquafeeds: a review. *Aquaculture research*. 38(6), 551-579.
- Gill, H. S., Weatherley, A. H., Lee, R., & Legere, D.** (1989). Histochemical characterization of myotomal muscle of five teleost species. *Journal of fish biology*, 34(3), 375-386.
- Goodacre, R., Vaidyanathan, S., Dunn, W. B., Harrigan, G. G., & Kell, D. B.** (2004). Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends in biotechnology*, 22(5), 245-252.
- Harrell, R. M., Kerby, J. H., and Minton, R. V.** (1990). *Culture and propagation of striped bass and its hybrids*. Striped Bass Committee, Southern Division, American Fisheries Society.
- Harrell, Reginal M., and Donald W. Webster.** (1997). An overview of *Morone* culture. *Developments in Aquaculture and Fisheries Science*. 3, 1-10.
- Hatae, K., Yoshimatsu, F., & Matsumoto, J. J.** (1990). Role of muscle fibers in contributing firmness of cooked fish. *Journal of food science*. 55(3), 693-696.
- Hawke, Thomas J., and Daniel J. Garry** (2001). Myogenic satellite cells: physiology to molecular biology. *Journal of Applied Physiology*. 91(2): 534-551.
- Hodson, R. G. and M. Hayes.** (1989b). Hybrid Striped Bass pond production of foodfish. *Southern Regional Aquaculture Center*. Publication Number 303.

- Hurling, R., Rodell, J. B., & Hunt, H. D.** (1996). Fiber diameter and fish texture. *Journal of Texture Studies*, 27(6), 679-685.
- Johnston, I. A.** (1999). Muscle development and growth: potential implications for flesh quality in fish. *Aquaculture*, 177(1), 99-115.
- Johnston, I. A., Bower, N. I., & Macqueen, D. J.** (2011). Growth and the regulation of myotomal muscle mass in teleost fish. *Journal of Experimental Biology*, 214(10), 1617-1628.
- Johnston, I. A., Manthri, S., Alderson, R., Smart, A., Campbell, P., Nickell, D., Billy Robertson, Charles GM Paxton, and & Burt, M. L.** (2003). Freshwater environment affects growth rate and muscle fibre recruitment in seawater stages of Atlantic salmon (*Salmo salar L.*). *Journal of Experimental Biology*, 206(8), 1337-1351.
- Johnston, I.A., Alderson, R., Sandham, C., Dingwall, A., Mitchell, D., Selkirk, C., Nickell, D., Baker, R., Robertson, B., Whyte, D. and Springate, J.** (2000a). Muscle fiber density in relation to the colour and texture of smoked Atlantic salmon (*Salmo salar L.*). *Aquaculture*, 189(3), 335-349.
- Johnston, I.A., Lee, H.T., Macqueen, D.J., Paranthaman, K., Kawashima, C., Anwar, A., Kinghorn, J.R. and Dalmay, T.** (2009). Embryonic temperature affects muscle fibre recruitment in adult zebrafish: genome-wide changes in gene and microRNA expression associated with the transition from hyperplastic to hypertrophic growth phenotypes. *Journal of Experimental Biology*, 212(12), 1781-1793.
- Kadi, F., Charifi, N., Denis, C., Lexell, J., Andersen, J.L., Schjerling, P., Olsen, S. and Kjaer, M.** (2005). The behaviour of satellite cells in response to exercise: what have we learned from human studies?. *Pflügers Archiv*, 451(2), 319-327.

- Kadi, F., Schjerling, P., Andersen, L. L., Charifi, N., Madsen, J. L., Christensen, L. R., & Andersen, J. L.** (2004). The effects of heavy resistance training and detraining on satellite cells in human skeletal muscles. *The Journal of physiology*, 558(3), 1005-1012.
- Kelly, Anita M., and Christopher C. Kohler.** (1996). Sunshine bass performance in ponds, cages, and indoor tanks. *The Progressive fish-culturist*. 58(1), 55-58.
- Kerby, J. H., & Harrell, R. M.** (1990). Hybridization, genetic manipulation, and gene pool conservation of striped bass.
- Kerby, Jerome Howard.** (1993). The hybrid striped bass and its hybrids. *Culture of nonsalmonid freshwater fishes*. Pages 252 – 255 in Robert R. Stickney editor. Boca Raton, FL: CRC Press, Inc.
- Listrat, A., Lebret, B., Louveau, I., Astruc, T., Bonnet, M., Lefaucheur, L., Picard, B. and Bugeon, J.** (2016). How muscle structure and composition influence meat and flesh quality. *The Scientific World Journal*, 2016.
- Liu, F. G., & Liao, I. C.** (1999). Effect of Feeding Regimen on the Food Consumption, Growth, and Body Composition in Hybrid Striped Bass *Morone saxatilis* × *M. chrysops*. *Fisheries science*. 65(4), 513-519.
- Lochmann, Rebecca.** (2015). Feeds and feeding of hybrid striped bass. *Southern Regional Aquaculture Center*. Publication Number 3001.
- Lodish, H., & Zipursky, S. L.** (2001). Molecular cell biology. *Biochemistry and Molecular Biology Education*, 29, 126-133.
- Ludwig, G.M.** (2004). Hybrid striped bass: Fingerling production in ponds. *Southern Regional Aquaculture Center*. Publication Number 302.

Mauro A. (1961) Satellite cell of skeletal muscle fibers". *J Biophys Biochem Cytol.* 9, 493-495.

Mommsen, T. P. (2001). Paradigms of growth in fish. *Comparative biochemistry and physiology part B: Biochemistry and molecular biology.* 129(2), 207-219.

NOAA (National Oceanic and Atmospheric Administration). (December 16, 2004).

Retrieved April 26, 2006, from

<http://www.oceanservice.noaa.gov/topics/navops/mapping/welcome.html>

Noga, E. J., Kerby, J. H., King, W., Aucoin, D. P., & Giesbrecht, F. (1994). Quantitative comparison of the stress response of striped bass (*Morone saxatilis*) and hybrid striped bass (*Morone saxatilis* x *Morone chrysops* and *Morone saxatilis* x *Morone americana*). *American journal of veterinary research.* 55(3), 405-409.

Pannérec A, Formicola L, Besson V, Marazzi G, and Sassoon DA. (2013) Defining skeletal muscle resident progenitors and their cell fate potentials. *Development.* 140, 2879-2891.

Rawles, Steven D., T. Gibson Gaylord, G. Scott Snyder, and Donald W. Freeman. (2012).

The influence of protein and energy density in commercial diets on growth, body composition, and nutrient retention of sunshine bass, *morone chrysops* ♀ x *morone saxatilis* ♂, reared at extreme temperature. *Journal of the World Aquaculture Society.* 41(s2), 165-178.

Rowlerson, A., & Veggetti, A. (2001). 5. Cellular mechanisms of post-embryonic muscle growth in aquaculture species. *Fish physiology,* 18, 103-140.

Rowlerson, A., Mascarello, F., Radaelli, G., & Veggetti, A. (1995). Differentiation and growth of muscle in the fish *Sparus aurata* (L): II. Hyperplastic and hypertrophic

- growth of lateral muscle from hatching to adult. *Journal of Muscle Research & Cell Motility*, 16(3), 223-236.
- Rudacille, J. B., & Kohler, C. C.** (2000). Aquaculture performance comparison of sunshine bass, palmetto bass, and white bass. *North American Journal of Aquaculture*, 62(2), 114-124.
- Stickland, N. C., White, R. N., Mescall, P. E., Crook, A. R., & Thorpe, J. E.** (1988). The effect of temperature on myogenesis in embryonic development of the Atlantic salmon (*Salmo salar* L.). *Anatomy and embryology*, 178(3), 253-257.
- Stickney, Robert R.** (1993). Culture of nonsalmonid freshwater fishes. Boca Raton, FL: CRC Press, Inc.
- Suresh, A. V., & Sheehan, R. J.** (1998). Muscle fibre growth dynamics in diploid and triploid rainbow trout. *Journal of Fish Biology*, 52(3), 570-587.
- Turano, Marc J., Russell J. Borski, and Harry V. Daniels.** (2007). Compensatory growth of pond-reared Hybrid Striped Bass, *Morone chrysops* × *Morone saxatilis*, Fingerlings. *Journal of the World Aquaculture Society*. 38(2), 250-261.
- USDA (United States Department of Agriculture).** (2019). 2017 Census of Agriculture, 2018 Census of Aquaculture, 3(2). Publication AC-17-SS-2
- Valente, L. M., Moutou, K. A., Conceição, L. E., Engrola, S., Fernandes, J. M., & Johnston, I. A.** (2013). What determines growth potential and juvenile quality of farmed fish species?. *Reviews in Aquaculture*, 5(s1), S168-S193.
- Vélez, E. J., Lutfi, E., Azizi, S., Perelló, M., Salmerón, C., Riera-Codina, M., Ibarz, A., Fernández-Borràs, J., Blasco, J., Capilla, E. and Navarro, I** (2016). Understanding fish muscle growth regulation to optimize aquaculture production. *Aquaculture*.

- Weatherley, A. H.** (1990). Approaches to understanding fish growth. *Transactions of the American Fisheries Society*, 119(4), 662-672.
- Weatherley, A. H. and Gill, H. S.** (1987a). The biology of fish. *Academic press*, London.
- Weatherley, A. H., & Gill, H. S.** (1982). Influence of bovine growth hormone on the growth dynamics of mosaic muscle in relation to somatic growth of rainbow trout, *Salmo gairdneri* Richardson. *Journal of Fish Biology*, 20(2), 165-172.
- Weatherley, A. H., & Gill, H. S.** (1987b). Growth increases produced by bovine growth hormone in grass pickerel, *Esox americanus vermiculatus* (Le Sueur), and the underlying dynamics of muscle fiber growth. *Aquaculture*, 65(1), 55-66.
- Weatherley, A. H., Gill, H. S., & Lobo, A. F.** (1988). Recruitment and maximal diameter of axial muscle fibres in teleosts and their relationship to somatic growth and ultimate size. *Journal of Fish Biology*. 33(6), 851-859.
- Webster, Carl D., Kenneth R. Thompson, Ann M. Morgan, Ebony J. Grisby, and Siddhartha Dasgupta.** (2001). Feeding frequency affects growth, not fillet composition, of juvenile sunshine bass *Morone chrysops* × *M. saxatilis* grown in cages. *Journal of the world aquaculture society*. 32(1), 79-88.
- Whitehurst, David K., and Robert E. Stevens.** (1990). History and overview of striped bass culture and management. Culture and propagation of striped bass and its hybrids. *American Fisheries Society, Southern Division, Striped Bass Committee*, Bethesda, Maryland. 1-5.
- Woiwode, J. G., & Adelman, I. R.** (1983). Growth, food conversion efficiency and survival of the hybrid white × striped bass as a function of temperature. *The aquaculture of*

striped bass: proceedings. Maryland Sea Grant Program, Publication UM-SG-MAP-84-01, University of Maryland, College Park, 143-150.

Table 1.1. Growth of white axial muscle (removed from a site dorso-lateral and slightly posterior to the dorsal fin) fibers (measured under the microscope for the number and “diameter” of each fiber) of ten species of freshwater teleosts from five families (*Cyprinidae*, *Centrarchidae*, *Percidae*, *Salmonidae*, *Esocidae*) possessing widely different growth rates and ultimate sizes have been studied. The dynamics of muscle increase (i.e., increase in fiber numbers and/or diameter) appears to determine the ability for rapid somatic growth and large ultimate size in teleosts. The largest and fastest growing species (smallmouth bass, lake whitefish, rainbow trout, muskellunge) show evidence of sustained recruitment of muscle fibers to a large size, in contrast to the smaller and slower growing species (bluntnose minnow, longnose dace, grass pickerel). Pumpkinseed, bluegill, common bully, and yellow perch are all intermediate in fiber growth dynamics, growth, and ultimate size between the smaller and larger species (Weatherley et al., 1988).

No	Species	Family	Muscle Growth		Body Growth (Fork Length)
			Hyperplasia (Increase in muscle fiber number)	Hypertrophy (Increase in muscle fiber diameter)	
1	Longnose dace (<i>Rhinichthys cataractae</i>)	Cyprinidae	Ceases at small body length	Progression of fiber diameter	Slow-growing
2	Bluntnose minnow (<i>Pimephales notatus</i>)		Ceases at small body length	Progression of fiber diameter	Slow-growing
3	Smallmouth bass (<i>Micropterus dolomieu</i>)	Centrarchidae	Sustained recruitment	Progression of fiber diameter	Fast-growing
4	Bluegill (<i>Lepomis macrochirus</i>)		Ceases eventually at larger body length	Progression of fiber diameter	Intermediate- growing
5	Pumpkinseed (<i>Lepomis gibbosus</i>)		Ceases eventually at larger body length	Progression of fiber diameter	Intermediate- growing
6	Yellow perch (<i>Perca flavescens</i>)	Percidae	Ceases eventually at larger body length	Progression of fiber diameter	Intermediate- growing
7	Rainbow trout (<i>Salmo gairdneri</i>)	Salmonidae	Sustained recruitment	Progression of fiber diameter	Fast-growing

Table 1.1 (continued)

No	Species	Family	Muscle Growth		Body Growth (Fork Length)
			Hyperplasia (Increase in muscle fiber number)	Hypertrophy (Increase in muscle fiber diameter)	
8	Lake whitefish <i>(Coregonus clupeaformis)</i>		Sustained recruitment	Progression of fiber diameter	Fast-growing
9	Common bully <i>(Gobiomorphus cotidianus)</i>	Esocidae	Ceases eventually at larger body length	Progression of fiber diameter	Intermediate- growing
10	Muskellunge <i>(Esox masquinongy)</i>		Sustained recruitment	Progression of fiber diameter	Fast-growing
11	Grass pickerel <i>(Esox americanus)</i>		Ceases at small body length	Progression of fiber diameter	Slow-growing

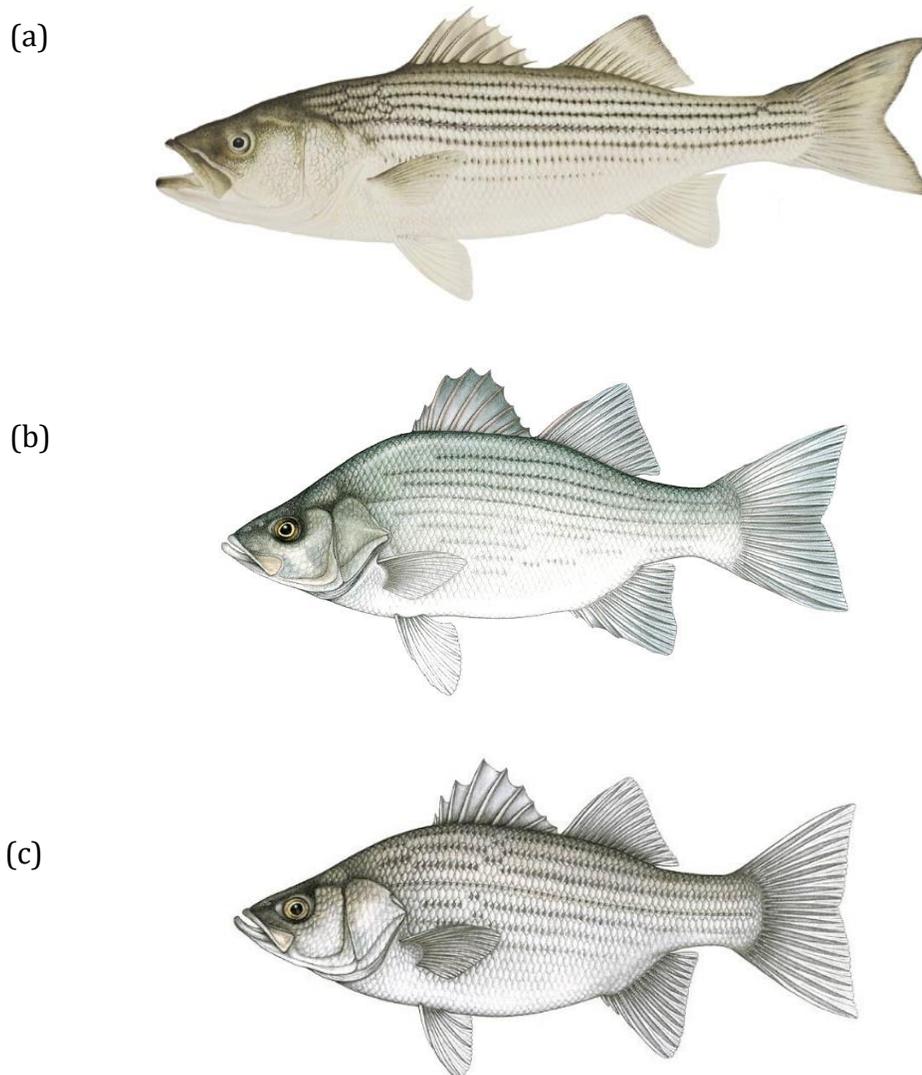


Figure 1.1. (a) Striped bass (*Morone saxatilis*). (b) White bass (*Morone chrysops*). (c) Sunshine cross (♂ Striped bass (*M. saxatilis*) x ♀ White bass (*M. chrysops*)), it can be distinguished from the striped bass by broken rather than solid horizontal stripes on the body (from AquaSol, Inc, 2016).

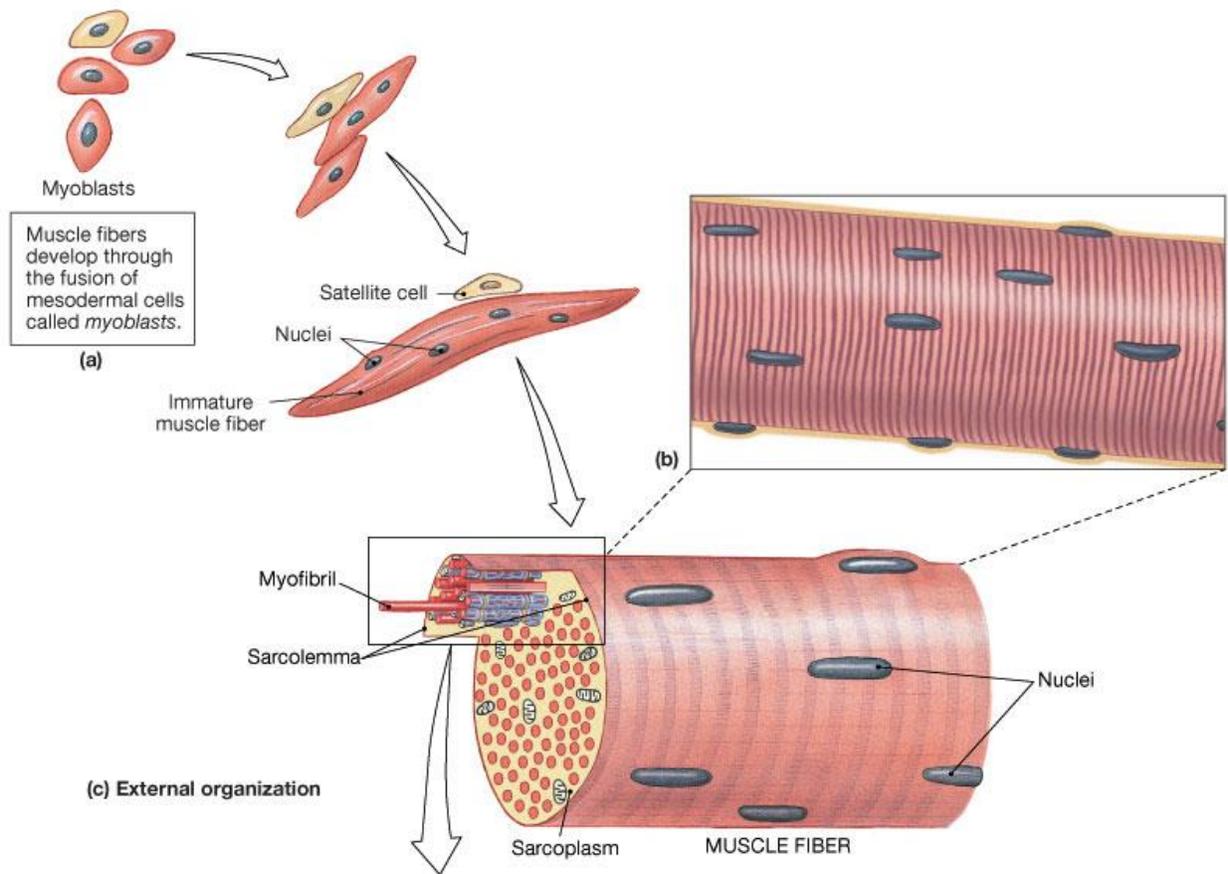


Figure 1.2. Satellite cells occupy a sub-laminal position in adult skeletal muscle. Elongated, spindle-shaped, quiescent myoblasts lying in close contact with adult muscle. The satellite cells can be distinguished from the myonuclei by a surrounding basal lamina and more abundant myofibril. The satellite cells are thought to play a role in muscle repair and regeneration (from Hawke et al., 2001).

Chapter II: White Muscle and Liver Metabolites Involved In Enhanced Growth Response of Hybrid Striped Bass

Abstract

In the aquaculture industry it is common to encounter a subgroup of individuals from the cohort that fail to reach the desired market size. The aim of this study was to identify discrete molecular biomarkers, or metabolites, in hybrid striped bass (HSB) (striped bass x white bass) white skeletal muscle with the specific objective of gaining insight into metabolic pathways associated with high growth rate performance and muscle accretion. Metabolites are small molecules (typically less than 1 kDa in size) that include all of the components involved in the collective chemical reactions that occur within and outside of cells and therefore underlie the physiological response of the animal. Farming of HSB is presently the fourth largest form of finfish aquaculture in the U.S., behind only catfish, salmonids and tilapia. HSB were reared in tanks and ponds according to standard two-phase culture procedures until harvest at market size. The average weight of the cohort (721 ± 128 , all values are given as average \pm standard deviation; $N= 238$) was calculated and individuals from the top and bottom 10% in terms of body size ($N=10$ fish each group) were selected as representatives of fish that grow good (total length: 397 ± 3 mm, weight: 956 ± 23 g) or poor (total length: 333 ± 6 mm, weight: 482 ± 26 g), respectively. These cutoffs were chosen as they represent a body size that is typically larger and smaller than that desired at market (680 to 907 g) and therefore represent the extremes of the growth performance distribution. White muscle samples were excised from the left side of HSB just ventral to the anterior margin of the first dorsal fin and posterior to the head. Transverse cross-sections were taken through the tissue with the purpose of evaluating

differences in muscle fiber number, distribution, and diameter. Most of the muscle fibers in the good growth group were at a hyperplastic growth phase, whereas fish in the poor growth group had switched to hypertrophic growth. Liver and muscle tissues were collected (N=18 sample each tissue) for a global quantitative metabolomics panel analysis where a total of 464 and 469 metabolites, respectively were identified, and analyzed by decision tree analysis. The results from this study assessing liver and muscle identified several metabolomic differences, including changes in metabolites related to energetics (e.g., glycolysis and TCA cycle), amino acid catabolism, and biomarkers of stress and inflammation.

Introduction

An Overview of Metabolites and Muscle Growth

Metabolites are the byproducts of metabolism, they are small molecules (typically less than 1 kDa in size) catalyzed by various enzymes and include all of the components involved in the collective chemical reactions that occur within and outside of cells in an organism and that also derive from nutrition (Harris, 2014; Nicholson et al., 2002). The formation of metabolites critically depends on enzymatic activities, and as such they define chemical intermediates along defined “pathways”, which are then used to represent the series of chemical reactions that occur within the cell (Gatlin et al., 2007; 2008; Harris, 2014). Metabolites are classified into two key categories, based on their level of contribution in metabolic processes: 1) Primary metabolites, which are synthesized by the cell, and are crucial for growth (e.g., amino acids, fatty acids, monosaccharides, and nucleotides) and 2) Secondary metabolites, which are compounds produced by an

organism that are not required for primary metabolic processes, although they can have other important functions (e.g., pigments and food additives with applications in aquaculture) (Harris, 2014; Holmes, 2008). Primary metabolites (e.g., amino acids) underlie the entire metabolism or physiology of an organism and represent the end point of all gene and protein activities (e.g., growth) (Nicholson et al., 2002). For example, amino acids are the building blocks of proteins and also regulate the pace of growth metabolism. In general, growth metabolism in vertebrates is constructed around the theory that protein synthesis is the direct fuel for the somatic growth of the animal (Houlihan et al., 2008). In fishes, the majority of somatic growth is invested in the mass of muscle tissue, mostly the white fibers (i.e., myofibrils), which account for two-thirds of the muscle protein synthetic activities (Mommsen, 2001; Kwasek et al., 2015; Weatherley, 1990; Weatherley et al., 1988). Muscle growth predominantly occurs in fishes by the formation of new myofibrils (hyperplasia) and the subsequent volumetric growth of these myofibrils (hypertrophy) (Gill et al., 1989; Weatherley et al., 1988). Thus, myofibril proteins may be considered the most important contributor to fish somatic growth and production of the fillet.

An Overview of Technology Platform for Metabolomics

There are multiple ways to separate, identify, and measure metabolites. They can be analyzed by standard tools of chemicals analysis such as molecular spectroscopy (MS) and tandem mass spectrometry (MS/MS) (Roessner and Beckles, 2009). The resolution, sensitivity and selectivity of these technologies can be enhanced or modified by coupling them to gas chromatography (GC) or liquid chromatography (LC) steps (Goodacre et al., 2004; Roessner and Beckles, 2009). The technologies commonly exploited for different

metabolomic strategies are shown in (Table 2.1). Generally, the technology platform of choice depends on the type of sample to be analyzed (Goodacre et al., 2004).

Metabolomics is a novel technology that provides a global characterization and quantification of all metabolites in a biological sample (Szpunar, 2005). Importantly, the collective responses of metabolites in the fish can be used to better understand its physiological responses (e.g., growth performance). For example, the importance of certain amino acids as building blocks of protein and other metabolites, such as lipids and carbohydrates on tissue metabolism and growth of white muscle, can be simultaneously evaluated in a single sample using metabolomics approaches.

Specific Literature Review on Metabolomics and Proteomics of Fish Growth

Metabolomics investigations have not yet been conducted on hybrid striped bass (HSB), however a related proteomics approach using 1-D electrophoresis and tandem mass spectrometry was used by Kwasek et al. (2012), on yellow perch (*Perca flavescens*). These two approaches use mass spectrometry to identify and measure the abundance of compounds, however the major difference is that the focus of metabolomics is on small molecules (metabolites) from various biochemical classes and proteomics only resolves peptides derived from proteins. Kwasek et al. (2012) compared the protein expression in skeletal muscle myofibrils associated with differences between fast- and slow-growing yellow perch in order to identify those muscle proteins in fish exhibiting different growth capabilities. The study identified several proteins, of which 10 had different staining intensities between fast- and slow-growing yellow perch, including several important metabolic enzymes (Table 2.2). The results of the study contributed to the identification of

different expression patterns of proteins in fast- and slow-growing yellow perch associated with their muscle growth and indicate that carbohydrate and glycolysis pathways are potentially important indicators of growth status. In another study by Terova et al., (2014), the proteome profile of European sea bass (*Dicentrarchus labrax*) muscle was analyzed using 2-D electrophoresis and tandem mass spectrometry with the aim of providing a more detailed characterization of its specific protein expression profile. A highly populated and well-resolved 2-DE map of the sea bass muscle tissue was generated, and the corresponding protein identity was provided for a total of 49 abundant protein spots (Table 2.3). Upon Ingenuity Pathway Analysis (IPA), the 49 protein identities obtained were classified according to their cellular localization and biological function (Figure 2.1). Over 60% of the proteins were found to be localized in the cytoplasm, 16% were localized in the nucleus and 14% in the muscle filaments (Figure 2.1B). About 51% of the identified proteins were enzymes involved in glycolytic process, 25% were structural proteins, 7% were binding proteins, and a low percentage were transport and biosynthesis proteins (Figure 2.1A). The proteins mapped in the sea bass muscle profile were mostly related to glycolysis and to the muscle myofibril structure, together with other biological activities crucial to fish muscle metabolism. Collectively, these proteomic studies in fishes point to major metabolic pathways involved with growth, including glycolysis, biosynthesis, and high energy phosphate compound production in the muscle.

Therefore, it is hypothesized that small molecules identified in the muscle metabolomics study of hybrid striped bass would likely correspond to some of these pathways and they may be shown to be important indicators of growth status or growth potential (i.e., biomarkers). Little is known about the global metabolite profiles in the liver

of fishes and this study is one of the first characterizations. We also evaluated which tissue (liver or muscle) provides more information regarding the growth status of the hybrid striped bass. Since the muscle is the actual tissue undergoing active growth, localized metabolite markers may be more indicative of growth status of the fish compared to those of the liver, however we verified this contention to be false.

Study Design

HSB were reared common garden in tanks and ponds according to standard two-phase aquaculture procedures until harvest at market size (D'Abramo and Frinsko, 2008). The average weight and total length of the cohort were calculated and individuals from the top and bottom 10% in terms of body size (N=10 fish each group) were selected as representatives of fish that grow good or poor, respectively. These cutoffs were chosen as they represent a body size that is typically larger and smaller than that desired at market and therefore represent the extremes of the distribution. Samples of the liver and muscle were collected from each fish for metabolomics analysis to identify the differences of key small molecules important to growth in hybrid striped bass.

Samples of muscle were also collected for histological analysis of white muscle morphometrics to evaluate the hyperplasia and hypertrophy in each HSB. Muscle fiber number and diameter has been used to evaluate the state of hyperplasia and hypertrophy in other fishes (Weatherley et al., 1988). HSB white muscle and liver tissue samples were evaluated for metabolite profiling to identify those which differ between the good growth and poor growth fish groups. Several hundreds of different named compounds (metabolites) were expected to be identified in the samples and a machine learning

decision tree (Random Forest) was used to analyze the data. Machine learning is a powerful approach to identifying patterns within data, in this case those metabolites that are related to growth performance of hybrid striped bass. Additionally, the data also provided the first characterization of muscle metabolite profiles of HSB white muscle and liver, which may be involved in various biological process and functions attributed to muscle growth. This study will set the foundation for future HSB research, which aims to develop novel molecular and/or genetic methods and protocols for monitoring growth dynamics (i.e., biomarkers or growth). Furthermore, combined with the application of genetic improvement technology and gene expression studies (i.e., transcriptomics, discussed below in **Objective 2, Chapter III**), manipulation of HSB growth and body size may be achieved by understanding and controlling certain signaling pathways.

Materials & Methods

Experimental Animals and Tissue Sampling

Domestic HSB (Sunshine Bass) were produced by crossing the domestic striped bass (*Morone saxatilis* male) with the domestic white bass (*Morone chrysops* female) and these phase I fingerlings were reared in ponds at the North Carolina State University Pamlico Aquaculture Field Laboratory (NCSU PAFL) in Aurora, NC, USA according to standard two-phase culture procedures until harvest at market size (Harrell et al., 1990; Hodson and Sullivan, 1993; Harrell, 1997; Sullivan et al., 2003; McGinty and Hodson, 2008; Reading et al., 2018). The average weight and total length of the cohort were calculated and individuals from the top and bottom 10% in terms of body size (N=10 fish each group) were selected as representatives of fish that grow poor (poor-growth) or good (good-

growth), respectively. These cutoffs were chosen as they represent a body size that is typically larger and smaller than that desired at market and therefore represent the extremes of the normal distribution curve. Selected individuals (N=4 from each growth-group of fish) from the good-growth groups and poor-growth groups were quickly sampled to mitigate any influence of capture stress factors. Fish were euthanized by immersion in a solution of eugenol and tricaine methane sulfonate (MS-222), according to standard hatchery procedures (Harrell et al., 1990; Harrell, 1997).

White muscle tissue was excised from the left side of each HSB just ventral to the anterior margin of the first dorsal fin and posterior to the head (poor-growth, N=5; good-growth, N=5). An approximately 3.5 mm³ thick muscle block was fixed in Bouin's fluid for 3 days (Suresh et. al., 1998) and dehydrated in 50% alcohol for one day. Fixed tissues were sent to the NCSU College of Veterinary Medicine Histology Laboratory (Raleigh, NC, USA) where they were further dehydrated through an ethanol series, embedded in paraffin and cross sectioned at 5 µm perpendicular to the muscle fibers. The sections were routinely stained in haematoxylin and eosin and mounted on microscope slides for muscle morphometry analysis.

Muscle and liver tissues (N=9 from each growth group for each tissue) were sampled from 10 fish in the poor-growth and good-growth groups. Nine of the ten samples were sent to Metabolon Inc. (Morrisville, NC, USA) for Metabolomics analysis (see *Metabolomics Analysis* below).

Evaluations of Growth Parameters

The total length, wet weight, gonad weight, and liver weight were measured and Gonadosomatic index (GSI) was calculated for each of the two growth groups. The following formula was used to calculate GSI (Komourdjian et. al., 1978):

$$GSI = \frac{\text{Gonad Weight (g)}}{\text{Total Body Weight (g)}} \times 100$$

Hepatosomatic index (HSI) which is the calculation of the liver weight in grams expressed as a percent of total body weight in grams was calculated using the following formula (Komourdjian et. al., 1978):

$$HSI = \frac{\text{Liver Weight (g)}}{\text{Total Body Weight (g)}} \times 100$$

Fulton's condition factor (K) was calculated by dividing the wet weight (W) of the fish in grams by the cube length (L³) of the fish in centimeter, the factor 100 is then used to bring K close to unity (Fulton, 1902). The following formula was used to calculate the K:

$$K = 100 \frac{\text{Wet Weight (W) (g)}}{\text{Cube Total Length (L}^3\text{) (cm}^3\text{)}}$$

Growth was measured at week 51.6 by sampling the fish in each tank and obtaining a group wet weight for approximately 50 fish from three tanks. The specific growth rate (SGR) was calculated based on average group weights of fish from each growth group at the end of the culturing period. Out of the 50 fish, 20 individuals (10 from each group growth performance) were individually weighed. The SGR was calculated using the following formula (Turano, 2006; Turano et al., 2007):

$$SGR = \frac{100 \times (\ln \text{Weight}_f - \ln \text{Weight}_i)}{(\text{Time}_f - \text{Time}_i)}$$

Muscle Histology Analysis

Muscle fiber morphometric data were enumerated using similar methodology to Weatherley & Gill (1987) to evaluate growth by hypertrophy (i.e., fiber area) and hyperplasia (i.e., fiber number) in hybrid striped bass. Representative images of muscle fibers for enumeration were collected in triplicate using an Olympus CH light microscope (4x) connected to a Celestron digital microscope camera (10x). All images were evaluated in a blind fashion where the scorer was unaware of the treatment groups. Briefly, three representative and random images were acquired at the same magnification from each tissue section and considered to be technical replicates. All fibers within the field of view were enumerated from the technical replicates to calculate the average total number of muscle fibers in fish from each of the good-growth and poor-growth groups. The cross-sectional areas of individual fibers within each tissue section also were determined through calibration with a stage micrometer using Image J imaging software (Fiji, NIH, Bethesda, MD, USA). In this case only those fibers that were entirely contained within the field of view were measured (i.e., those marginal fibers or those that were partially transected by the margin of view were excluded to avoid any measuring bias). Assuming the muscle fibers to be circular, the diameter of each was calculated as a geometric derivative of its area using Image J. The presence of muscle fibers $\leq 20 \mu\text{m}$ was evaluated as an indicator of growth by hyperplasia (Weatherley et al., 1988) and these fibers also were enumerated as described above, however only those fibers that were entirely within the field of view were counted.

Metabolomic Analysis

To identify biomarkers and gain better understanding of the metabolic pathways in the hybrid striped bass associated with the growth performance (i.e., body weight and muscle accretion, potential): 18 liver samples (poor-growth n=9, good-growth n=9) and 18 white muscle samples (poor-growth n=9, good-growth n=9) were selected from each HSB growth group for metabolomic analysis. All chosen samples were submitted to Metabolon where they were immediately inventoried and stored at -80°C. A unique identifier was assigned to the project, by the Metabolon LIMS system, and was used to track the project through handling, tasks, results, and analysis.

Initial preparation was conducted using the automated MicroLab STAR® system from Hamilton Company. Several recovery standards were added prior to the first step in the extraction process for Quality Control (QC) purposes. To remove protein, dissociate small molecules bound to protein or trapped in the precipitated protein matrix, and to recover chemically diverse metabolites, proteins were precipitated with methanol under vigorous shaking for 2 min (Glen Mills GenoGrinder 2000) followed by centrifugation. The resulting extract was divided into five fractions: two for analysis by two separate reverse phase (RP)/Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectroscopy (UPLC-MS/MS) methods with positive ion mode electrospray ionization (ESI), one for analysis by RP/UPLC-MS/MS with negative ion mode ESI, one for analysis by Hydrophilic interaction liquid chromatography (HILIC)/UPLC-MS/MS with negative ion mode ESI, and one sample was reserved for backup. Samples were placed briefly on a TurboVap® (Zymark) to remove the organic solvent. Prior to analysis, extracted muscle and liver tissues were stored overnight under nitrogen before preparation for analysis.

Raw data were extracted, peaks-identified (the area-under-the-curve) and QC processed using Metabolon's hardware and software. Compounds were identified by comparison to Metabolon's library entries of purified standards or recurrent unknown entities. All molecules within the library contain recorded information, such as retention time/index, mass to charge ratio, and chromatographic data (including mass spectrometer data). Identification of metabolites must meet the retention time/index, mass to charge ratio, and chromatographic standards. Unidentified metabolite information was stored in the Metabolon database for possible future identification through technological advancement, although it is not reported here.

After identification and analyses were completed, metabolites were curated to ensure accuracy and consistency. Normalized values of raw area counts were recorded for each of the metabolites identified. Any missing values within the data set were metabolites that fell below the mass spectrometers level of detection. These counts were then scaled, using the natural logarithm, so the median was equal to one which reigned the effects of any potential outliers. Missing values were imputed with the minimum value for that metabolite across all muscle and liver tissue samples.

Data Analysis

Growth Parameters

Statistical differences in mean total length, wet weight, GSI and HSI between fish from the two growth group were evaluated with Oneway Student's t-test at an alpha level of 0.05 using Statistical Analysis System – JMP (SAS JMP®, 11.0.0; SAS Institute Inc., Cary, NC, USA).

Muscle Histology

Statistical analyses were conducted to account for any significant muscle morphometric differences between the two growth groups. The mean number of fibers, the mean fiber diameters (μm), and the mean muscle fibers that were $\leq 20 \mu\text{m}$ (as an indicator of hyperplastic growth) were compared between the two growth groups with One-way Student's t-test at an alpha level of 0.05 using Statistical Analysis System – JMP Pro. 14.0 (JMP®, 11.0.0; SAS Institute Inc., Cary, NC, USA). In addition, a linear regression analysis was used to assess the relationship between the average fiber number and the average fiber diameter (μm) between the two growth groups through their corresponding average fish wet weights (g) using the following linear equation:

$$y = mx + b$$

Metabolomics

Two types of statistical analysis were performed to identify differences in concentration of metabolites between the two growth groups: (1) significance tests (Welch's Two Sample t-test) and (2) classification analysis. Standard statistical analyses were performed in ArrayStudio on log transformed data. For those analyses not standard in ArrayStudio, the programs R (<http://cran.r-project.org/>) or JMP were used.

Welch's Two-sample t-test was used on liver and muscle tissue samples to determine significant differences in metabolite concentrations between the two growth groups. A principal component analysis (PCA) was conducted to identify if the fish samples from the poor-growth and good-growth groups could be separated based solely on the metabolic signatures. To determine which of the identified metabolites made the largest contribution to the classification, a "metabolite importance" measure known as the "Mean

Decrease Accuracy” (MDA) was computed. The MDA was determined by the Random Forest machine learning analysis, which provided an “importance” rank ordering of metabolites (e.g., a sensitivity analysis), and 30 metabolites were ranked as the top most important for differentiating fish between the good and poor-growth groups. Those top 30 metabolites for each of the liver and muscle were assigned as potentially worthy of further investigation (i.e., akin to “statistically” significant).

Results

Experimental Animals, Tissue Sampling, and Evaluations of Growth Parameters

The weight class of bottom and top 10% of the fish in terms of body size in grams, which represents fish that grow poor (poor-growth) or good (good-growth), respectively are shown in Figure 2.2. The average total length, average wet weight, average GSI, average HSI, average K, and average SGR for the poor- and good-growth groups of fish growth parameters data are shown in Table 2.4. There were significant differences between the poor- and good-growth groups in regard to average total length (One-way Student’s t-test, $p = 0.0001$) (Figure 2.3A), average wet weight (One-way Student’s t-test, $p = 0.0001$) (Figure 2.3B), average GSI (One-way Student’s t-test, $p = 0.0147$) (Figure 2.3C), average HSI (One-way Student’s t-test, $p = 0.0037$) (Figure 2.3D), average K (One-way Student’s t-test, $p = 0.0001$) (Figure 2.2E), and average SGR (One-way Student’s t-test, $p = 0.0001$) (Figure 2.3F), indicating that the fish were of different sizes (i.e., growth condition) in the two growth groups at the time of sampling.

Muscle Histology Analysis

An example of the muscle histology images is shown in Figure 2.4. There were no significant differences in average total muscle fiber number or in average muscle fiber diameter between growth groups (One-way Student's t-test, $\alpha=0.05$) (Figure 2.5A, 2.5C). There was a significant difference in number of muscle fibers that were $\leq 20 \mu\text{m}$ in diameter (One-way Student's t-test, $p=0.0274$), such that fish from the good growth group had more smaller muscle fibers (Figure 2.5B). A linear regression between the average total number of muscle fibers and fish wet weight of the good and poor growth groups showed a very little correlation ($R^2=0.0858$ and 0.0185 , respectively) (Figure 2.6A). A linear regression between the average muscle fiber diameter and fish wet weight of the good and poor growth groups showed a weak correlation ($R^2=0.1813$ and 0.2165 , respectively) (Figure 2.6B). The observed trend was that larger fish generally had smaller muscle fibers and this trend was observed in fish from both growth groups (i.e., the largest fish in both the poor- and good- growth groups generally had smaller muscle fibers).

Metabolomics analysis

A total of 464 metabolites were identified in the HSB liver tissue samples and 469 metabolites in muscle tissue samples. Metabolites were identified, organized, and had their raw area counts provided for each sample (Table 2.5 and Table 2.6, respectively; Supplemental file available at North Carolina State University library online repository).

Following log transformation and imputation of missing values, if any, with the minimum observed value for each compound, Welch's Two-sample t-test was used to identify metabolites that differed significantly between the two fish growth group means.

A summary of the numbers of metabolites determined to have statistical significance ($p \leq 0.05$), as well as those approaching significance ($0.05 < p < 0.10$) identified among liver and muscle tissue samples are shown in Table 2.7 and Table 2.8, respectively. A separate analysis using only the higher-powered female cohort was also included to aid discrimination of potential gender effects on the data set in Table 2.7 and Table 2.8, respectively, as described below.

In the Principal Component Analysis (PCA) for liver and muscle tissue samples, data tended to separate by gender (Figure 2.7, 2.8), respectively, with a secondary separation in liver by growth group (i.e., poor-growth, good-growth) (Figure 2.7). Statistical contrasts for the higher-powered female cohort [good-growth (female) vs poor-growth (female)] were also included to aid discrimination of gender effects on the dataset; however, a cursory assessment suggested similar trends in the two contrasts in both liver and muscle datasets.

Random Forest analysis was moderately effective at classifying samples in liver and muscle, with a predictive accuracy of 89% (Table 2.9) and 72%, (Table 2.10), respectively. The analysis identified the top thirty most important metabolites in the liver tissue and muscle tissue samples in relation to the two fish growth groups (Figure 2.9, 2.10), respectively. The metabolic super pathway to which each metabolite of each tissue sample belonged are provided within the legends of Figure 2.9 and Figure 2.10. In liver tissue, lipids made up the greatest percentage of those identified (30.0%). Other represented super pathways were amino acids (20.0%), carbohydrates and energy (13.3%), cofactors and vitamins, nucleotides, and peptides (13.3%), and finally, xenobiotics (3.3%). In muscle tissue, lipids also made up the greatest percentage of those identified (46.67%). Other

represented super pathways were amino acids (20.0%), carbohydrates, energy, nucleotides, and xenobiotics (6.67%), and cofactors and vitamins, and peptides (3.33%).

Discussion

HSB in good- and poor-growth performance groups differed significantly in average weight, length, GSI, HSI, K, and SGR (Figures 2.3A-F), indicating that the fish were of different sizes and growth condition, or metabolism. There was a significant difference between fish in the good and poor growth groups in regard to muscle fibers of ≤ 20 μm in diameter. HSB from the good-growth group had more of these small muscle fibers in the tissue sections analyzed (Figure 2.5C). Larger fish trended towards having muscle fibers of smaller diameter, however the statistical relationship between average muscle fiber diameter and wet fish weight was weak (Figure 2.6B). Collectively, these findings suggest that the HSB from the good-growth group have a greater degree of hyperplasia compared to fish in the poor-growth group. The smaller HSB generally had a greater proportion of larger, hypertrophic muscle fibers. This trend indicates that the larger fish have retained a longer period of hyperplastic muscle growth compared to the smaller fish, which may have had a precocious switch to hypertrophy. This precocious switch in fish muscle from the poor-growth group to hypertrophic muscle growth might have contributed to the overall poor-growth performance observed in this study.

The PCA plots for both liver (Figure 2.7) and muscle (Figure 2.8) tissues of HSB showed a tendency to form sub-clusters that might suggest diverse metabolic inputs into growth potential. The Random Forest machine learning analysis for both the liver (Figure 2.9) and muscle (Figure 2.10) tissues highlighted similar metabolites, including those

related to lipid metabolism (e.g., carnitine conjugates of dicarboxylate fatty acids such as an isobar of pimeloylcarnitine/2-methyladipoylcarnitine, adipoylcarnitine), tricarboxylic acid cycle (e.g., citrate, alpha-ketoglutarate), and redox homeostasis (e.g., gamma-glutamylglutamate, cystathionine, anserine). Several metabolites related to glucose metabolism were also highlighted in the liver Random Forest analysis (e.g., glucosamine-6-phosphate, sedoheptulose-7-phosphate, and glucose 6-phosphate).

Glucose can be utilized to support a variety of physiological processes, including energy generation, fatty acid synthesis, protein glycosylation, and nucleotide biogenesis. Several comparisons in this study showed changes in glycolytic metabolites suggestive of a shift in metabolic functions. Glucose was significantly elevated in muscle in the fish from the good-growth group (Figure 2.11A). This could reflect a change in glucose uptake or use. Other glycolytic metabolites as a class showed a non-significant trend toward increase in the fish in the good-growth group, which would be consistent with increased glucose availability and utilization. However, pyruvate levels were similar between fish from the good- and poor-growth groups (Figure 2.11B), although lactate showed a mild, non-significant increase in fish from the good-growth group (Table 2.6. Supplemental file available at North Carolina State University library online repository). Elevated lactate levels from the fish of the good-growth group may be related to active metabolic rate of the white muscle, which has already been shown to have significantly greater hyperplastic growth compared to the fish from the poor-growth group (Figure 2.4B). Interestingly, nucleotide sugars (UDP-N-acetylglucosamine, UDP-N-acetylgalactosamine) showed a trend, although non-significant, toward elevation in the fish from the poor-growth group (Figure 2.11C, 2.11D). This could suggest a shift in glucose use for hexosamine synthesis, which has

been shown to function in nutrient sensing that might lead to insulin resistance in mammals (Buse, 2006).

Carbon can flow into the tricarboxylic acid (TCA) cycle from a number of sources, including carbohydrates (entering as pyruvate through glycolysis) and lipids (entering as keto acids through beta-oxidation) via conversion of acetyl-CoA to citrate, glutamine (entering as alpha-ketoglutarate), and branched-chain amino acids (entering as acetyl-CoA and succinyl-CoA) (Figure 2.12). Levels of three keto acids in the muscle suggest a shift in inputs into TCA cycle between fish from the two different growth groups (Figures 2.13A-C). The significant increases in butyrylcarnitine and the elevated propionylcarnitine levels in the fish from the good-growth group could suggest an increased pool size due to declining catabolic use. Concurrently, decreases in succinylcarnitine in fish from the good-growth group compared to those from the poor-growth group may suggest declining branched chain amino acid (BCAA) metabolism to support energetics in muscle. The BCAA include leucine, isoleucine and valine. This observation would be consistent with a shift towards anabolism in the muscle, which is observed as hyperplasia in the fish from the good-growth group (Figure 2.4B). Interestingly, dietary supplementation with leucine promotes weight gain in rainbow trout (Choo et al., 1991), which may suggest dietary BCAA supplementation could positively affect growth in hybrid striped bass as well as this pathway was shown to be important by metabolomics analysis here.

Reduced levels of alpha-ketoglutarate in muscle of fish from the good-growth group (Figure 2.14A) could suggest decreased glutaminolysis and input of glutamate into the TCA cycle. The significant increases in acetylphosphate (Figure 2.14B) on the other hand could suggest increased pyruvate input into the TCA cycle, although citrate levels did not differ

between fish from the growth groups (Figure 2.14C) (Campbell, 2004). These muscle metabolites collectively could suggest that fish from the good-growth group primarily metabolize glucose into pyruvate for TCA cycle or lactate for gluconeogenesis. These pathways appeared to be impaired in fish from the poor-growth group, in particular in regards to lipid metabolism discussed below.

While thiamin (Vitamin B₁) was below the threshold of detection in muscle and liver, a related product, thiamin diphosphate, formerly known as thiamin pyrophosphate, showed a non-significant trend towards increase in muscle of fish from the good-growth group compared to the poor-growth group. The biological function of thiamin pyrophosphate, which is the active form of thiamin, is to serve as a cofactor for several enzymes involved primarily in carbohydrate catabolism. The enzymes are important for biosynthesis of a number of cell constituents, including neurotransmitters, and for production of reducing equivalents used in oxidant stress defense and biosynthesis of pentoses used as nucleic acid precursors. It is also known to function as a coenzyme involved in carbohydrate metabolism, thus making keto analogues from amino and fatty acid metabolism available for production of energy (Shepard and Broderick, 2010). Generally, pyruvate binds to thiamin pyrophosphate on pyruvate dehydrogenase enzyme and undergoes decarboxylation (O'Brien et al., 1980). Regulation of the of pyruvate dehydrogenase activity in skeletal muscle plays an important role in fuel selection and glucose homeostasis. Activation of the complex promotes catabolism of glucose, whereas inactivation conserves substrates for hepatic glucose production by gluconeogenesis (Wu et al., 1999). Decreased levels of pyruvate and elevated levels of thiamin pyrophosphate in muscle of fish from the good-growth group, with concurrent lower levels of liver glucose

and significantly elevated levels of muscle glucose in fish from the good-growth group may suggest activation of the pyruvate dehydrogenase enzyme complex which promotes disposal of glucose from the liver to the blood (i.e., muscle). In contrast, decreased levels of thiamin pyrophosphate in muscle of fish from the poor-growth group with elevated levels of pyruvate and increased levels of liver glucose, parallel with a significant decrease in muscle glucose in fish from the poor-growth group suggests some sort of metabolic deficiency where pyruvate is not being catalyzed. This may be due to dysfunction of the pyruvate dehydrogenase complex. Overall, this suggests thiamin deficiency in fish of the poor-growth group, which can be treated with thiamin supplementation in the prepared diet.

The cause of such thiamin deficiency in fish from the poor-growth group remains unknown, however, it could be that thiamin is heat denatured during the feed extrusion process or that poor-growth fish perhaps do not get access to enough feed, and hence have marginal thiamin intake. To make up for food deprivation, these fish may have foraged on crustaceans, small prey fish, and other live food sources during the pond stage of rearing, all of which may contain thiaminases that actually degrade dietary thiamin that may have been included in the prepared feed that they also consumed, thus exacerbating the apparent thiamin deficiency. Many studies have shown that thiaminases are present in the viscera of some fish species, including white bass (*Morone chrysops*) (Greig and Gnaedinger, 1971). From a genetic point of view, it could be that poor-growth HSB lack the gene(s) responsible for mitigating thiaminase metabolism, which might lead to poor-growth performance reported here. Literature shows that animals suffering from thiamin deficiency due to thiaminases generally respond well if treated early, by intra-muscular or

injections of thiamin, and/or by removing the source of the thiaminase from the diet. This is usually sufficient to bring around full recovery. However, in our study and since thiaminase metabolism could have been inherited, future studies should target genes that are responsible for activating or inhibiting metabolomic activity of thiaminases (e.g. *thi20* trifunctional hydroxymethylpyrimidine kinase/phosphomethylpyrimidine kinase/thiaminase).

Further, nicotinic acid was considered one of the most important predictors of growth in this study (Figure 4.9). Nicotinic acid also known as niacin (vitamin B₃) is stable to thermal processing, however carnivorous fishes such as HSB have a relatively high dietary requirement because they are very inefficient converters of tryptophan to niacin and serotonin, the calming neurotransmitter antagonist to cortisol. Niacin deficiency places a higher demand on NAD(H) and NADP(H), which are required for a very large number of enzymes in essentially all metabolic pathways, especially metabolism of carbohydrates, branch-chain amino acids, fats (i.e., cholesterol, fatty acids), and central roles in energy metabolism. Normally, the liver synthesizes niacin from the essential amino acid tryptophan, but the synthesis is extremely slow in carnivorous fishes. Deficiency in nicotinic acid is generally found in individuals with restricted diets and/or substantial gastrointestinal disease, which can be both severe as well as fatal. In our data, tryptophan was significantly elevated in liver of fish from the poor-growth group and it was undetected in muscle of fish from either growth groups. Although cholesterol was below the threshold of detection in the muscle of both groups, it was elevated in liver of fish from the poor-growth group along with related products including the bile acids cholate and taurocholate as previously discussed. This suggests that fish from the poor-growth group were unable to

convert tryptophan to niacin and/or may suffer from dietary niacin deficiency. As above with thiamin, inclusion of supplementary niacin in the prepared diet may improve growth performance of the fish.

Additionally, it could be that the rate-limiting step in converting the essential amino acid tryptophan to niacin, in the liver of fish from the poor-growth group, is Kynureninase, an enzyme that has a high Michaelis constant (K_m) for Vitamin B₆ (Mayfield, 2013). K_m indicates the affinity of an enzyme for a given substrate; the lower the K_m value, the higher the enzyme's affinity for the substrate, hence requires a greater concentration of substrate; i.e., Vitamin B₆, to achieve maximum reaction rate (Snell and Haskell, 1970). In our data, the higher levels of BCAA in the liver of fish from the poor-growth group compared to the fish from the good-growth group, as previously discussed (Figure 2.24), indicates marginal vitamin B₆ deficiency, which can cause some deficit in assimilating the BCAA metabolism, consequently will affect the protein metabolism necessary for muscle. Simultaneously, transamination and liver-oxidation of BCAA, another chemical reaction that transfers an amino group to a ketoacid to form new amino acids (Romo et al., 2011), is also dependent on vitamin B₆ and thiamin pyrophosphate as cofactors (Danner et al., 1978) both enzymes could be the underlying reason for the inefficient metabolism of tryptophan found in fish liver from the poor-growth group. Transamination is commenced by enzymes called transaminases or aminotransferases. Aminotransferases introduce an amino group into a keto-sugar; i.e., ketoacid, to form new amino acid substrate (Sweet et al., 2004). For the reaction to complete, aminotransferases require manifestation of aldehyde containing coenzyme, pyridoxal 5'-phosphate (PLP), a derivative of pyridoxine; i.e., the active form of B₆ vitamer (Parra and Hellmann, 2018). Vitamin B₆ comprises three forms pyridoxine,

pyridoxal, and pyridoxamine. All three forms of B₆ can be converted to the coenzyme PLP. Vitamin B₆ in its coenzyme form is involved in more than 100 enzyme reactions, many concerned with protein metabolism. It can be measured by the concentration of its different blood forms, i.e., muscle (Pitkin et al., 2000; McCormick, 2006). Henceforth, If there is a deficiency in vitamin B₆, the keto acids of these BCAA cannot be formed through transaminase, and subsequently decarboxylated, which is thiamin pyrophosphate dependent, to form the TCA substrates (Chen et al., 2019). Accordingly, BCAA will be build-up, i.e., store in the liver (Waterlow, 2006). Our data support such vitamin B₆ defacing, whereby pyridoxine, pyridoxal, and pyridoxamine (also known as pyridoxamine phosphate) concentrations were all significantly elevated in the liver of fish from the poor-growth group and declined in the muscle, parallel with the significant increase in most of the BCAA concentrations in the liver as previously discussed (Figure 2.24). Vitamin B₆ deficiency can be treated as above, with the other vitamin B types, by the inclusion of supplementary B₆ different forms in the prepared diet, which may improve the fish's growth performance.

Cortisol levels in the muscle tissue of fish in the poor-growth group were significantly elevated compared to those in the good-growth group (Figure 2.15A). Cortisol also was the most important muscle metabolite identified by Random Forest analysis (Figure 2.10). Cortisol is a stress hormone that could possibly indicate stress in the fish from the poor-growth group. The significant decreases in the pro-inflammatory eicosanoids 12-HETE and 12-HEPE (along with a non-significant decrease in 15-HETE) in the muscle tissue from fish of the good-growth group could suggest reduced inflammation, which may positively affect growth (Figures 2.15B-D). The hormone cortisol is released in

response to inflammation and it has anti-inflammatory actions. Chronic high cortisol levels have been associated with decreased growth (muscle wasting) and poor feed conversion in rainbow trout (Gregory and Wood, 1999). Increased stress levels also negatively affect feeding behavior in the closely related sea bass (*Dicentrarchus labrax*) (Leal et al., 2011). Follow-up studies correlating cortisol levels with growth in hybrid striped bass could possibly identify some upstream factors that may influence regulation of growth and energetics in fish muscle. While cholesterol, a precursor of cortisol, was below the threshold of detection in muscle, related products including the bile acids cholate and taurocholate, and 7-hydroxycholesterol (Figures 2.15E-G) showed a non-significant trend toward increase in the muscle of the poor-growth group fish. Additionally, omega-3 and omega-6 polyunsaturated fatty acids (PUFAs) can be utilized to produce anti-inflammatory and pro-inflammatory metabolites. These compounds may possibly improve growth and reduce inflammation if added to the diet.

Cholate and taurocholate showed a similar elevation in the liver of fish in the poor-growth group (Figure 2.16A, 2.16B). Cholesterol levels in the liver were not different between the fish from the different growth groups (Figure 2.16C). This suggests overall that reduced cortisol levels in the good-growth group fish may be associated with decreased cholesterol availability or perhaps some dysfunctional cholesterol metabolism in the poor-growth group fish. For example, if cholesterol metabolism is insufficient or inappropriate, then this may lead to an inability to effectively clear cortisol from the blood following a stress event, which is why it may remain elevated in fish from the poor-growth group (Figure 2.17). Cortisol is typically cleared from the blood by the liver (McCann and Fulton, 1975). Persistent elevated cortisol levels may lead to muscle wasting and poor

muscle growth. The significant increases in choline phosphate, and the elevated levels of glycerophosphorylcholine (GPC), and glycerophosphoethanolamine (GPE), in the muscle tissues from fish of the good-growth group could suggest increased phospholipid turnover or remodeling (Figures 2.18A-C), which would be expected during hyperplasia. Further studies assessing cholesterol levels in the plasma may identify whether these fishes might have been exposed to less stress or experience high cholesterol demand (and decreased availability) that may lead to poor muscle growth.

Glucose levels in the liver tissue of fish in the good-growth group were decreased compared to those in the poor-growth group (Figure 2.19A), along with the significant increases in glucose-6-phosphate, fructose-6-phosphate and dihydroxyacetone phosphate (DHAP) (Figures 2.19B-D) suggestive of increasing glycolytic use. In the absence of changes in glucose import, 3-carbon glycolytic intermediates tend to decrease as glycolytic use increases. A trend toward increase in phosphoenolpyruvate (PEP) (Figure 2.19E) and non-significant increase in 3-phosphoglycerate (Figure 2.19F) in liver are consistent with decreasing glycolytic use in the fish from the poor-growth group. The glycolytic end-product pyruvate showed a trend toward increase in the liver from fish of the good-growth group (Figure 2.20A), while lactate was not significantly changed between the different growth groups, suggestive of increasing glycolytic input into the TCA cycle. Elevated levels of sedoheptulose-7-phosphate and the purine precursor AICA ribonucleotide (AICAR) in the liver of the fish from the good-growth group (Figure 2.20B, 2.20C) could suggest increased glucose use to support nucleotide demand. Significant decreases in hypoxanthine and xanthine in the liver from the good-growth fish group, with a trend toward decrease in allantoin, could suggest increasing purine demand (Figure 2.20C-F). This is consistent with

trends toward non-significant decreases in adenosine 3'-monophosphate (3'-AMP) and adenosine 3',5'-diphosphate (ADP), although guanosine 5'- monophosphate (5'-GMP) was significantly elevated (Figures 2.20G-I). In muscle of fish from the good-growth group, there was a significant increase in xanthine and significant decreases in adenosine 2'-monophosphate (2'-AMP) and guanine compared to fish from the poor-growth group (Figures 2.21A-C). Additionally, there was a trend for muscle inosine and allantoin levels to be non-significantly lower in fish from the good-growth group (Figure 2.21D, 2.21E). Collectively these findings further support purine nucleotide demand in fish from the good-growth group. These nucleotides could be used as a substrate for RNA or DNA synthesis in rapidly dividing cells such as those observed in the hyperplastic muscle (Figure 2.4B). The RNA and DNA content of tissues has been used as a reference of cell growth or proliferation in a number of previous studies (Davis et al., 2015). A trend toward increase in products of glycogen breakdown (e.g., maltohexaose, maltopentaose, maltotetraose) in the liver of fish from the good-growth group (Figures 2.22A-C) could suggest increased glycogen breakdown. Glucose made available from glycogen could be used for a variety of the aforementioned pathways.

In liver, the significantly elevated levels of citrate in the fish of the good-growth group could suggest increasing input of acetyl-CoA into the TCA cycle (Figure 2.23A, 2.23B), which could be consistent with increased glycolytic use as previously discussed. Further, fish from the poor-growth group had elevated levels of acetyl-CoA and very low levels of citrate, indicating a decreased input into the TCA cycle (Figure 2.23B, 2.23A). A non-significant increase in acetylphosphate, which is produced from the conversion of pyruvate to acetyl-CoA, in the fish of the good-growth group further supports increased

glycolytic input into the TCA cycle (Figure 2.23C). Similarly, the significantly elevated levels of alpha-ketoglutarate in the fish of the good-growth group could indicate increased glutaminolysis, as glutamine was also significantly decreased in the liver of fish from the good-growth group (Figure 2.23D, 2.23E), along with several other amino acids discussed below. Overall this was also consistent with increased TCA cycle use among fish in the good-growth group, which indicates active metabolism.

Signs of significantly increased amino acid demand in liver were observed in the fish from the good-growth group. A number of amino acids (e.g., aspartate, glutamine, histidine, isoleucine, leucine, methionine, proline, serine, threonine, tryptophan, tyrosine, valine) were decreased in liver, although not all were significant (Figures 2.24A-L). This may be expected given the high metabolic activity of the hyperplastic muscle growth of the fish from the good-growth group. The depletion of these amino acid might be related to the incorporation into protein in the growing muscle tissue and/or their use for production of energy to support active metabolism (Platell et al., 2000). The significantly decreased levels of urea, ornithine, and citrulline in the liver of fish from the good-growth group (Figures 2.25A-C) could reflect incorporation of these amino acids into proteins in light of the observed decreases in amino acid levels. Interestingly, creatine and creatinine were both significantly elevated in the liver of fish from the good-growth group (Figures 2.25D, 2.25E), which could suggest a redirection of arginine toward creatine production to supplement high energetic demand. Metabolites derived from the BCAA isoleucine, leucine or valine catabolism can enter gluconeogenesis or the TCA cycle for energy production. These first two amino acids from the BCAA (i.e., leucine and isoleucine) were significantly decreased (Figures 2.24D, 2.24E) and, although not statistically significant, valine was also

found to be depleted (Figure 2.24L) in the liver of the fish from the good-growth group, suggestive of changes in amino acid availability or catabolic use. Consistent with catabolic use, the keto-acids 4-methyl-2-oxopentanoate, 3-methyl-2-oxovalerate, and 3-methyl-2-oxobutyrate were also significantly elevated in the fish from the good-growth group (Figures 2.26A-C). While signs of changing BCAA catabolism were observed, succinate and succinylcarnitine (a surrogate reporter for succinyl-CoA) were not significantly changed in liver between the different growth groups (Figure 2.26D, 2.26E), however the data suggest that fish from the good-growth group do catabolize BCAA for energy production in the liver (Figure 2.27) and may also use free amino acids for protein polymerization in the muscle.

Fatty acid omega-oxidation may serve as a rescue pathway when beta-oxidation is impaired and it may also supplement beta-oxidation at times of extreme oxidative demand (Papamandjaris et al., 1998; Wanders et al., 2011). While dicarboxylate fatty acids (DFAs) themselves were below the threshold of detection (in both liver and muscle tissues), changes in DFA-carnitine conjugates could suggest subtle differences in beta-oxidative efficiency or fatty acid use in good-growth group fish compared to the poor-growth group. Decreases in carnitine-dicarboxylate fatty acid conjugates with the significant decrease in adipoylcarnitine and an isobar of pimeloylcarnitine/3-methyladipoylcarnitine in the liver of the good-growth group fish could suggest changes in fatty acid omega-oxidation (Figures 2.28A-C). Adipoylcarnitine also showed a non-significant decrease in muscle of fish from good-growth group (Figure 2.29). Collectively, low levels of omega- or beta-oxidation compounds in the tissues of the good-growth group fish suggests that they are not stressed and growing well. Conversely, elevated levels of these compounds in fish from the poor-growth group may suggest exposure to some form of stress and/or metabolic dysfunction.

In conclusion, the results from this global metabolomic study assessing liver and muscle from good- or poor-growth hybrid striped bass identified several metabolomic differences, including changes in metabolites related to energetics (e.g., glycolysis and TCA cycle), amino acid catabolism, and biomarkers of stress and inflammation between HSB that exhibited good- and poor-growth performance. Samples tended to separate by gender, with samples in the liver forming secondary populations by group, which is suggestive of more profound metabolomic changes in the liver over muscle and this is consistent with observed significant changes in the metabolites. Further studies assessing markers known to contribute to growth (e.g., plasma growth hormone, microbiome differences, efficiency of feed conversion) could provide additional criteria by which to assess changes in this and similar datasets as they become available. A number of metabolic differences were observed in the liver including signs of increased glycolytic and glutaminolytic input into the TCA cycle, which is suggestive of increased energy demand in fish from the good-growth group compared to the poor-growth fish group. Decreased catabolic use of branched-chain amino acids in muscle could suggest shifts toward anabolic growth along with increased glucose availability. This also was associated with non-significant changes in glycolytic metabolites suggestive of increased use in fish from the good-growth group. Finally, increases in eicosanoids and cortisol in muscle could suggest increased stress and inflammation in fish from the poor-growth group, which has been shown to have negative effects on fish growth. Fish from the poor-growth group also had elevated metabolites related to fatty acid omega-oxidation, which further supports this.

References

- Alamgir, A. N. M.** (2018). Secondary Metabolites: Secondary Metabolic Products Consisting of C and H; C, H, and O; N, S, and P Elements; and O/N Heterocycles. In *Therapeutic Use of Medicinal Plants and their Extracts: Volume 2* (pp. 165-309). Springer, Cham.
- Andersen, S. M., Waagbø, R., & Espe, M.** (2016). Functional amino acids in fish health and welfare. *Frontiers in Bioscience*, 8, 143-169.
- Buse, M. G.** (2006). Hexosamines, insulin resistance, and the complications of diabetes: current status. *American Journal of Physiology-Endocrinology And Metabolism*. 290(1), E1-E8.
- Campbell, I.** (2004). Intermediary metabolism. *Anaesthesia & intensive care medicine*. 5(4), 141-143.
- Chen, P. Y. T., Li, B., Drennan, C. L., and Elliott, S. J.** (2019). A reverse TCA cycle 2-oxoacid: ferredoxin oxidoreductase that makes CC bonds from CO₂. *Joule*, 3(2), 595-611.
- Choo, P. S., Smith, T. K., Cho, C. Y., & Ferguson, H. W.** (1991). Dietary excesses of leucine influence growth and body composition of rainbow trout. *The Journal of nutrition*. 121(12), 1932-1939.
- Cunha, T. J.** (2012). *Horse feeding and nutrition*. Academic Press.
- D'Abramo, L.R. and M.O. Frinsko.** (2008). Hybrid Striped Bass: Pond Production of Food Fish. Southern Regional Aquaculture Center. Stoneville, Mississippi. Publication Number 303.

- Danner, D. J., Lemmon, S. K., and Elsas II, L. J.** (1978). Substrate specificity and stabilization by thiamine pyrophosphate of rat liver branched chain α -ketoacid dehydrogenase. *Biochemical medicine*, 19(1), 27-38.
- Davis, W. J., Lehmann, P. Z., and Li, W.** (2015). Nuclear PI3K signaling in cell growth and tumorigenesis. *Frontiers in cell and developmental biology*, 3, 24.
- Fiehn, O.** (2001). Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comparative and functional genomics*. 2(3), 155-168.
- Fulton, T. W.** (1902). The rate of growth of sea fishes. *20th Annual Report of the Fishery Board of Scotland*. 1901(3), 326-446.
- Gatlin, Delbert M., F Houlihan, D., Boujard, T., & Jobling, M. (Eds.).** (2008). *Food intake in fish*. John Wiley & Sons.
- Gatlin, Delbert M., Frederic T. Barrows, Paul Brown, Konrad Dabrowski, T. Gibson Gaylord, Ronald W. Hardy, Eliot Herman, Gongshe Hu, Åshild Krogdahl, Richard Nelson, Kenneth Overturf, Michael Rust, Wendy Sealey, Denise Skonberg, Edward J Souza, David Stone, Rich Wilson, and Eve Wurtele.** (2007). Expanding the utilization of sustainable plant products in aquafeeds: a review. *Aquaculture research*. 38(6), 551-579.
- Gill, H. S., Weatherley, A. H., Lee, R., & Legere, D.** (1989). Histochemical characterization of myotomal muscle of five teleost species. *Journal of fish biology*. 34(3), 375-386.
- Gómez-Requeni, P., de Vareilles, M., Kousoulaki, K., Jordal, A. E. O., Conceição, L. E., & Rønnestad, I.** (2011). Whole body proteome response to a dietary lysine imbalance

- in zebrafish *Danio rerio*. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*. 6(2), 178-186.
- Goodacre, R., Vaidyanathan, S., Dunn, W. B., Harrigan, G. G., & Kell, D. B. (2004).** Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends in biotechnology*. 22(5), 245-252.
- Gregory, T. R., & Wood, C. M. (1999).** The effects of chronic plasma cortisol elevation on the feeding behaviour, growth, competitive ability, and swimming performance of juvenile rainbow trout. *Physiological and Biochemical Zoology*. 72(3), 286-295.
- Greig, R. A., & Gnaedinger, R. H. (1971).** *Occurrence of thiaminase in some common aquatic animals of the United States and Canada* (No. 631). US Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service.
- Guyton, J. R. (2004).** Extended-release niacin for modifying the lipoprotein profile. *Expert opinion on pharmacotherapy*, 5(6), 1385-1398.
- Harrell, R.J. (Ed). (1997).** Striped Bass and Other Morone Culture. *Developments in Aquaculture and Fisheries Science* (Vol. 30). Elsevier, Amsterdam, The Netherlands.
- Harrell, R.M., Kerby, J.H., Minton, R.V. (Eds.). (1990).** Culture and Propagation of Striped Bass and Its Hybrids. *American Fisheries Society*, Bethesda, MD. 7-16.
- Harrigan, G. G., & Goodacre, R. (Eds.). (2012).** Metabolic profiling: its role in biomarker discovery and gene function analysis. *Springer Science & Business Media*.
- Harris, E. (2014).** Biochemical facts behind the definition and properties of METABOLITES.
- Hodson, R.G., and C.V. Sullivan. (1993).** Induced maturation and spawning of domestic and wild Striped Bass, *Morone saxatilis* (Walbaum), broodstock with implanted GnRH analogue and injected hCG. *Aquaculture Research*. 24(3), 389-398.

- Holmes, E., Wilson, I. D., & Nicholson, J. K.** (2008). Metabolic phenotyping in health and disease. *Cell*. 134(5), 714-717.
- Houlihan, D., Boujard, T., & Jobling, M. (Eds.)**. (2008). Food intake in fish. *John Wiley & Sons*.
- Kirkland, J. B., and Meyer-Ficca, M. L.** (2018). Niacin. In *Advances in food and nutrition research* (Vol. 83, pp. 83-149). Academic Press.
- Komourdjian, M. P., Burton, M. P., and Idler, D. R.** (1978). Growth of rainbow trout, *Salmo gairdneri*, after hypophysectomy and somatotropin therapy. *General and comparative endocrinology*, 34(2), 158-162.
- Kurutas, E. B.** (2015). The importance of antioxidants which play the role in cellular response against oxidative/nitrosative stress: current state. *Nutrition journal*, 15(1), 1-22.
- Kwasek, K. A.** (2012). *The nutritional and genetic effects on body growth, reproduction and molecular mechanisms responsible for muscle growth in yellow perch Perca flavescens* (Doctoral dissertation, The Ohio State University).
- Kwasek, K., Wick, M., & Dabrowski, K.** (2015). Muscle Protein Characteristic and Its Association with Faster Growth in Percids and Other Teleosts. In *Biology and Culture of Percid Fishes* (pp. 339-352). Springer Netherlands.
- Leal, E., Fernández-Durán, B., Guillot, R., Ríos, D., & Cerdá-Reverter, J. M.** (2011). Stress-induced effects on feeding behavior and growth performance of the sea bass [*Dicentrarchus labrax*]: a self-feeding approach. *Journal of Comparative Physiology B*. 181(8), 1035-1044.

- Lindon, J. C., Holmes, E., & Nicholson, J. K.** (2003). Peer Reviewed: So What's the Deal with Metabonomics?. *Analytical Chemistry*. 75(17), 384-A.
- Magnier-Gaubil, Clarice, Jean-Marc Herbert, Rozenn Quarck, Béla Papp, Elisabeth Corvazier, Frank Wuytack, Sylviane Lévy-Tolédano, and Jocelyne Enouf.** (1996). Smooth Muscle Cell Cycle and Proliferation Relationship Between Calcium Influx And Sarco-Endoplasmic Reticulum Ca²⁺ ATPase Regulation. *Journal of Biological Chemistry*. 271(44), 27788-27794.
- Mayfield, J. A.** (2013). *The flavin-dependent N-hydroxylating ornithine monooxygenase from Aspergillus fumigatus and the heme-dependent O₂-generating chlorite dismutase from Dechloromonas aromatica: Studies of oxygen consumption and production by two microbial enzymes.* University of Notre Dame.
- McCann, V. J., and Fulton, T. T.** (1975). Cortisol metabolism in chronic liver disease. *The Journal of Clinical Endocrinology & Metabolism*. 40(6), 1038-1044.
- McCormick, D. B.** (2006). Vitamin B-6. Present Knowledge in Nutrition. *Present knowledge in nutrition, 9th edn.* International Life Sciences Institute, Washington, DC.
- McGinty, A.S. and Hodson, R.G.** (2008). Hybrid striped bass: hatchery phase. *Southern Regional Aquaculture Center*.
- Mommsen, T. P.** (2001). Paradigms of growth in fish. *Comparative biochemistry and physiology part B: Biochemistry and molecular biology*. 129(2), 207-219.
- Nakagawa, T., & Nagayama, F.** (1989). Enzymatic properties of fish muscle aldolase. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*. 92(2), 405-410.

- Neinast, M.D., Jang, C., Hui, S., Murashige, D.S., Chu, Q., Morscher, R.J., Li, X., Zhan, L., White, E., Anthony, T.G. and Rabinowitz, J.D** (2019). Quantitative analysis of the whole-body metabolic fate of branched-chain amino acids. *Cell metabolism*, 29(2), 417-429.
- Nicholson, J. K., Connelly, J., Lindon, J. C., & Holmes, E.** (2002). Metabonomics: a platform for studying drug toxicity and gene function. *Nature reviews Drug discovery*. 1(2), 153-161.
- Nicholson, J. K., Lindon, J. C., & Holmes, E.** (1999). 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica*, 29(11), 1181-1189.
- O'Brien, T. A., Kluger, R., Pike, D. C., & Gennis, R. B.** (1980). Phosphonate analogues of pyruvate. Probes of substrate binding to pyruvate oxidase and other thiamin pyrophosphate-dependent decarboxylases. *Biochimica et Biophysica Acta (BBA)-Enzymology*, 613(1), 10-17.
- Papamandjaris, A. A., MacDougall, D. E., & Jones, P. J.** (1998). Medium chain fatty acid metabolism and energy expenditure: obesity treatment implications. *Life sciences*. 62(14), 1203-1215.
- Parra, M., Stahl, S., & Hellmann, H.** (2018). Vitamin B₆ and its role in cell metabolism and physiology. *Cells*, 7(7), 84.
- Pitkin, R. M., Allen, L. H., Bailey, L. B., and Bernfield, M.** (2000). Dietary Reference Intakes for Thiamin, riboflavin, niacin, vitamin B₆, folate, vitamin B₁₂, Pantothenic acid, biotin and choline. *Washington, DC*.

- Platell, C., Kong, S. E., McCauley, R., & Hall, J. C.** (2000). Branched-chain amino acids. *Journal of gastroenterology and hepatology*. 15(7), 706-717.
- Pocket, K.** (2006) No. 15. 'Omics' Sciences: Genomics, Proteomics, and Metabolomics. International Service for the Acquisition of Agribiotech Applications ISAAA.
- Reading, B.J., McGinty, A.S., Clark, R.W., Hopper, M.S., Woods III, L.C., and Baltzegar, D.A.** (2018). Genomic Enablement of Temperate Bass Aquaculture (Family Moronidae), In Breeding and Culture of Perch and Bass (Wang, H., and Liang, X., Eds.). *Science China Press (Chinese Academy of Sciences)*.
- Reddish, J. M., St-Pierre, N., Nichols, A., Green-Church, K., & Wick, M.** (2008). Proteomic analysis of proteins associated with body mass and length in yellow perch, *Perca flavescens*. *Proteomics*. 8(11), 2333-2343.
- Riaz, M. N., Asif, M., & Ali, R.** (2009). Stability of vitamins during extrusion. *Critical Reviews in Food Science and Nutrition*, 49(4), 361-368.
- Roessner, U., & Beckles, D. M.** (2009). Metabolite measurements. In *Plant metabolic networks* (pp. 39-69). Springer New York.
- Romo, A. J., and Liu, H. W.** (2011). Mechanisms and structures of vitamin B₆-dependent enzymes involved in deoxy sugar biosynthesis. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1814(11), 1534-1547.
- Shepard, E. M., & Broderick, J. B.** (2010). S-Adenosylmethionine and iron-sulfur clusters in biological radical reactions: the radical SAM superfamily.
- Sidransky, H.** (2001). *Tryptophan: biochemical and health implications*. CRC Press.
- Snell, E. E., and Haskell, B. E.** (1970). The metabolism of vitamin B₆. In *Comprehensive biochemistry* (Vol. 21, pp. 47-71). Elsevier.

- Sullivan, C.V., Hiramatsu, N., Kennedy, A.M., Clark, R.W., Weber, G.M., Matsubara, T. and Hara, A.** (2003). Induced maturation and spawning: opportunities and applications for research on oogenesis. *Fish Physiology and Biochemistry*. 28(1-4), 481-486. <https://doi.org/10.1023/B:FISH.0000030635.92568.0a>
- Suresh, A. V., & Sheehan, R. J.** (1998). Muscle fibre growth dynamics in diploid and triploid rainbow trout. *Journal of Fish Biology*. 52(3), 570-587.
- Sweet, C. R., Ribeiro, A. A., & Raetz, C. R.** (2004). Oxidation and transamination of the 3''-position of UDP-N-acetylglucosamine by enzymes from *Acidithiobacillus ferrooxidans* Role in the formation of lipid A molecules with four amide-linked acyl chains. *Journal of Biological Chemistry*, 279(24), 25400-25410.
- Szpunar, J.** (2005). Advances in analytical methodology for bioinorganic speciation analysis: metallomics, metalloproteomics and heteroatom-tagged proteomics and metabolomics. *Analyst*. 130(4), 442-465.
- Terova, G., Pisanu, S., Roggio, T., Preziosa, E., Saroglia, M., & Addis, M. F.** (2014). Proteomic profiling of sea bass muscle by two-dimensional gel electrophoresis and tandem mass spectrometry. *Fish physiology and biochemistry*. 40(1), 311-322.
- Turano, M. J.** (2006). *Effect of cyclic feeding on compensatory growth and water quality in hybrid striped bass, Morone chrysops × M. saxatilis culture* (Doctoral dissertation, PhD dissertation).
- Turano, Marc J., Russell J. Borski, and Harry V. Daniels.** (2007). Compensatory growth of pond-reared Hybrid Striped Bass, *Morone chrysops × Morone saxatilis*, Fingerlings. *Journal of the World Aquaculture Society*. 38(2), 250-261.

- Wanders, R. J., Komen, J., & Kemp, S.** (2011). Fatty acid omega-oxidation as a rescue pathway for fatty acid oxidation disorders in humans. *The FEBS journal*. 278(2), 182-194.
- Waterlow, J. C.** (2006). Measurement of whole body protein turnover by the end-product method. *Protein Turnover; CABI: Wallingford, UK*, 97-105.
- Weatherley, A. H.** (1990). Approaches to understanding fish growth. *Transactions of the American Fisheries Society*, 119(4), 662-672.
- Weatherley, A. H. and Gill, H. S.** (1987a). The biology of fish. *Academic press, London*.
- Weatherley, A. H., & Gill, H. S.** (1987b). Growth increases produced by bovine growth hormone in grass pickerel, *Esox americanus vermiculatus* (Le Sueur), and the underlying dynamics of muscle fiber growth. *Aquaculture*. 65(1), 55-66.
- Weatherley, A. H., Gill, H. S., & Lobo, A. F.** (1988). Recruitment and maximal diameter of axial muscle fibres in teleosts and their relationship to somatic growth and ultimate size. *Journal of Fish Biology*. 33(6), 851-859.
- Wu, P., Inskeep, K., Bowker-Kinley, M. M., Popov, K. M., & Harris, R. A.** (1999). Mechanism responsible for inactivation of skeletal muscle pyruvate dehydrogenase complex in starvation and diabetes. *Diabetes*, 48(8), 1593-1599.
- Zhang, S., Hulver, M. W., McMillan, R. P., Cline, M. A., & Gilbert, E. R.** (2014). The pivotal role of pyruvate dehydrogenase kinases in metabolic flexibility. *Nutrition & metabolism*, 11(1), 1-9.

Table 2.1. Classification of metabolomic approaches and the optimal technology used. The terminologies are still evolving and there can be overlaps in their definition, however, the following classification highlights the current options available for monitoring the metabolome (Goodacre et al., 2004).

No	Classification of Metabolomic	Definition	Optimal Technology
1	Metabolite target analysis	Analysis restricted to metabolites, for example, a particular enzyme system that would be directly affected by abiotic or biotic perturbation (Lindon et al., 2003)	<ul style="list-style-type: none"> - Gas chromatography (GC) - High performance liquid chromatography (HPLC) - Hydrophobic interaction chromatography (HILIC)
2	Metabolite profiling	Analysis focused on a group of metabolites, for example, a class of compounds such as carbohydrates, amino acids or those associated with a specific pathway (Fiehn, 2001)	<ul style="list-style-type: none"> - Mass spectrometry (MS/MS) - Gas chromatography (GC) - Parallel liquid chromatography (LC) - Electrospray ionization (ESI)
3	Metabolomics	Comprehensive analysis of the whole metabolome under a given set of conditions (Fiehn, 2001)	<ul style="list-style-type: none"> - Parallel liquid chromatography (LC) - Mass spectrometry (MS/MS) - Nuclear magnetic resonance (NMR)
4	Metabolic fingerprinting	Classification of samples on the basis of provenance of either their biological relevance or origin (Fiehn, 2001).	<ul style="list-style-type: none"> - Mass spectrometry (MS/MS) - Nuclear magnetic resonance (NMR)

Table 2.1 (continued).

No	Classification of Metabolomic	Definition	Optimal Technology
5	Metabolic profiling	Often used interchangeably with 'metabolite profiling'; metabolic fingerprinting is commonly used in clinical and pharmaceutical analysis to trace the fate of a drug or metabolite (Harrigan et al., 2012; Szpunar, 2005).	<ul style="list-style-type: none">- Gas chromatography (GC)- High performance liquid chromatography (HPLC)- Hydrophobic interaction chromatography (HILIC)- Nuclear magnetic resonance (NMR)
6	Metabonomics	Measure of the fingerprint of biochemical perturbations caused by disease, drugs, and toxins (Lindon et al., 2003; Nicholson et al., 1999).	<ul style="list-style-type: none">- Mass spectrometry (MS/MS)- Nuclear magnetic resonance (NMR)- Hydrophobic interaction chromatography (HILIC)

Table 2.2. Proteins associated with fast and slow growing yellow perch. Nine proteins/peptides were identified as predominantly metabolic enzymes, including: Alpha actinin, Glycogen phosphorylase, Phosphoglucose isomerase, ATP synthase, Enolase, Creatine kinase, Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), lactate dehydrogenase, and adenylate kinase. All of these proteins with the exception for sarcoendoplasmic reticulum calcium ATPase (SERCA) were found in bands that showed increased expression in fast-growing fish compared to slow-growing fish suggesting possible association of these proteins with faster growth of yellow perch when expressed at higher levels (Kwasek et al. 2012; Pocket, 2006).

No	Growth performance	Proteins expressed in the skeletal white muscle myofibrils	Biological function and or Pathway	Supporting Reference (s)	Other Reference Species
1	Fast-growing	Alpha actinin	ATP binding	(Terova et al., 2014)	European sea bass (<i>Dicentrarchus labrax</i>)
2		Glycogen phosphorylase	Carbohydrate metabolism, glycogen metabolism	(Terova et al., 2014)	European sea bass (<i>Dicentrarchus labrax</i>)
3		Phosphoglucose isomerase	Involved in glycolytic path ways which increase ATP to support rapid growth	(Gomez-Requeni et al., 2011; Nagakawa and Nagayama, 1989)	Zebrafish (<i>Danio rerio</i>) Red sea bream (<i>Pagrus major</i>) Pacific mackerel (<i>Scomber japonicas</i>) Carp (<i>Cyprinus carpio</i>)
4		ATP synthase	ATP synthesis, hydrogen ion transport, rule base, ion transport, transporter	(Terova et al., 2014)	European sea bass (<i>Dicentrarchus labrax</i>)

Table 2.2. (continued).

No	Growth performance	Proteins expressed in the skeletal white muscle myofibrils	Biological function and or Pathway	Reference (s)	Species
5		Enolase	Involved in glycolytic path ways which increase ATP to support rapid growth	(Gomez-Requeni et al., 2011; Nagakawa and Nagayama, 1989; Reddish et al., 2008)	Zebrafish (<i>Danio rerio</i>) Red sea bream (<i>Pagrus major</i>) Pacific mackerel (<i>Scomber japonicas</i>) Carp (<i>Cyprinus carpio</i>) Yellow perch (<i>Perca flavescens</i>)
6		Creatine kinase	Glycolysis (kinase, transferase)	(Terova et al., 2014)	European sea bass (<i>Dicentrarchus labrax</i>)
7		GAPDH	Involved in glycolytic path ways which increase ATP to support rapid growth	(Gomez-Requeni et al., 2011; Nagakawa and Nagayama, 1989)	Zebrafish (<i>Danio rerio</i>) Red sea bream (<i>Pagrus major</i>) Pacific mackerel (<i>Scomber japonicas</i>) Carp (<i>Cyprinus carpio</i>)
8		Lactate dehydrogenase	Glycolysis (oxidoreductase)	(Terova et al., 2014)	European sea bass (<i>Dicentrarchus labrax</i>)
9		Adenylate kinase	Glycolysis (kinase, transferase)	(Terova et al., 2014)	European sea bass (<i>Dicentrarchus labrax</i>)
10	Slow-growing	SERCA	Pumping Ca ²⁺ to maintain muscle cell homeostasis and proliferation	(Magnier-Gaubil et al., 1996)	Human (<i>Homo sapiens</i>)

Table 2.3. List of the 10 most important proteins that are directly related to muscle metabolic activities out of the 49 proteins identified, which were extracted from frozen muscle tissues of European sea bass obtained by 2-D electrophoresis (2-DE) and tandem mass spectrometry (MS/MS) analysis. The protein identities are listed according to their accession number obtained from the UniProt database (Terova et al., 2014).

No	Identified protein	Species	Location	Biological function
1	Myosin heavy chain	Common carp <i>(Cyprinus carpio)</i>	Cytoplasm, thick filament	Muscle contraction
2	Myosin heavy chain	Red jungle fowl <i>(Gallus gallus)</i>	Cytoplasm, thick filament	Muscle contraction
3	Myosin heavy chain	Red jungle fowl <i>(Gallus gallus)</i>	Cytoplasm, thick filament	Muscle contraction
4	Myosin heavy chain	Red jungle fowl <i>(Gallus gallus)</i>	Cytoplasm, thick filament	Muscle contraction
5	Desmin	Cattle <i>(Bos taurus)</i>	Cytoplasm, intermediate filament	Muscle protein
6	Desmin	Zebra Fish <i>(Danio rerio)</i>	Cytoplasm, intermediate filament	Muscle protein
7	Troponin I	Human <i>(Homo sapiens)</i>	Cytoplasm, nucleus	Muscle protein
8	Myosin light chain 1	Chub mackerel <i>(Scomber japonicus)</i>	Cytoplasm, thick filament	Muscle contraction
9	Myosin regulatory light chain 2	Red jungle fowl <i>(Gallus gallus)</i>	Cytoplasm	Muscle protein

Table 2.3 (continued).

No	Identified protein	Species	Location	Biological function
10	Myosin light chain 2	Atlantic horse mackerel <i>(Trachurus trachurus)</i>	Cytoplasm	Muscle protein

Table 2.4. Hybrid striped bass growth parameters data for the poor-growth and good-growth groups. Values are provided for individual fish (Fish No.) along with mean values and SDE (standard error of the mean).

Fish No.	Growth Group	Total length (mm)	Wet Weight (g)	Gonadosomatic Index (GSI)	Hepatosomatic Index (HSI)	Condition Factor (K)	Specific Growth Rate (SGR)
1	Poor-Growth	360	470.4	5.49	2.30	1.01	2.29
2		330	516.7	14.64	4.94	1.44	2.48
3		342	511.0	14.98	3.44	1.28	2.45
4		345	537.4	6.30	2.91	1.31	2.55
5		330	461.2	9.04	2.67	1.28	2.26
6		319	391.5	11.81	3.26	1.21	1.94
7		335	537.7	14.97	4.86	1.43	2.55
8		294	299.8	3.27	2.54	1.18	1.42
9		338	525.6	7.07	4.12	1.36	2.51
10		340	569.6	17.03	3.22	1.45	2.66
Mean		333.3	482.09	10.46	3.43	1.29	2.31
SDE		5.54	25.77	1.53	0.29	0.04	0.12
1	Good-Growth	415	1100.4	16.83	4.89	1.54	3.94
2		380	875.0	13.05	5.31	1.59	3.50
3		400	929.5	13.56	4.80	1.45	3.61
4		392	914.8	14.62	5.43	1.52	3.58
5		402	1012.5	19.07	4.04	1.56	3.78
6		388	889.3	18.45	3.70	1.52	3.53
7		404	988.6	8.36	4.72	1.50	3.73
8		395	905.3	14.60	4.73	1.47	3.56
9		404	1017.9	19.99	3.93	1.54	3.79
10		394	929.0	16.75	4.26	1.52	3.61

Table 2.4 (continued).

Fish No.	Growth Group	Total length (mm)	Wet Weight (g)	Gonadosomatic Index (GSI)	Hepatosomatic Index (HSI)	Condition Factor (K)	Specific Growth Rate (SGR)
Mean		397.40	956.23	15.53	4.58	1.52	3.66
SDE		3.10	22.52	1.09	0.18	0.01	0.04

Table 2.7. Statistical comparisons between the poor- and good-growth groups identifying the number of significant metabolites found in the hybrid striped bass liver tissue using Welch’s Two-sample t-tests. Two thresholds were used: 1) for statistical significance, a $p \leq 0.05$ was used as a statistical cutoff, and 2) $0.05 < p < 0.10$ was used to determine a trend towards significance. The chart describes the total number of significantly different metabolites (109 in the good-growth group compared to the poor-growth group) with an indication of those metabolites that were elevated (30) and decreased (79) in the good-growth group. A gender effect was noted in the principle component analysis (Figure 2.7), so a statistical contrast was also provided for the female cohort. (F): denotes only female samples were compared.

Welch’s Two-Sample t-test	Growth Groups (Liver)	
	<u>Good-Growth</u> Poor-Growth	<u>Good-Growth (F)</u> Poor-Growth (F)
Total Metabolites; $p \leq 0.05$	109	88
Metabolites (↑ ↓)	30 79	24 64
Total Metabolites; $0.05 < p < 0.10$	43	49
Metabolites (↑ ↓)	18 25	13 36

Table 2.8. Statistical comparisons between the poor- and good-growth groups identifying the number of significant metabolites found in the hybrid striped bass muscle tissue using Welch’s Two-sample t-tests. Two thresholds were used: 1) for statistical significance, a $p \leq 0.05$ was used as a statistical cutoff, and 2) $0.05 < p < 0.10$ was used to determine a trend towards significance. The chart describes the total number of significantly different metabolites (50 in the good-growth group compared to the poor-growth group) with an indication of those metabolites that were elevated (23) and decreased (27) in the good-growth group. A gender effect was noted in the principle component analysis (Figure 2.8), so a statistical contrast was also provided for the female cohort. (F): denotes only female samples were compared.

Welch’s Two-Sample t-test	Growth Groups (Muscle)	
	<u>Good-Growth</u> Poor-Growth	<u>Good-Growth (F)</u> Poor-Growth (F)
Total Metabolites; $p \leq 0.05$	50	51
Metabolites (↑ ↓)	23 27	24 27
Total Metabolites; $0.05 < p < 0.10$	43	49
Metabolites (↑ ↓)	20 15	14 20

Table 2.9. Confusion matrix showing the predictive accuracy of the Random Forest analysis (Figure 2.7), conducted on hybrid striped bass liver tissues (N=9). Predictive accuracy was 89% and calculated in the shown table by comparing the predicted group versus actual group (Class Error).

Actual Group	Predicted Group		Class Error
	Good-Growth	Poor-Growth	
Good-Growth	9	0	100%
Poor-Growth	2	7	22.2%
Predictive Accuracy = 89%			

Table 2.10. Confusion matrix showing the predictive accuracy of the Random Forest analysis (Figure 2.8), conducted on hybrid striped bass muscle tissues (N=9). Predictive accuracy was 72% and calculated in the shown table by comparing the predicted group versus actual group (Class Error).

Actual Group	Predicted Group		Class Error
	Good-Growth	Poor-Growth	
Good-Growth	6	3	33.3%
Poor-Growth	2	7	22.2%
Predictive Accuracy = 72%			

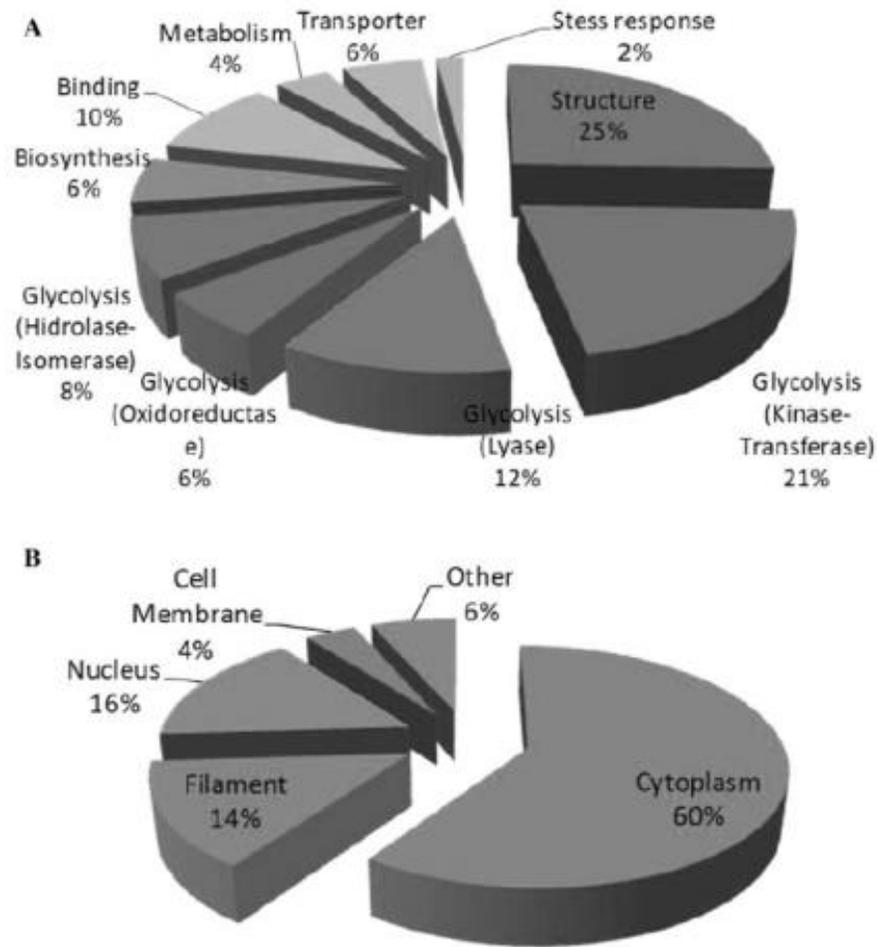


Figure 2.1. Pie chart distribution of protein identities directly related to muscle metabolic activities of 49 proteins identified related to muscle growth in European sea bass based on: A) Function of the protein, B) Location of the proteins (Terova et al., 2014).

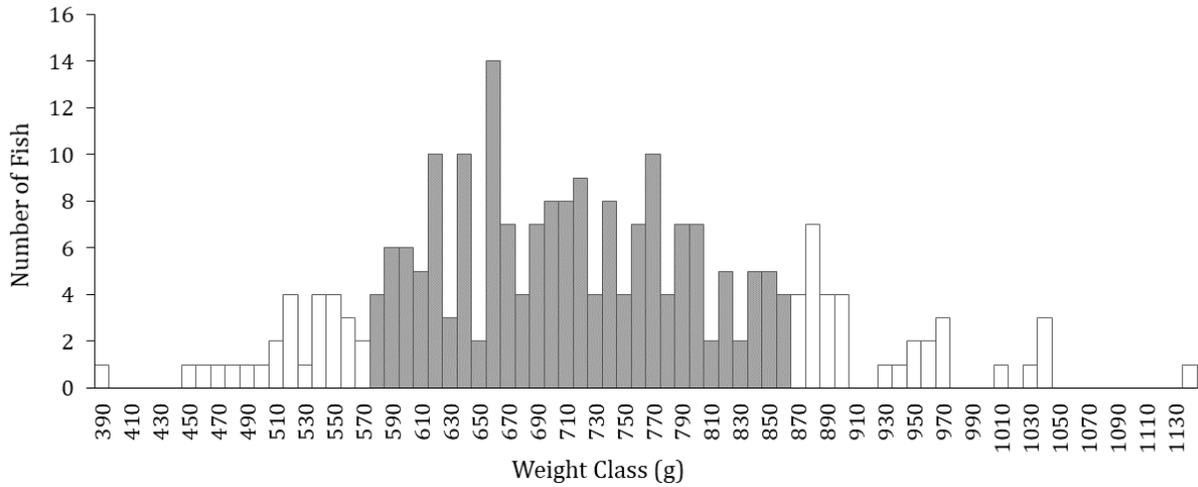
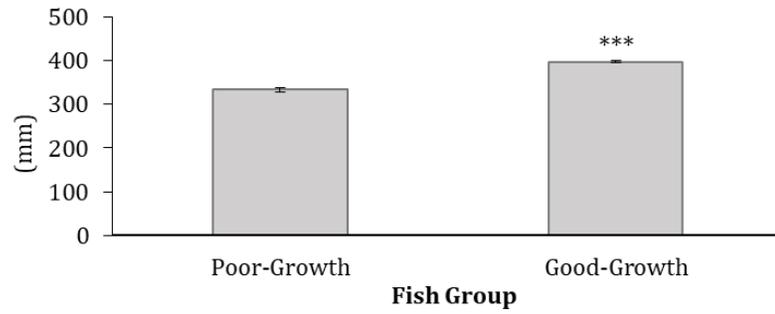


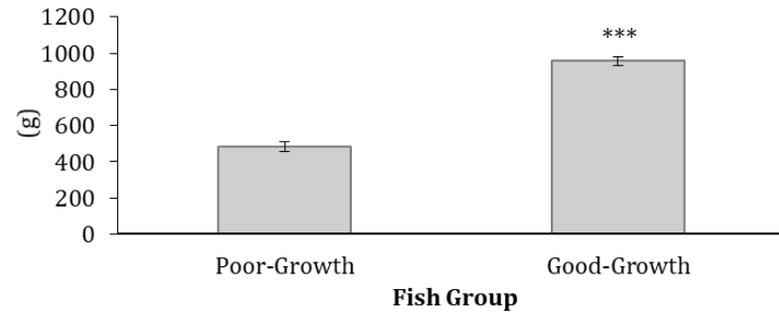
Figure 2.2. Bar charts showing the weight class (g) frequency in numbers of the hybrid striped bass. The white bars represent the top and bottom 10% in terms of body size (N=10 fish each group, 10 x 2 fish total) that were selected as representatives of fish that grow good (good-growth) or poor (poor-growth), respectively. These cutoffs were chosen as they represent a body size that is typically larger and smaller than that desired at market and therefore represent the extremes of the distribution.

Figure 2.3. Bar charts showing the (a) average total length (mm), (b) average wet weight (g), (c) average Gonadosomatic index (GSI), (d) average Hepatosomatic index (HSI), (e) average Condition Factor (K), and (f) Average Specific Growth Rate (SGR) of hybrid striped bass in each growth group (poor-growth or good-growth). Average values are represented by each bar and the standard error of the mean is shown by the brackets. There were ten fish in each group. There were significant differences between groups (One-way Student's t-test at an alpha level of 0.05; *** $p = 0.0001$; ** $p = 0.0037$; * $p = 0.0147$).

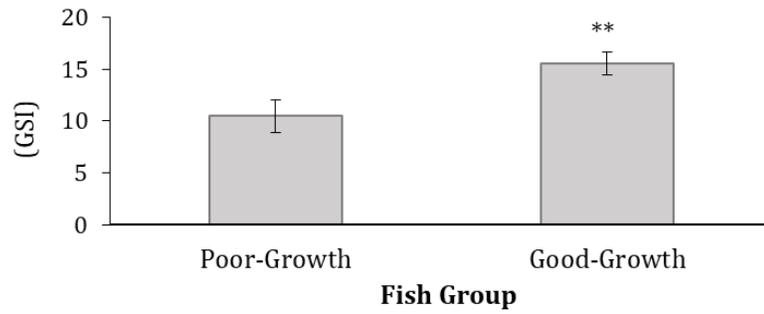
(a) Average total length (mm) of sunshine bass groups



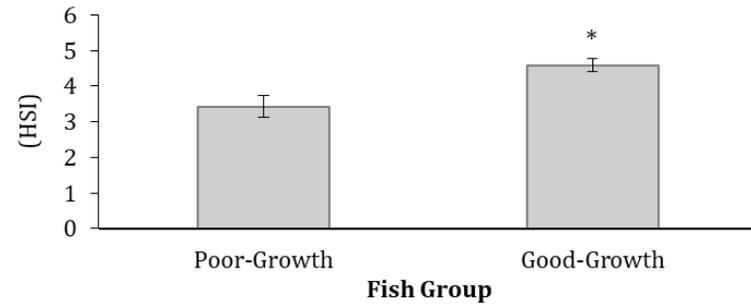
(b) Average wet weight (g) of sunshine bass groups



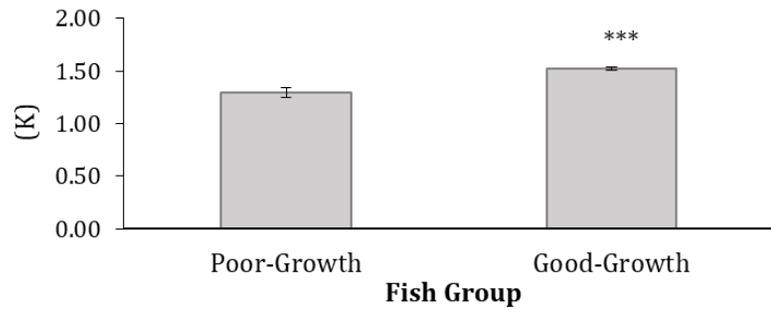
(c) Average Gonadosomatic index (GSI) of sunshine bass groups



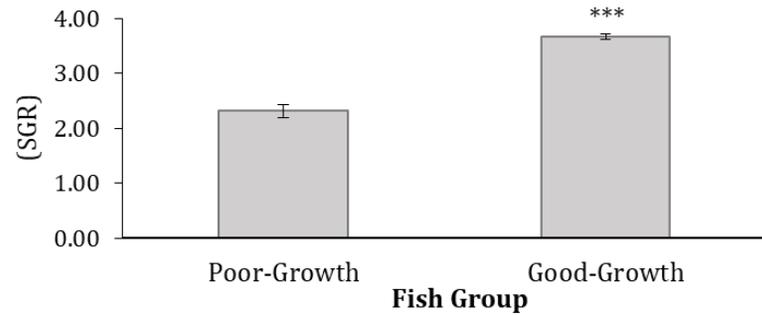
(d) Average Hepatosomatic index (HSI) of sunshine bass groups



(e) Average Condition Factor (K) of sunshine bass groups



(f) Average Specific Growth Rate (SGR) of sunshine bass groups



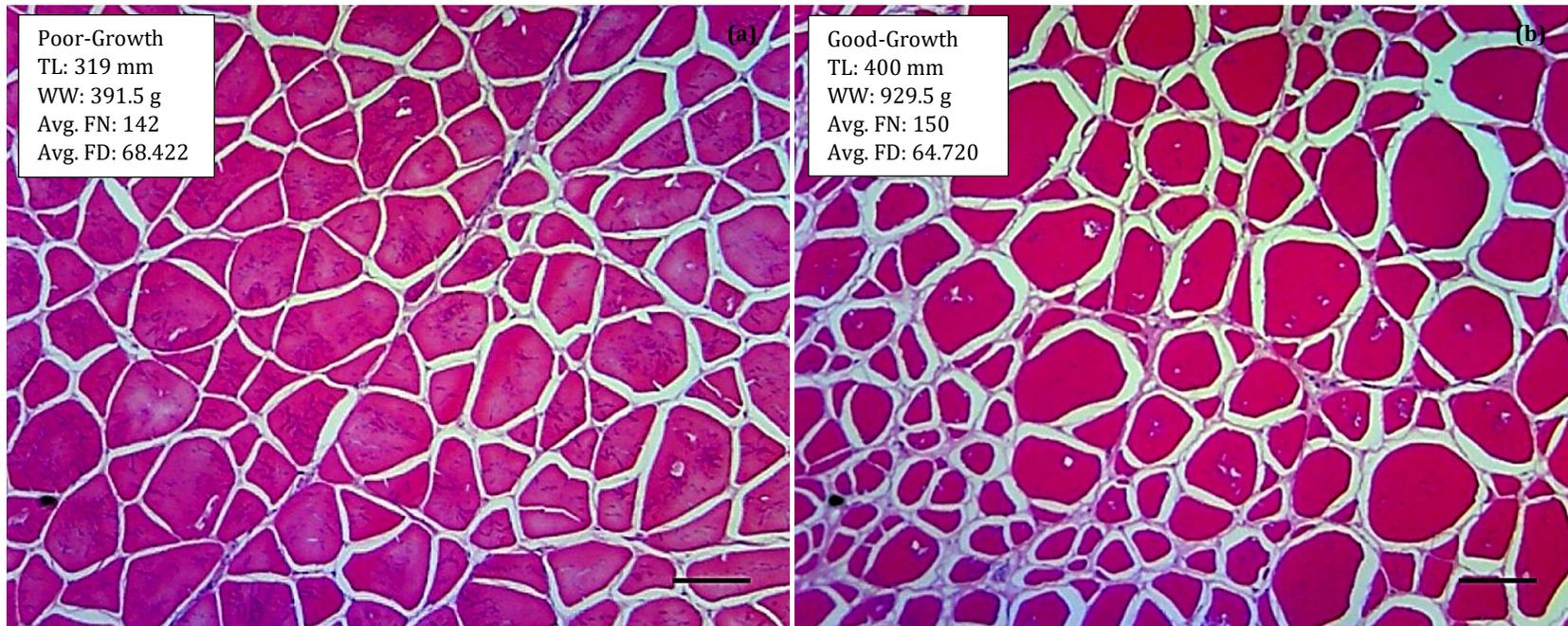
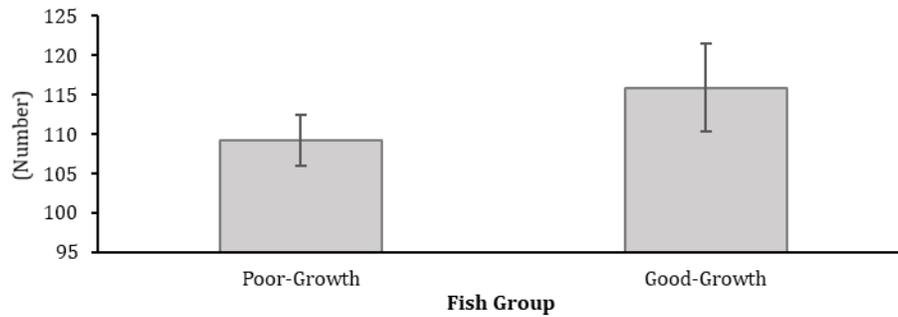
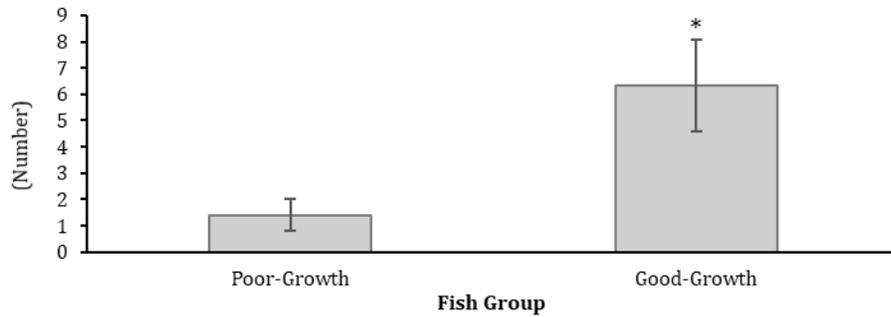


Figure 2.4. Representative histological cross-sections of hybrid striped bass (HSB) white muscle from fish in the poor- (a) and good- (b) growth groups. TL: Fish total length, WW: Fish Wet weight, Avg. FN: Average muscle fiber number, Avg. FD: Average muscle fiber diameter (μm). Comparing the two cross-sections, HSB in the poor-growth group have noticeably fewer, but larger muscle fibers, indicating hypertrophic growth status, whereas HSB in the good-growth group have more numerous but smaller muscle fibers, indicating hyperplastic growth. The bar corresponds to 100 μm .

(a) Average total muscle fiber number of sunshine bass groups



(b) Average muscle fiber number ($\leq 20 \mu\text{m}$) of sunshine bass groups



(c) Average muscle fiber diameter (μm) of sunshine bass groups

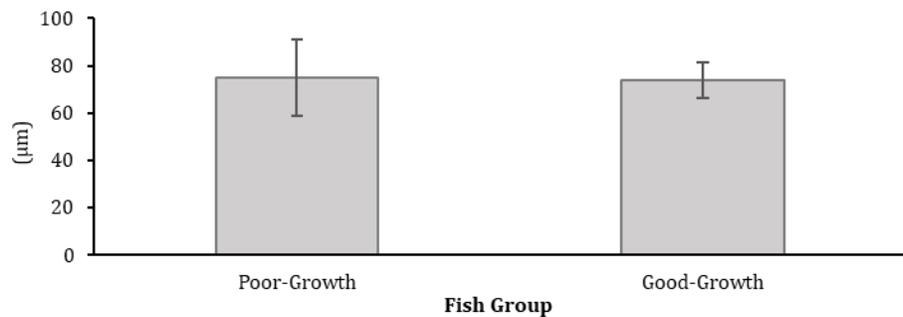
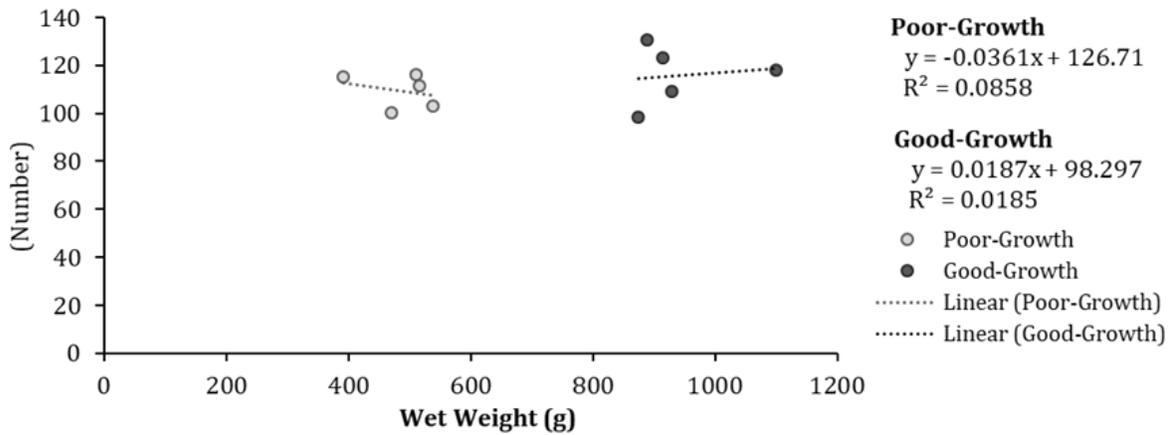


Figure 2.5. Bar charts showing the (a) average total muscle fiber number, (b) average muscle fiber number ($\leq 20 \mu\text{m}$), and (c) average muscle fiber diameter (μm) of sunshine hybrid striped bass in each growth group. The fish were divided into the poor-growth or good-growth groups. Average values are represented by each bar and the standard error of the mean is shown by the brackets. There were ten fish in each group. There were significant differences between groups (One-way Student's t-test at an alpha level of 0.05; * $p = 0.0274$).

(a) Average total muscle fiber number of sunshine bass groups



(b) Average muscle fiber diameter (μm) of sunshine bass groups

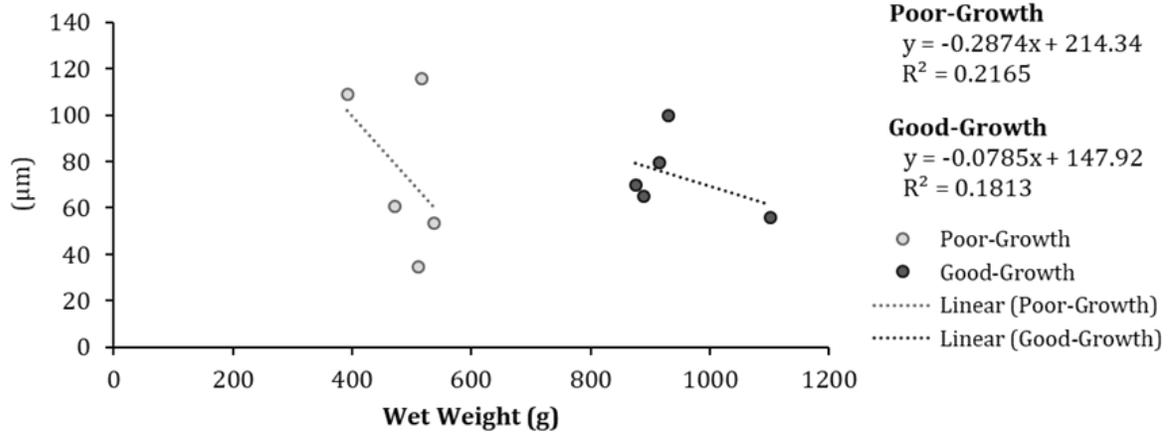


Figure 2.6. Linear regression graph between the: a) Average total muscle fiber number and hybrid striped bass wet weight (g) of the poor- and good- growth groups, b) Average muscle fiber diameter (μm) and the fish wet weight (g) of the poor- and good- growth groups. For each growth group, the linear regression equation is given as $y = mx + b$, where the slopes (m) of the lines represent the correlation between the variables in each growth groups, and correlation coefficient (R^2). A linear correlation between the Average total muscle fiber number and fish Wet Weight of the good- and poor-growth groups showed a weak correlation ($R^2=0.0858$ and 0.0185 , respectively) with little relationship. A linear correlation between the Average muscle fiber diameter and fish Wet Weight of the good- and poor-growth groups showed a weak correlation ($R^2=0.1813$ and 0.2165 , respectively).

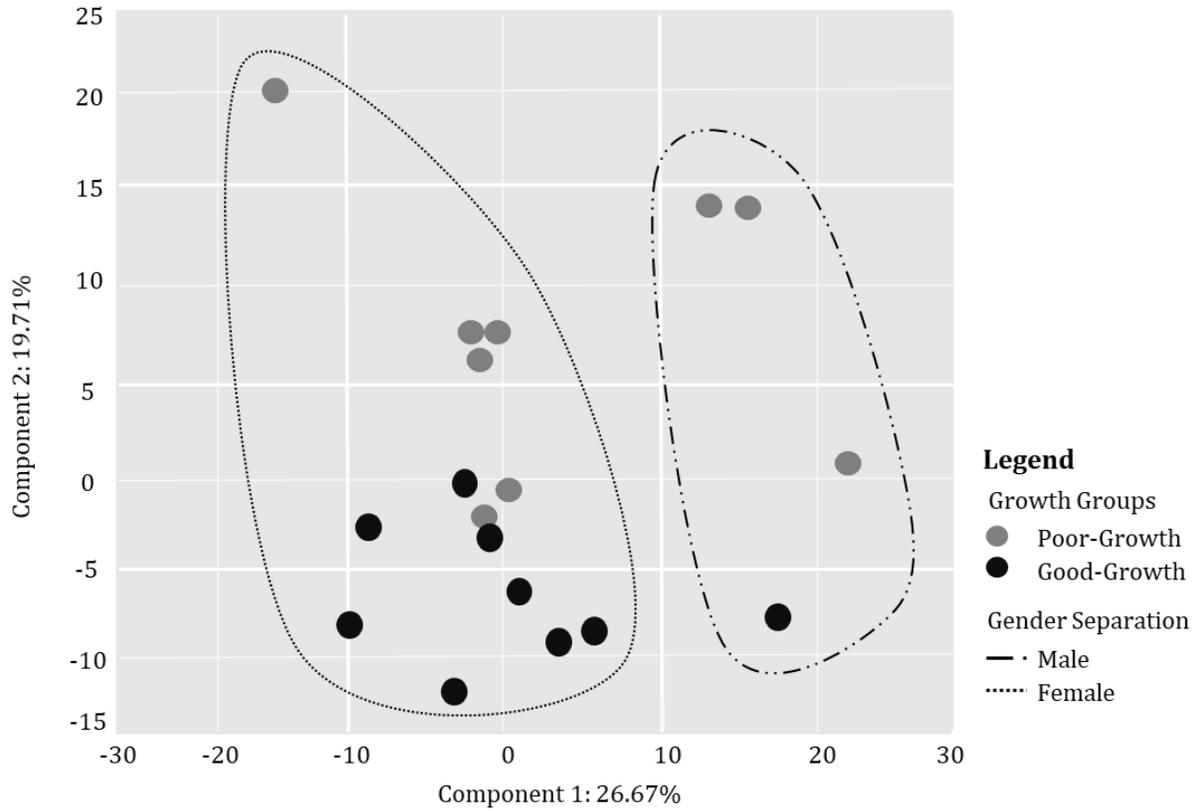


Figure 2.7. Principle Component Analysis (PCA) of hybrid striped bass liver tissue. Component 1 appeared to separate the growth groups based on gender. A statistical test is provided in Table 2.5 for female cohort to help discriminate gender-related effects in the dataset.

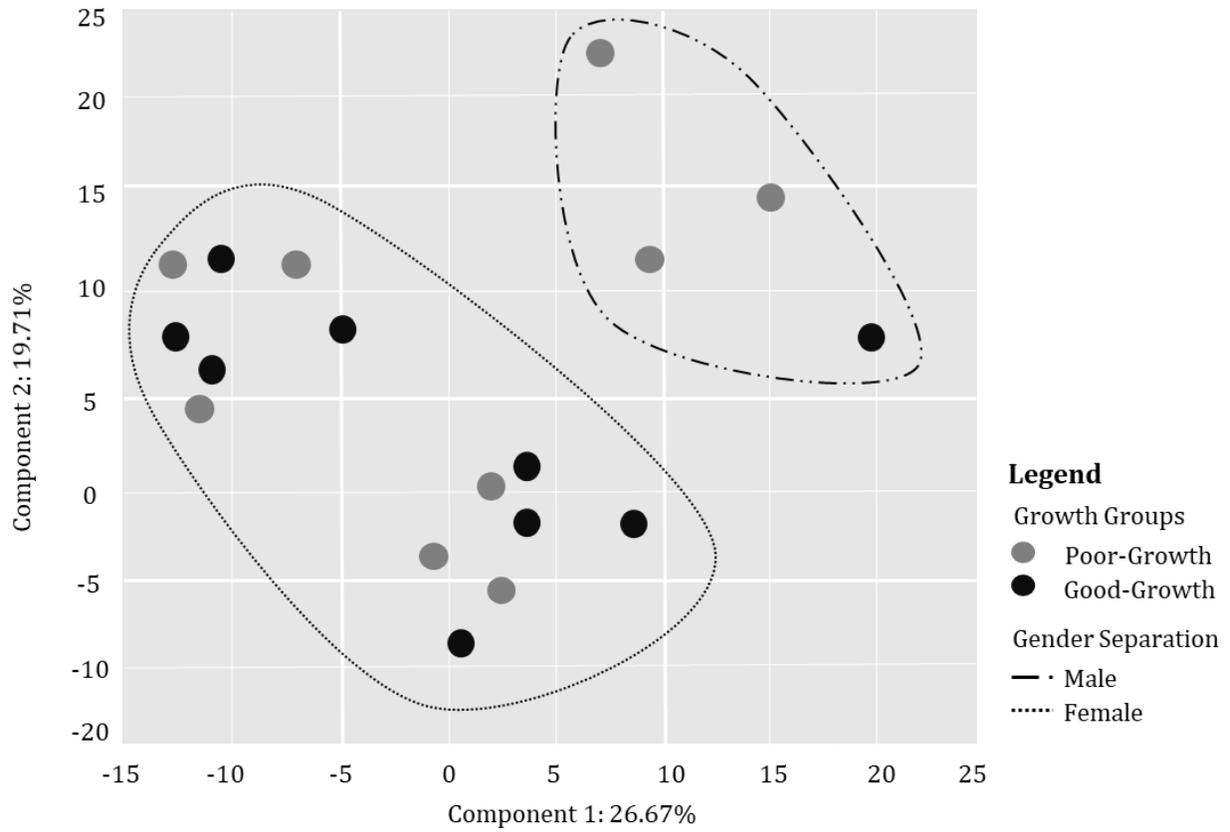


Figure 2.8. Principle Component Analysis (PCA) of hybrid striped bass muscle tissue. Component 1 appeared to separate the growth groups based on gender. A statistical test is provided in Table 2.6 for female cohort to help discriminate gender-related effects in the dataset.

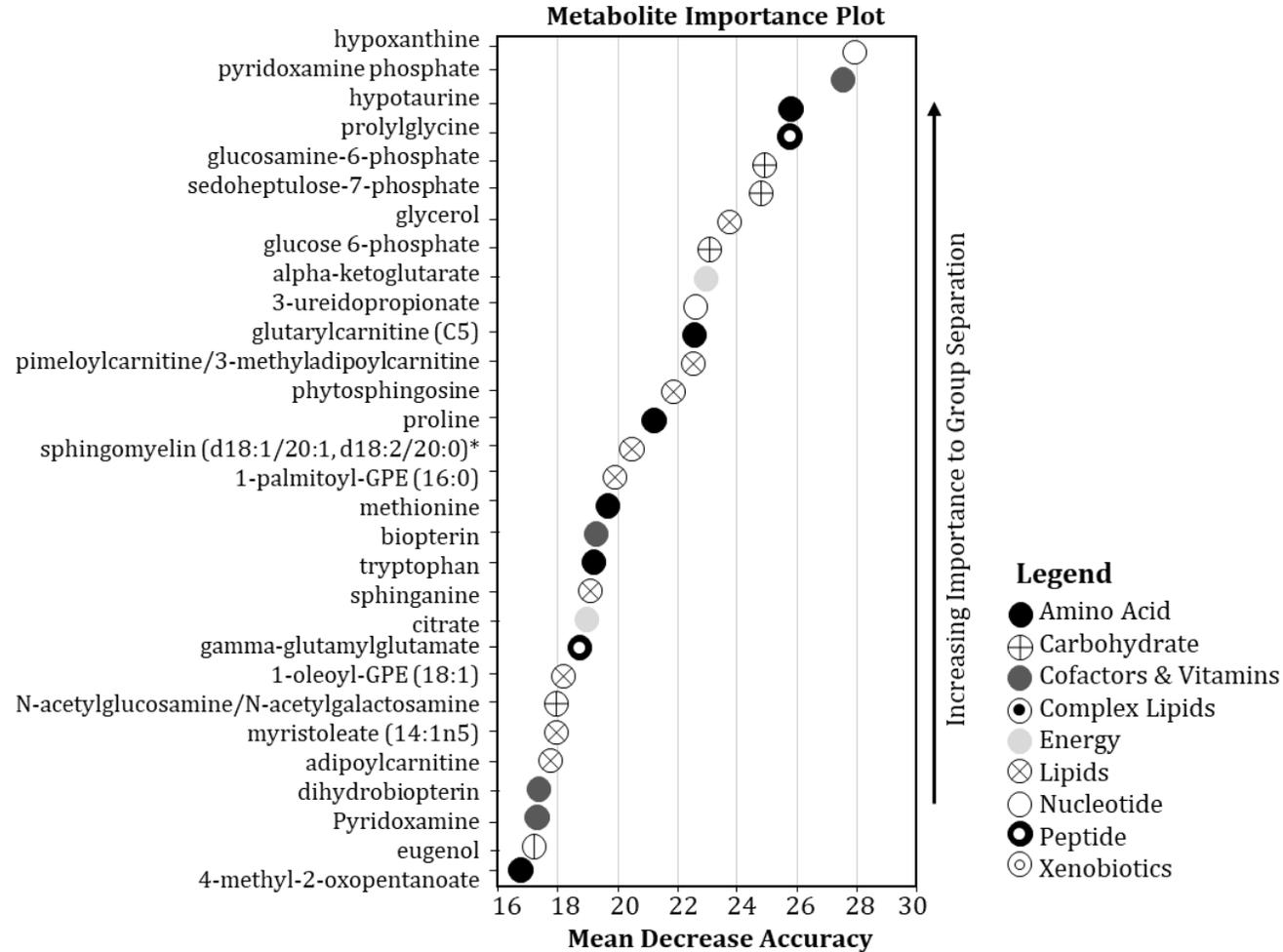


Figure 2.9. Random Forest analysis identifying the top thirty most important metabolites when comparing liver tissue of hybrid striped bass from the poor- and good-growth groups. Metabolites were ranked by increasing importance and mean decrease accuracy (sensitivity analysis) during Random Forest decision tree analysis. The legend describes the metabolic super-pathway to which each listed metabolite belongs. An asterisk designates metabolites with identifications that could not be confirmed based on a control standard, but Metabolon was confident in its identification.

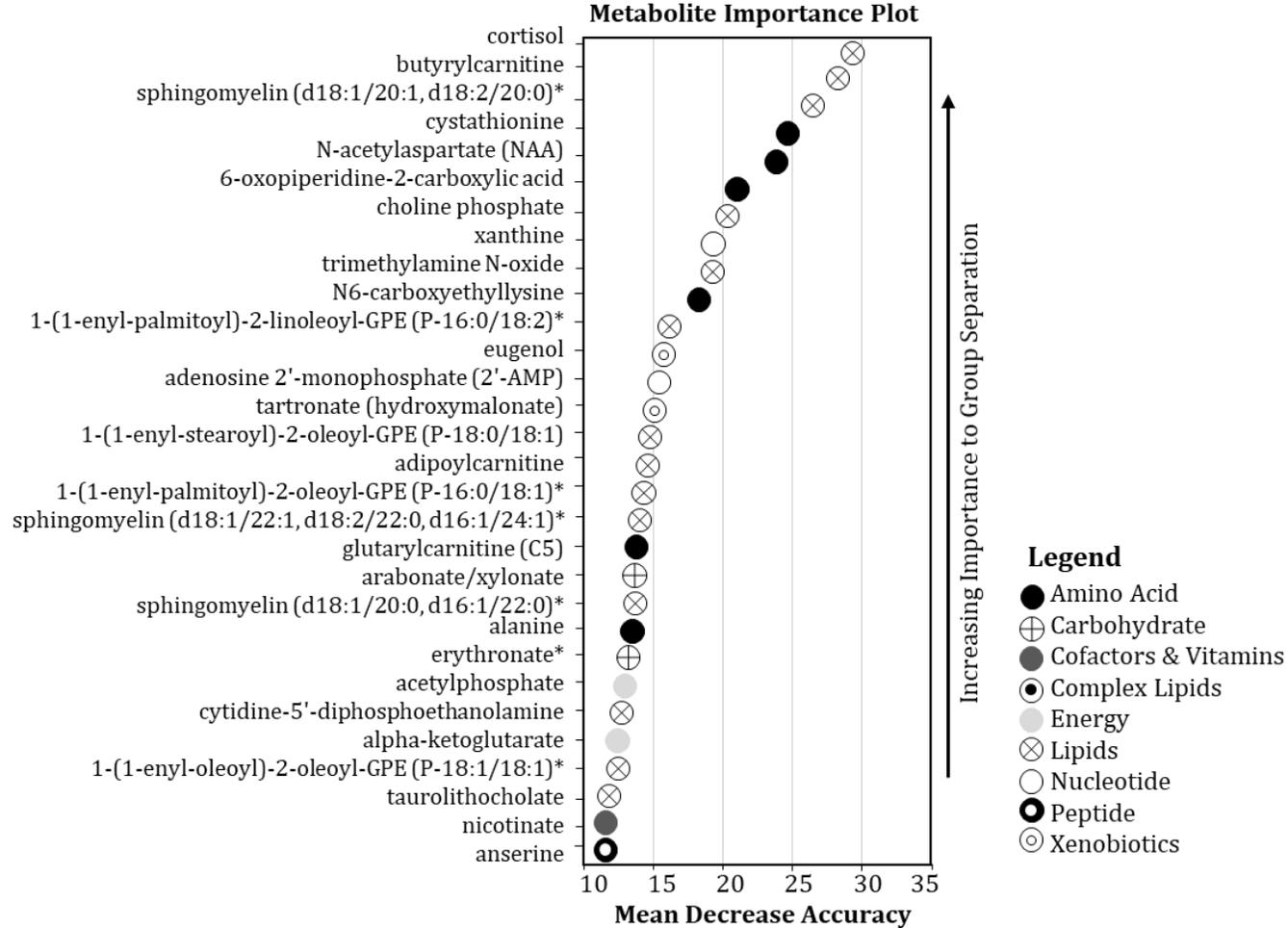


Figure 2.10. Random Forest analysis identifying the top thirty most important metabolites when comparing muscle tissue of hybrid striped bass from the poor- and good-growth groups. Metabolites were ranked by increasing importance and mean decrease accuracy (sensitivity analysis) during Random Forest decision tree analysis. The legend describes the metabolic super-pathway to which each listed metabolite belongs. An asterisk designates metabolites with identifications that could not be confirmed based on a control standard, but Metabolon was confident in its identification.

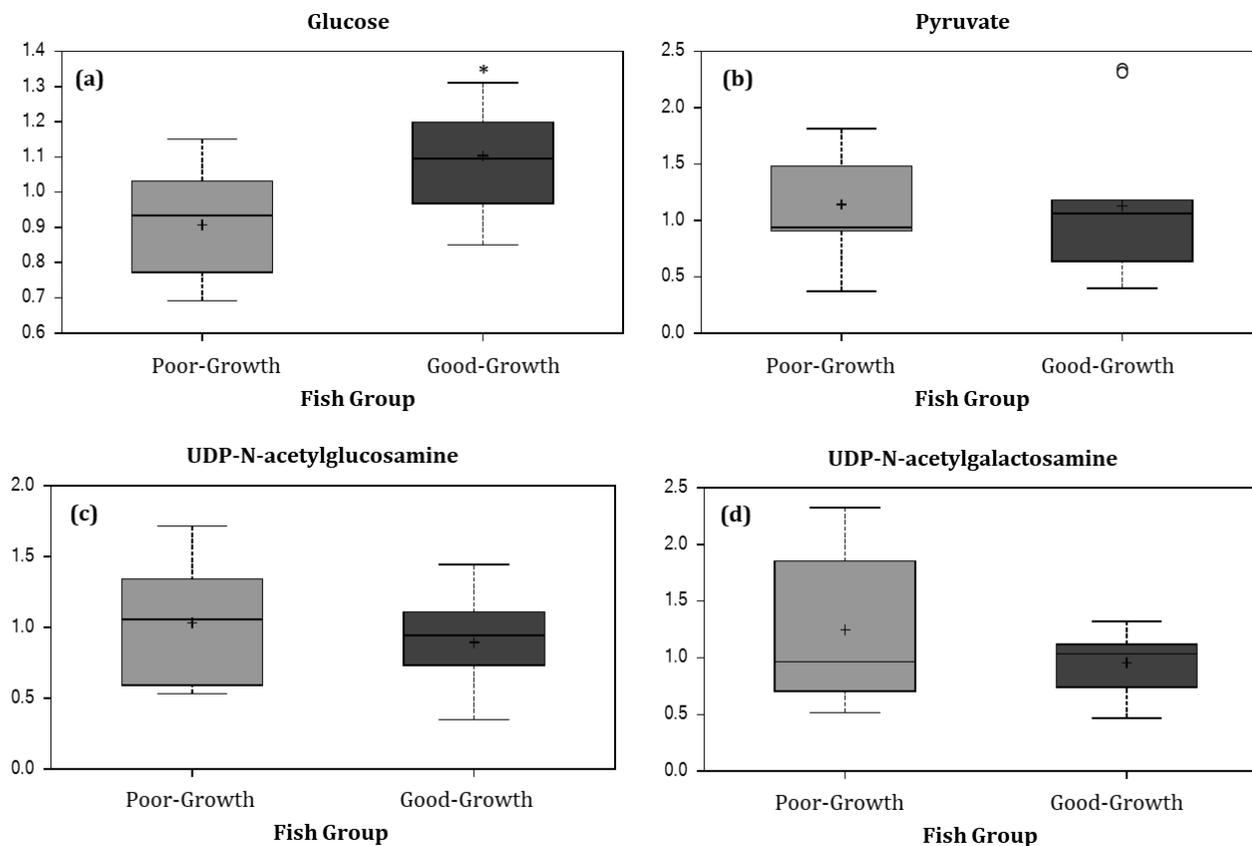


Figure 2.11. Boxplots showing (a) glucose, (b) pyruvate, (c) UDP-N-acetylglucosamine, and (d) UDP-N-acetylgalactosamine of scaled intensity range for poor- and good-growth groups of hybrid striped bass white muscle tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass muscle tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

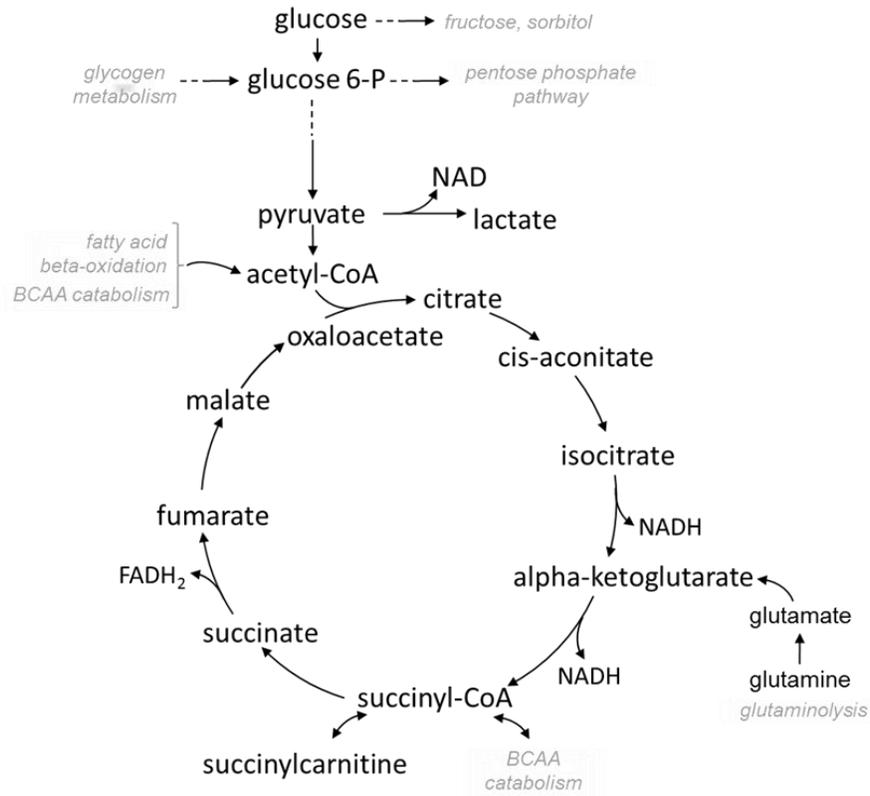


Figure 2.12. Schematic diagram of the tricarboxylic acid (TCA) cycle where the grey sub-pathways shown correspond to the metabolites identified as important for growth in the hybrid striped bass white muscle tissue.

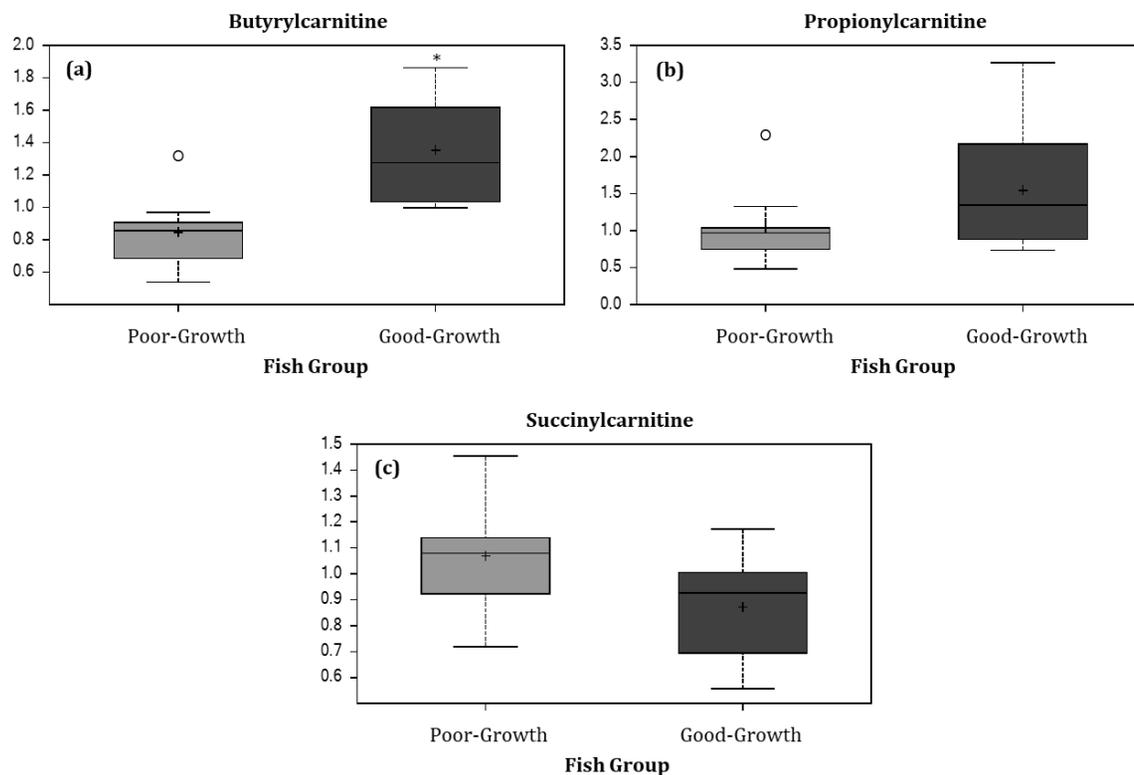


Figure 2.13. Boxplots showing (a) butyrylcarnitine, (b) propionylcarnitine, and (c) succinylcarnitine of scaled intensity range for poor- and good- growth groups of hybrid striped bass white muscle tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass muscle tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

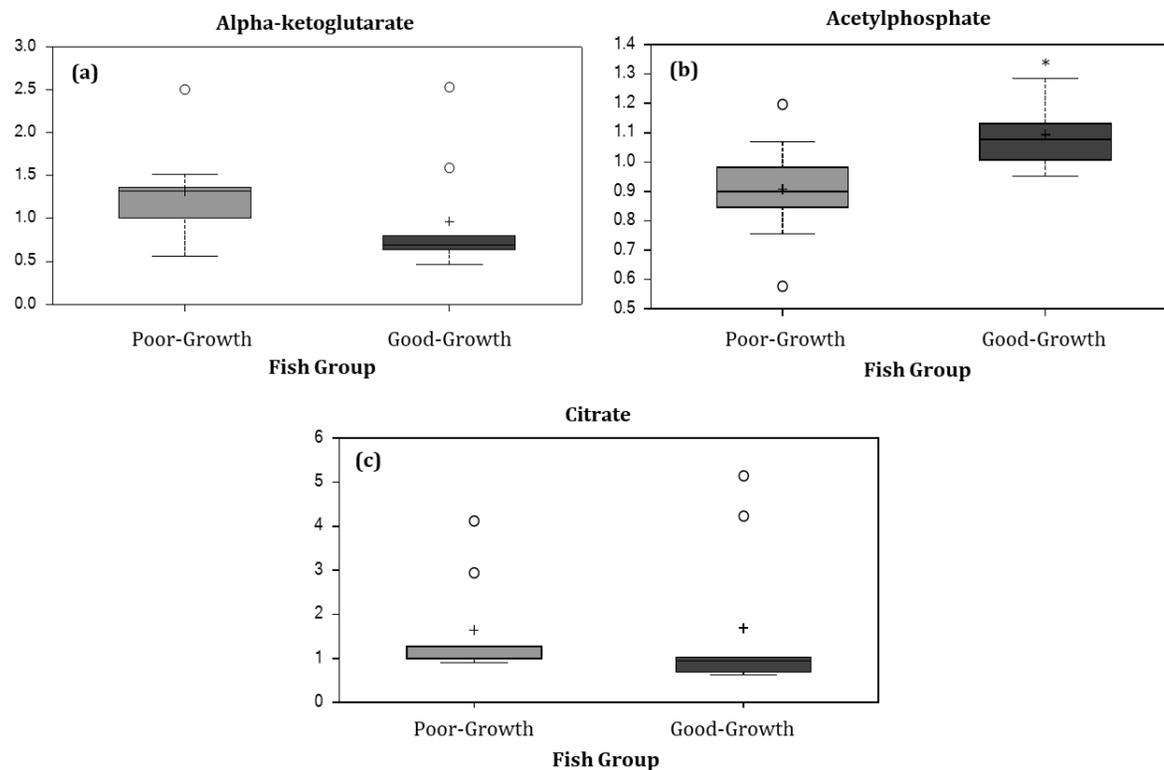


Figure 2.14. Boxplots showing (a) alpha-ketoglutarate, (b) acetylphosphate, and (c) citrate of scaled intensity range for poor- and good- growth groups of hybrid striped bass white muscle tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass muscle tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

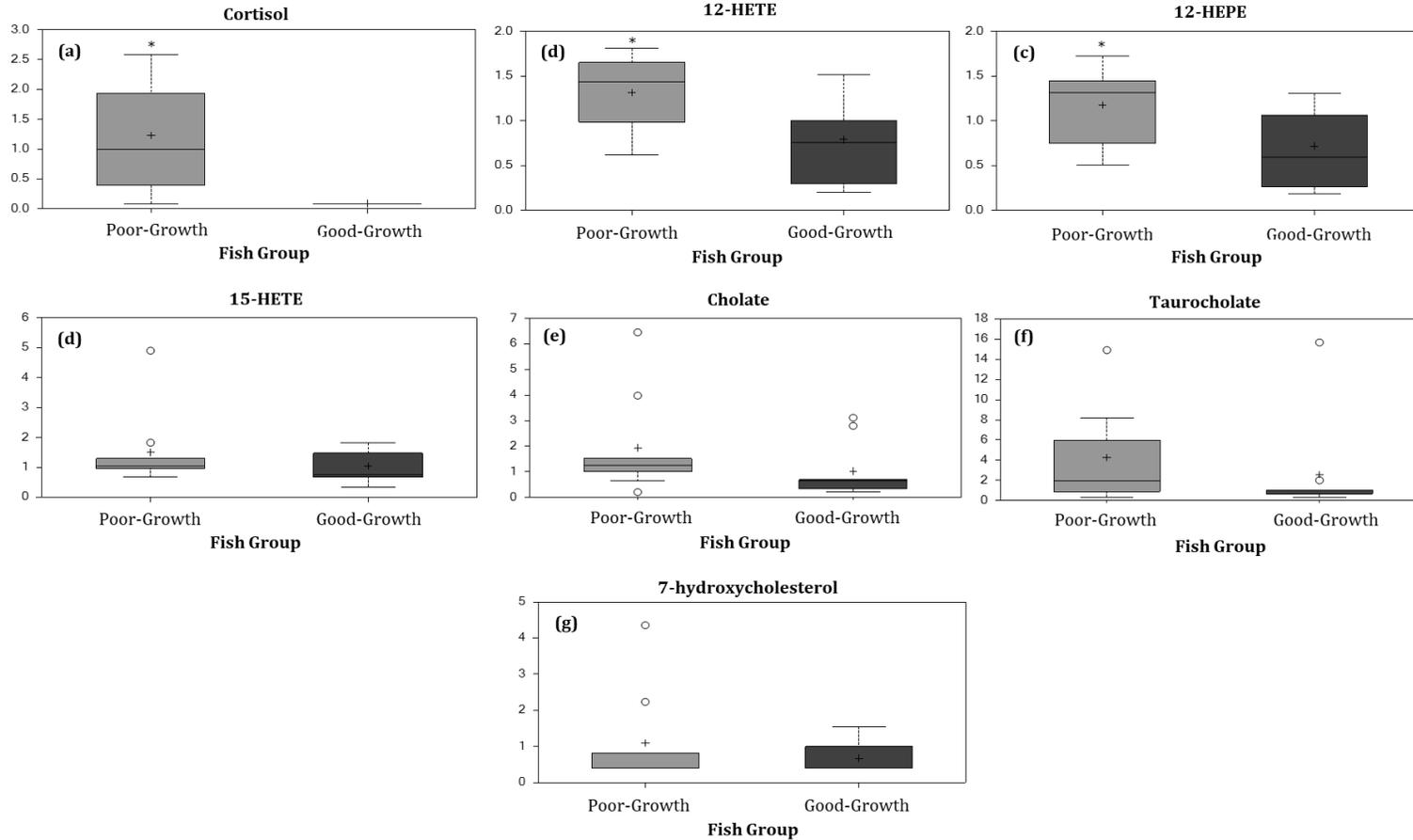


Figure 2.15. Boxplots showing (a) cortisol, (b) 12-HETE, (c) 12-HEPE, (d) 15-HETE, (e) chololate, (f) taurocholate, and (g) 7-hydroxycholesterol of scaled intensity range for poor- and good- growth groups of hybrid striped bass white muscle tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass muscle tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

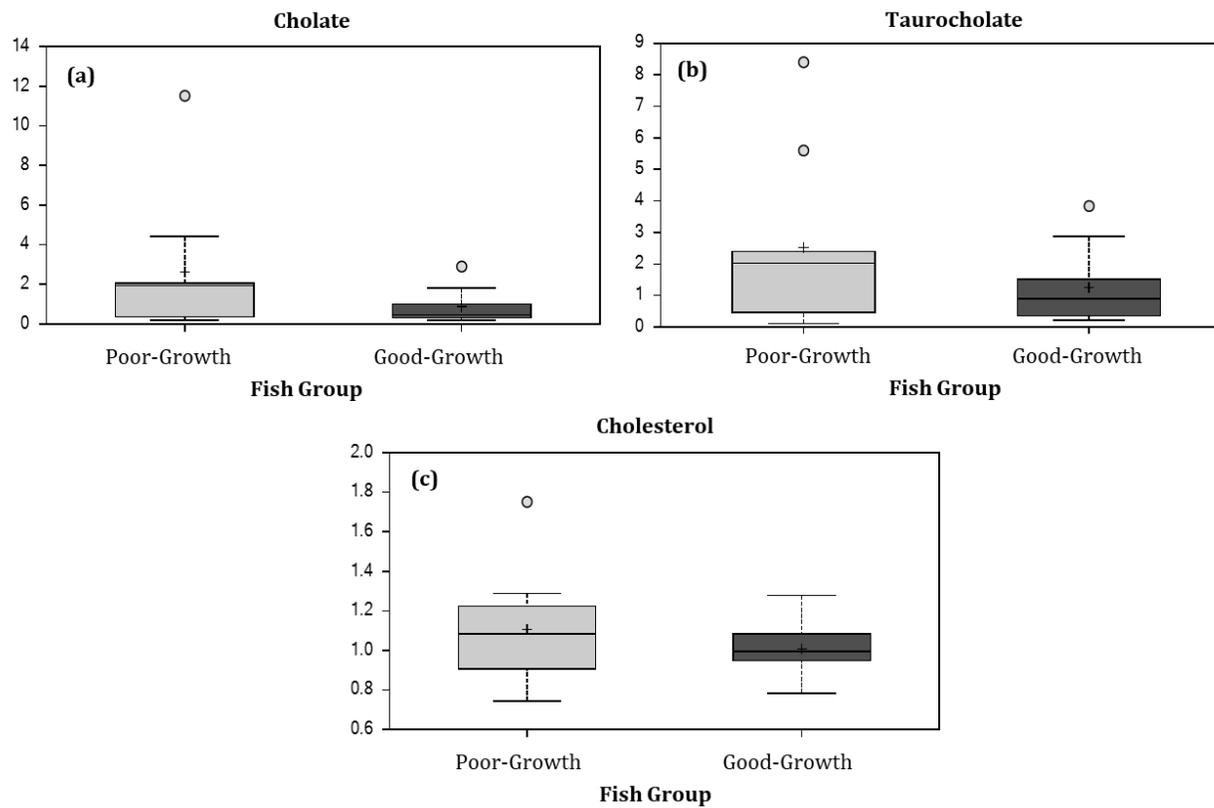


Figure 2.16. Boxplots showing (a) chololate, (b) taurocholate, and (c) cholesterol of scaled intensity range for poor- and good-growth groups of hybrid striped bass liver tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (there were no significant differences between groups, $\alpha = 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

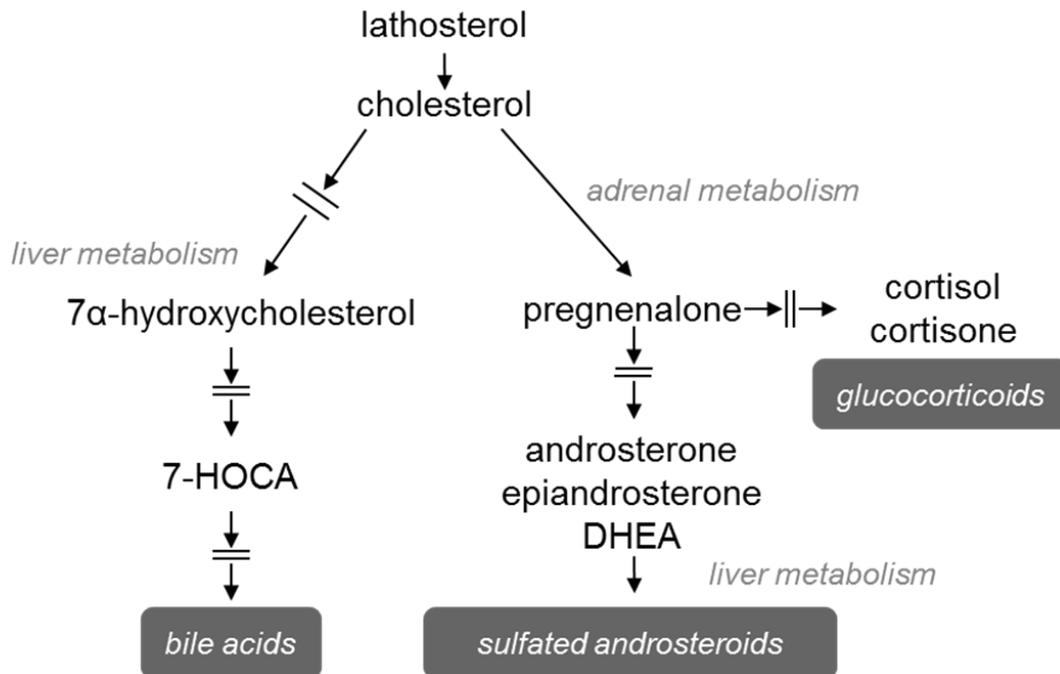


Figure 2.17. Schematic diagram of the cholesterol metabolic pathway. The schematic representation of the cholesterol biosynthetic pathway includes a number of cholesterol by-products including bile acids and cortisol. The pathway shown corresponds to the metabolites identified as important to growth in the hybrid striped bass liver tissue.

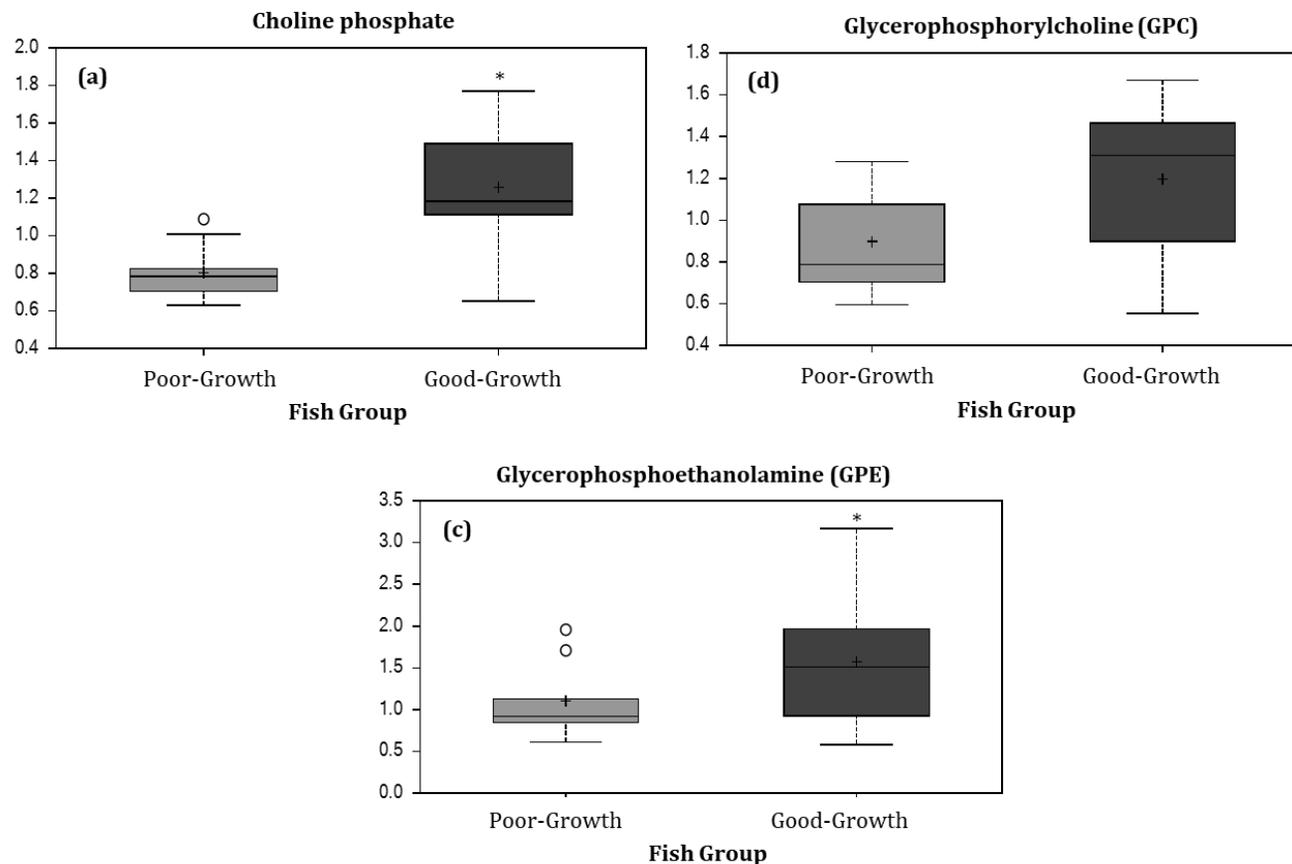


Figure 2.18. Boxplots showing (a) choline phosphate, (b) glycerophosphorylcholine (GPC), and (c) glycerophosphoethanolamine (GPE) of scaled intensity range for poor- and good- growth groups of hybrid striped bass white muscle tissue. Welch’s Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass muscle tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the “+” symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

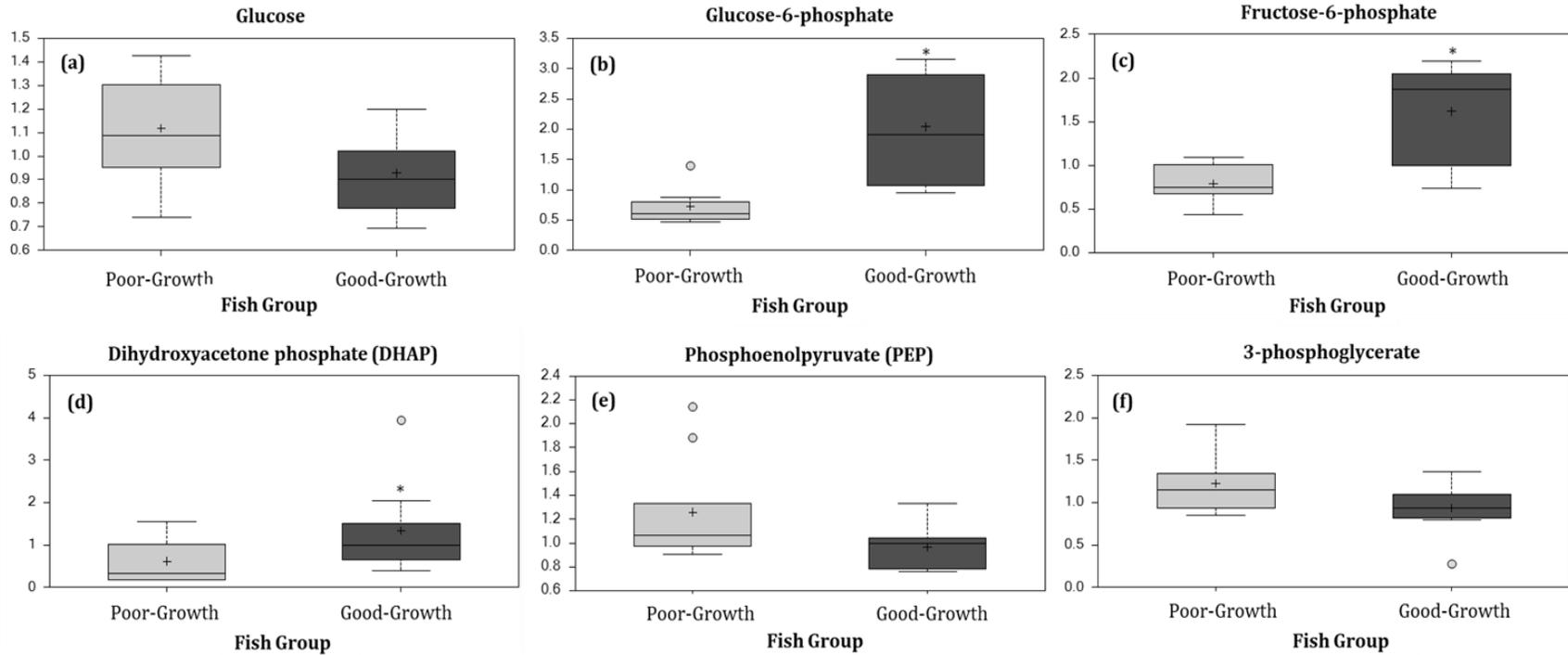


Figure 2.19. Boxplots showing (a) glucose, (b) glucose-6-phosphate, (c) fructose-6-phosphate, (d) dihydroxyacetone phosphate (DHAP), (e) phosphoenolpyruvate (PEP), and (f) 3-phosphoglycerate of scaled intensity range for poor- and good-growth groups of hybrid striped bass liver tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

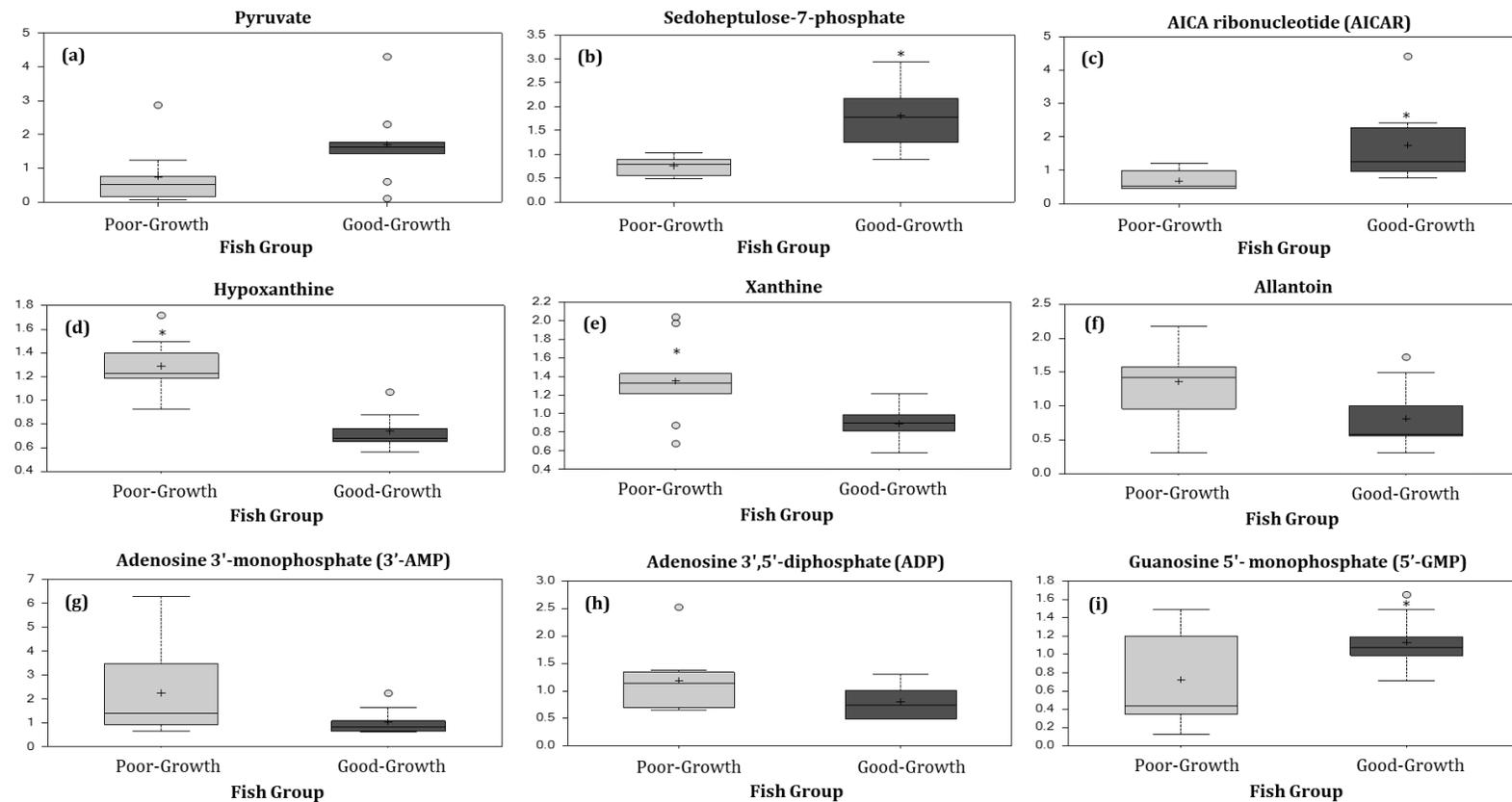


Figure 2.20. Boxplots showing (a) pyruvate, (b) sedoheptulose-7-phosphate, (c) AICA ribonucleotide (AICAR), (d) hypoxanthine, (e) xanthine, (f) allantoin, (g) adenosine 3'-monophosphate (3'-AMP), (h) adenosine 3',5'-diphosphate (ADP), and (i) guanosine 5'- monophosphate (5'-GMP) of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

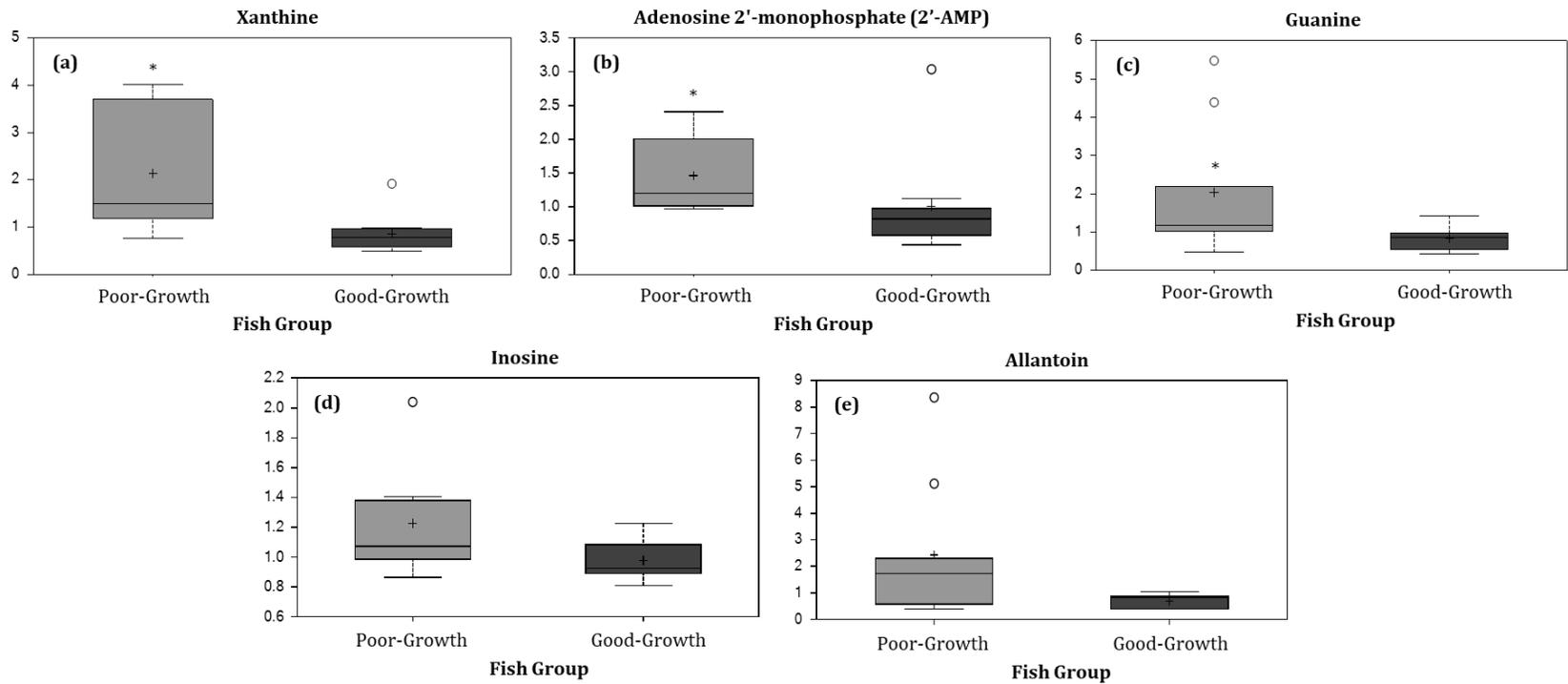


Figure 2.21. Boxplots showing (a) xanthine, (b) adenosine 2'-monophosphate (2'-AMP), (c) guanine, (d) inosine, and (e) allantoin of scaled intensity range for poor- and good- growth groups of hybrid striped bass white muscle tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass muscle tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

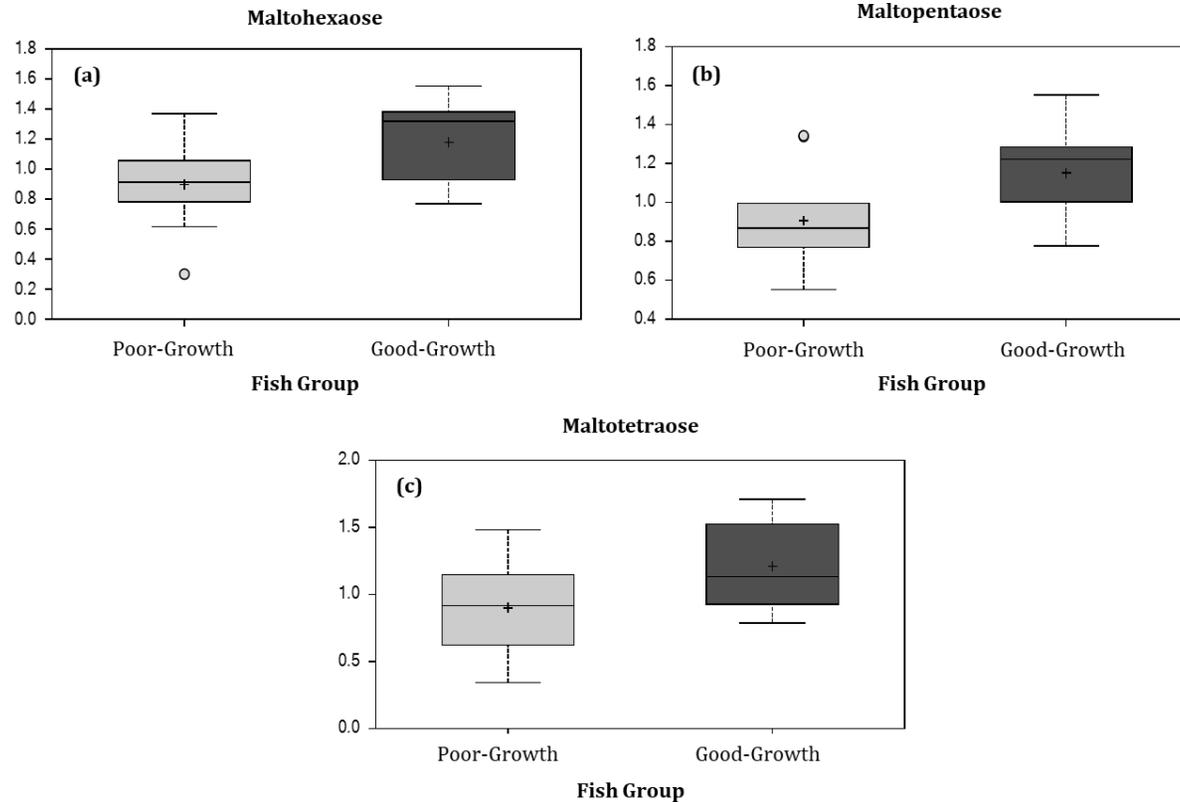


Figure 2.22. Boxplots showing (a) maltohexaose, (b) maltopentaose, and (c) maltotetraose of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch’s Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (there were no significant differences between groups, $\alpha = 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the “+” symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

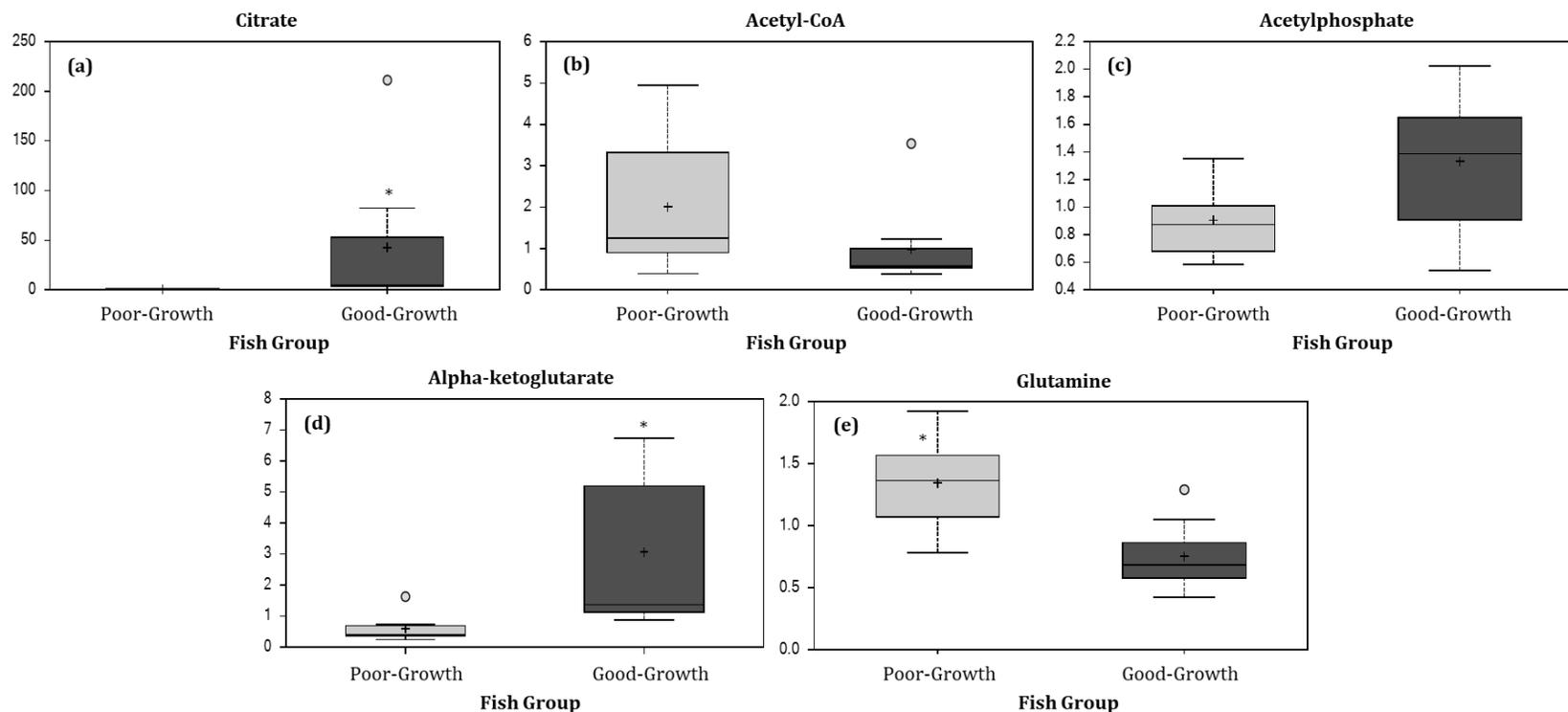


Figure 2.23. Boxplots showing (a) citrate, (b) acetyl-CoA, (c) acetylphosphate, (d) alpha-ketoglutarate, and (e) glutamine of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

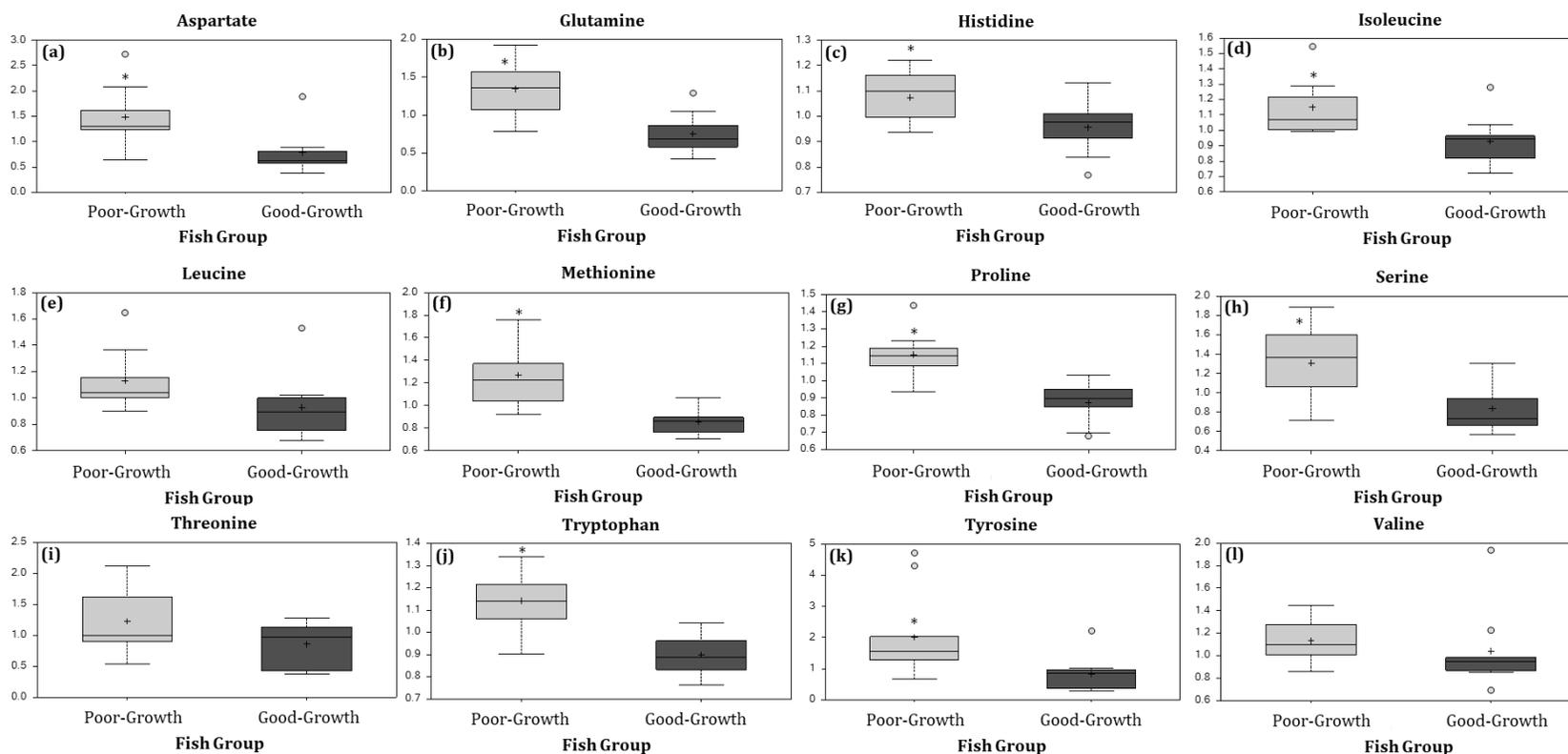


Figure 2.24. Boxplots showing a number of amino acids (a) aspartate, (b) glutamine, (c) histidine, (d) isoleucine, (e) leucine, (f) methionine, (g) proline, (h) serine, (i) threonine, (j) tryptophan, (k) tyrosine, and (l) valine of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

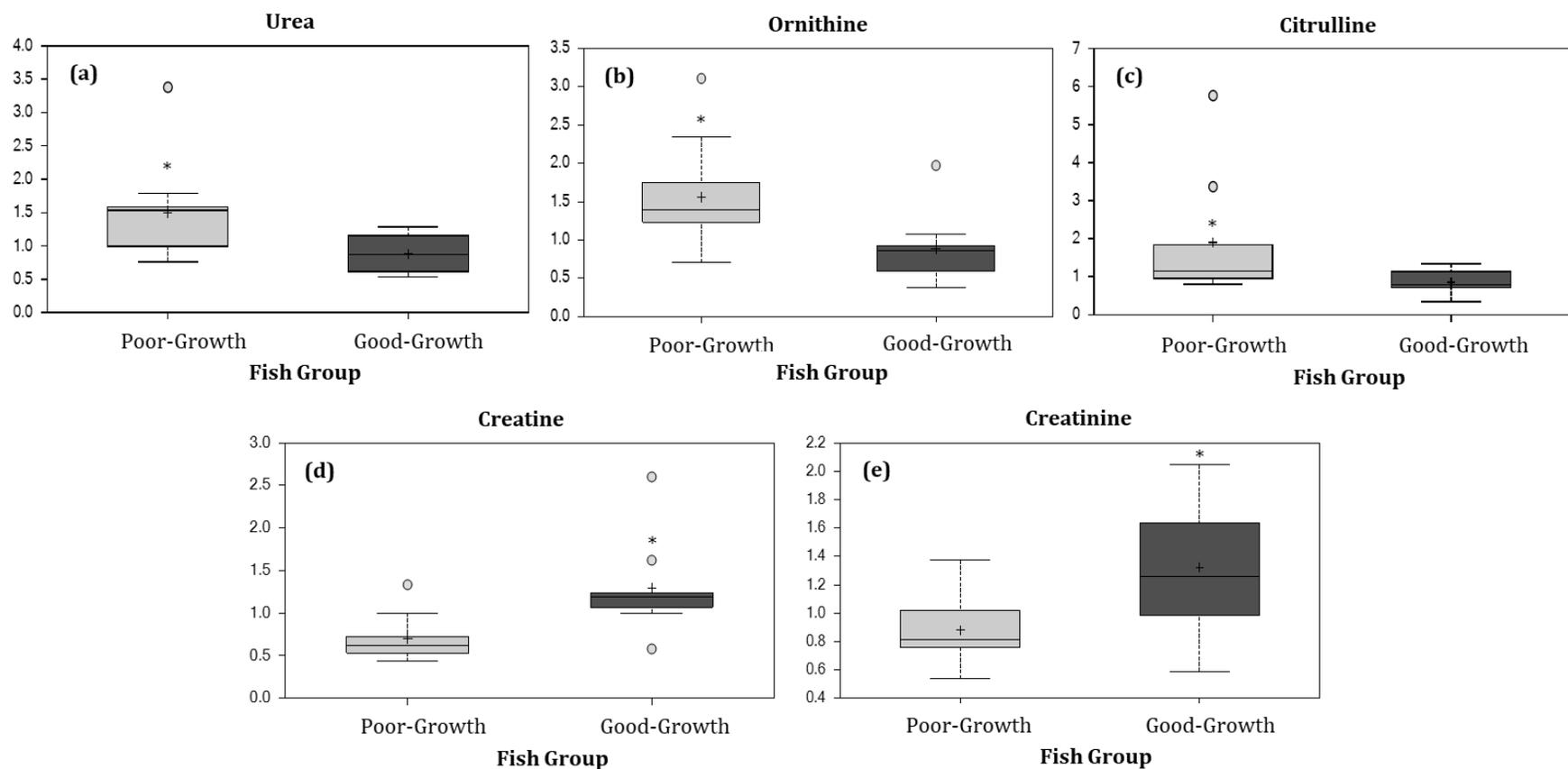


Figure 2.25. Boxplots showing (a) urea, (b) ornithine, (c) citrulline, (d) creatine, and (e) creatinine of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch's Two-sample t-tests comparison between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

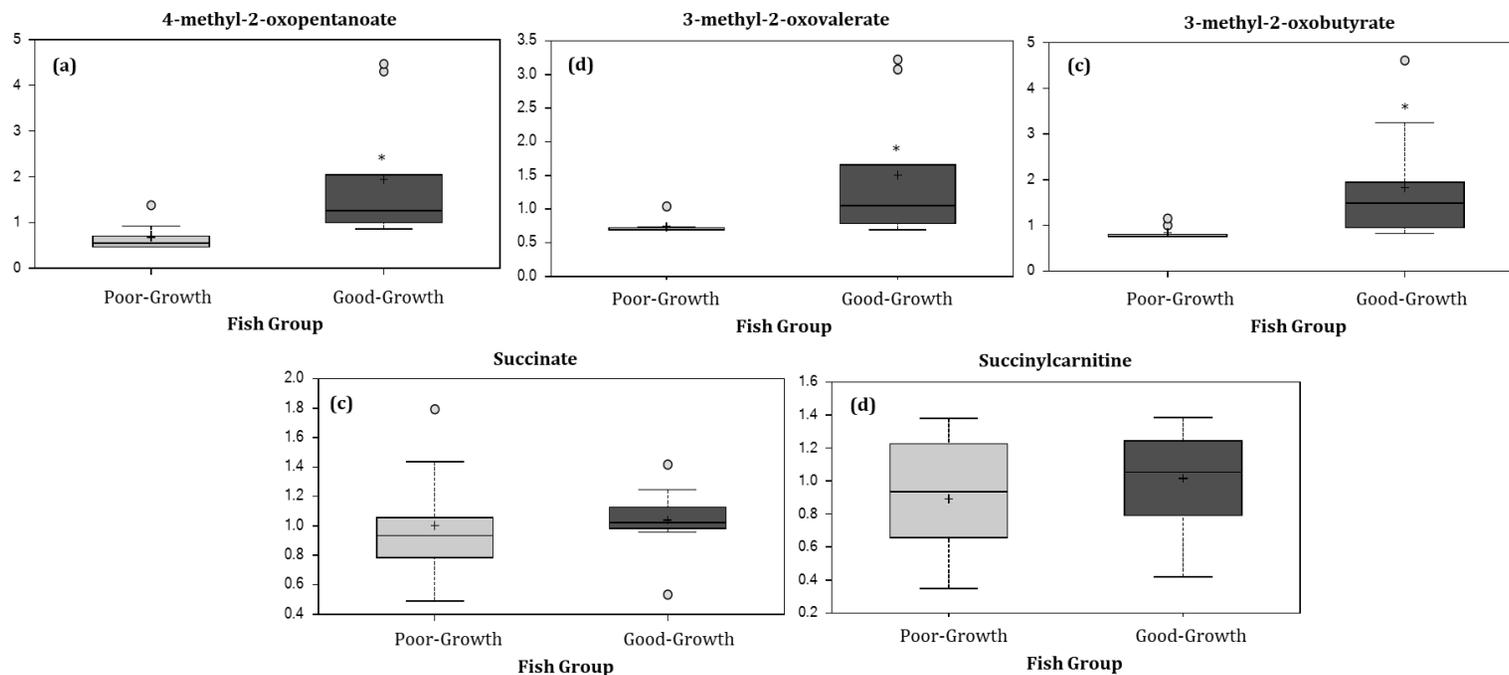


Figure 2.26. Boxplots showing (a) 4-methyl-2-oxopentanoate, (b) 3-methyl-2-oxovalerate, (c) 3-methyl-2-oxobutyrate, (d) succinate, and (e) succinylcarnitine of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch’s Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the “+” symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

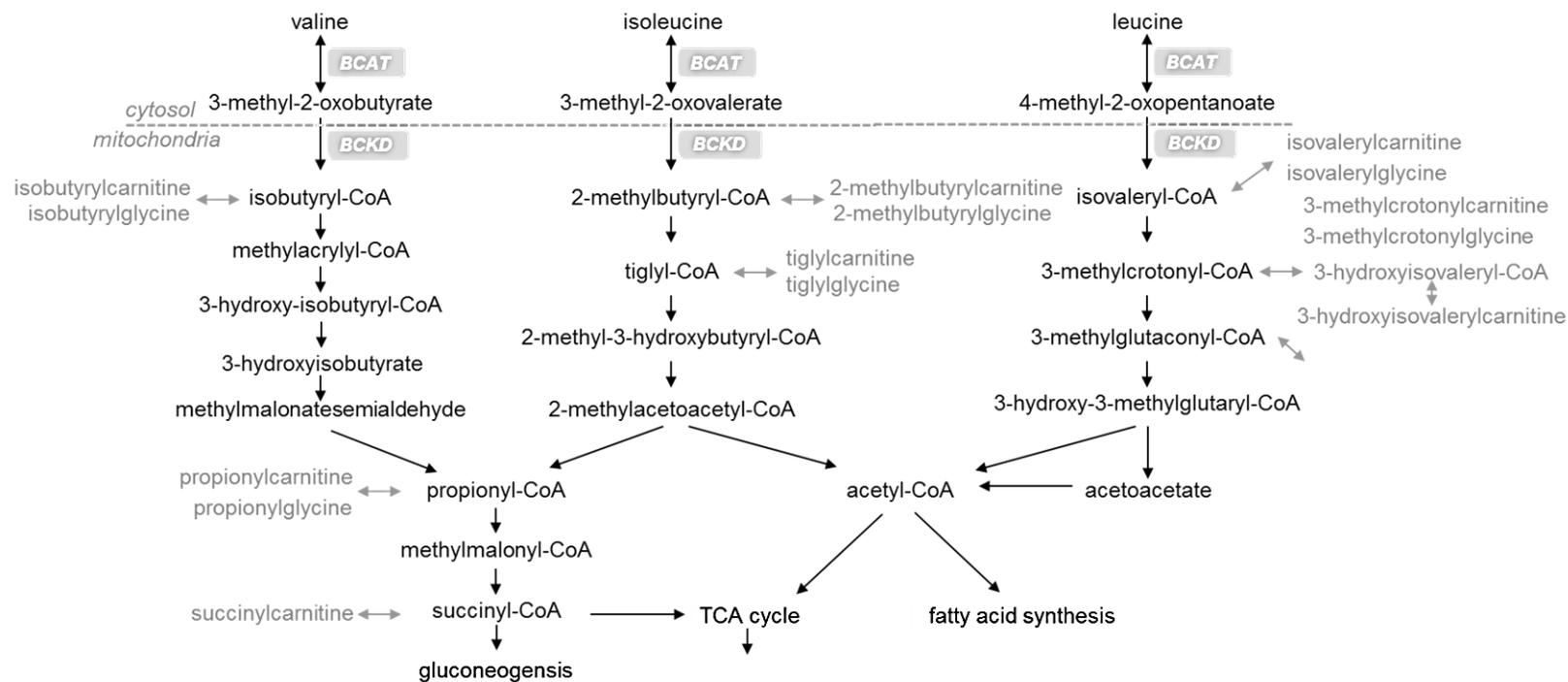


Figure 2.27. Schematic diagram for some of the branched chain amino acids (BCAA) (valine, isoleucine, leucine) catabolism pathways. Catabolized BCAA were assimilated for energy production through the tricarboxylic acid (TCA) cycle. Consistent with catabolic use, the keto-acids (4-methyl-2-oxopentanoate, 3-methyl-2-oxovalerate, and 3-methyl-2-oxobutyrate) were elevated. The pathway shown corresponds to metabolites identified as important to growth in the good-growth group hybrid striped bass liver tissue.

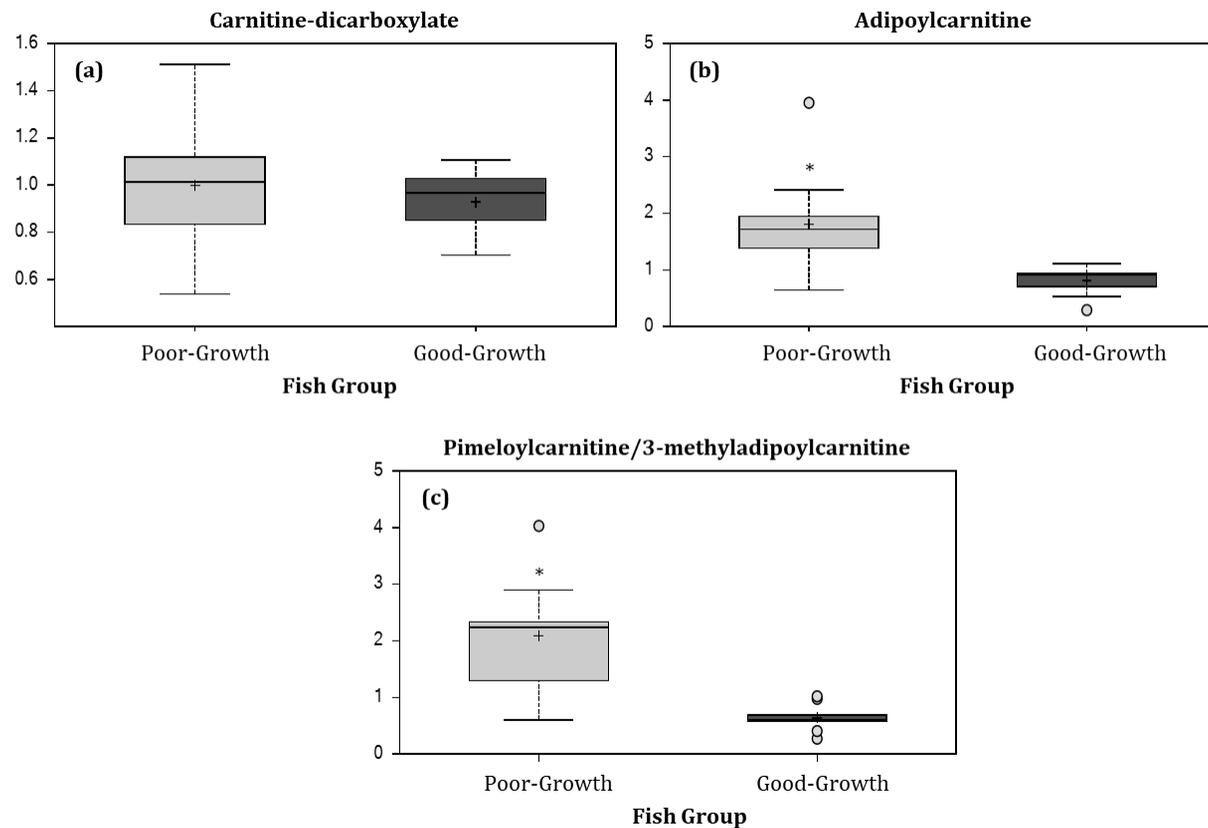


Figure 2.28. Boxplots showing (a) carnitine-dicarboxylate, (b) adipoylcarnitine, and (c) pimeloylcarnitine/3-methyladipoylcarnitine of scaled intensity range for poor- and good- growth groups of hybrid striped bass liver tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify the number of significant metabolites found in the hybrid striped bass liver tissue (* indicates significant difference between groups, $p \leq 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

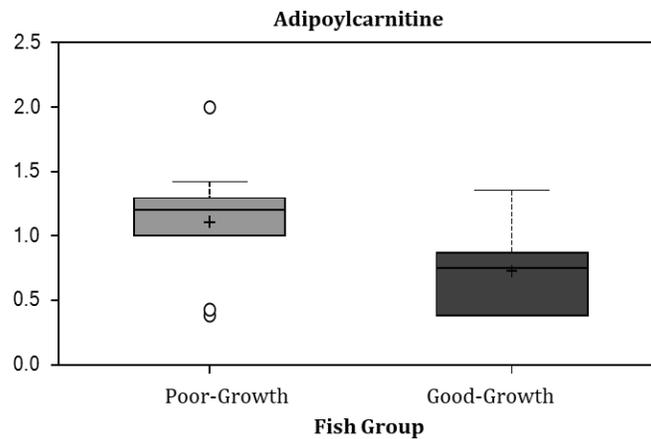


Figure 2.29. Boxplot showing adipoylcarnitine of scaled intensity range for poor- and good- growth groups of hybrid striped bass white muscle tissue. Welch's Two-sample t-test comparisons between the growth groups was conducted to identify significant metabolites found in the hybrid striped bass muscle tissue (there was no significant difference between groups, $\alpha = 0.05$). There were nine fish in each group. Median value is shown as the line through each box. Mean value is represented as the "+" symbol; potential outliers are indicated with a small circle. Upper and lower quartile range are the top and bottom of each box. Maximum and minimum distribution for each group is designated by the top and bottom whiskers.

Chapter III: White Muscle Genes Involved In Enhanced Growth Response of Hybrid Striped Bass

Abstract

In every cohort of cultured fish, a subgroup of individuals fails to reach the desired market size. The aim of this study was to provide an in-depth view of the transcriptome of white skeletal muscle of the hybrid striped bass (HSB) with the specific objective of providing an informative list of the most important genetic indicators involved in HSB growth performance. HSB fingerlings were reared common garden in tanks and ponds according to standard two-phase aquaculture procedures until harvest at market size (average weight: 721 ± 8 g, all values are given as average \pm standard deviation). Individuals representing the top and bottom 10% in terms of body size (N=10 fish each) were selected as representatives of fish that grow good (total length: 397 ± 3 mm, weight: 956 ± 23 g) or poor (total length: 333 ± 6 mm, weight: 482 ± 26 g). These cutoffs were chosen as they represent the extremes of the harvest distribution, including those fish that did not reach market size. White muscle samples were excised from the left side of HSB just ventral to the anterior margin of the first dorsal fin and posterior to the head. Transverse cross-sections were taken through the tissue with the purpose of evaluating differences in muscle fiber number, distribution, and diameter. Most of the muscle fibers in the good growth group were at a hyperplastic growth phase, whereas fish in the poor-growth group had switched to hypertrophic growth. RNA-seq analysis of HSB muscle was conducted using Illumina next generation sequencing. The transcriptomes were assembled from both *de novo* transcriptomes and existing whole-genome resources that have been developed for *Morone*. A novel Support Vector Machines based-machine learning analysis was used to

compare 72,893 different expressed genes between fish from the good- and poor- growth groups. SVMAttributeEval pattern recognition identified 150 as important gene transcripts and DAVID Gene Ontology (GO) enrichment organized these genes into three functional groups. The gene members in these groups share commonality in biological function related to growth performance, including regulatory functions such as cell signaling, and second messenger systems associated with the plasma membrane and cell proliferation and differentiation. Not surprisingly, genes involved with insulin-like growth factor and other important growth regulators also were identified.

Introduction

An overview of transcriptomics

Transcriptomics studies are a set of approaches to isolate all of the RNA transcripts (i.e. transcriptome) from a cell or a population of cells comprising a tissue such as muscle, and then to determine the presence and quantity of each particular gene-encoding messenger RNA (mRNA) (Gupta, 2014; McLean, 2013; Mortazavi et al., 2008; Wang et al., 2009). Namely, transcriptomics is viewed as a high-throughput technology for determining how the transcriptome changes with respect to various factors at a certain time point and at a given biological state (Gupta, 2014). For example, the regulation of gene expression that determines organismal growth, which has the ability to adapt rapidly and with tremendous variability in different tissues (e.g., muscle and liver) in response to stimuli (e.g., food consumption), is highly complex and underlies many fundamental biological processes (Sandvik et al., 2006). Transcriptomics and gene expression are powerful tools used extensively in many scientific fields, especially when it comes to comparing genes

expressed in certain groups of interest exposed to different treatments (Eun et al., 2008; Gibb et al., 2011). Findings of such studies not only advance our understanding of gene expression, but they also reveal gene transcripts that can be quantitatively assessed as molecular markers (i.e., biomarkers). Therefore, understanding the transcriptome is essential for interpreting the functional elements of the genome and revealing the molecular constituents of cells and tissues, and also for understanding development, such as muscle growth (Gupta, 2014).

An overview of technology platform for RNA sequencing (RNA-Seq)

By using the RNA sequencing (RNA-Seq) technique, mRNA can be quantitatively measured and hence used to evaluate the expression level of particular genes in a tissue, for example, muscle. In other words, by measuring the genes in the transcriptome, metabolic processes and hence, physiological pathways can be identified (Gupta, 2014; Mortazavi et al., 2008). RNA-Seq is a recently developed approach to transcriptome profiling that uses deep-sequencing technologies. It offers several key advantages: 1) no previous knowledge on gene sequence is necessary (Gupta, 2014), 2) can reveal the precise location of transcription boundaries, and the sequence variations in the transcript region (Wang et al., 2009), and 3) it has been shown to be highly accurate for quantifying gene expression levels (Nagalakshmi et al., 2008). Taking all of these advantages into account, RNA-Seq is the first sequencing-based method that allows the entire transcriptome to be surveyed in qualitative and quantitative manner, either by single-base resolution for annotation or digital gene expression levels at the genome scale and has the potential to rapidly and cost effectively expand sequence databases for non-model organisms, such as

hybrid striped bass (Estévez et al., 2012; Wang et al., 2009). With the power of RNA-Seq, especially when it is coupled to genetic analysis, it is expected that transcriptome analysis will help in genome annotation, expression analysis, functional analysis, and identification of candidate genes for important aquaculture traits. Thus, the implementation of transcriptomics through the RNA-Seq technique has been widely adopted throughout the life sciences (Mazurais et al., 2011; McLean, 2013) and aquaculture industry (Vélez et al., 2016).

Specific overview on myogenesis of fish growth

Myogenesis is a highly plastic process that may be influenced by internal and external signals arising from changing environmental conditions; swimming activity and nutritional inputs that are integrated to modify growth patterns, for example (i.e., different body sizes of fish from a rearing cohort) (Barbrier et al., 2016; Johnston, 2006). In all life history stages of fishes, myogenesis involves satellite cell proliferation, migration, fusion, terminal differentiation, and sarcomere assembly. Many transcription factors (e.g., *pax7* and *pax5*) and signaling molecules including myogenic protein factors (e.g., *myf5*, *myf6* and *myoD*), are required for the regulation of this complex processes in vertebrates (Figure 3.1) (Johnston et al., 2011; Harel, 2009).

An RNA-seq study was conducted by Palstra et al., (2013) to provide an in-depth view of the transcriptome of red and white skeletal muscle of exercised and non-exercised rainbow trout (*Oncorhynchus mykiss*) with the specific objective to quantify the transcriptomic effects of swimming-induced exercise. The results showed 1085 and 1228 novel gene sequences for red and white muscle, respectively, that included a number of

important molecules for skeletal muscle function. These included growth and myogenic factors and their receptors that participate in the regulation of myogenic proliferation and differentiation, such as myocyte enhancer factor 2C (*mef2c*), myogenic factor 6 also known as muscle-specific regulatory factor 4 (*myf6*), fibroblast growth factor 1, and others listed in Table 3.1. Of particular interest are components of the insulin-like growth factor (IGF) family, including the cation-independent mannose-6-phosphate receptor or IGF-2 receptor, which is known to transduce the myogenic differentiation-promoting effects of IGF-2 in muscle (Table 3.1) (Chen et al., 1995; Fauconneau, 2013). Another component of interest that has been identified was the cytoskeletal protein supervillin, which is also a gene involved in myogenic regulation (Ting et al., 2002). Estévez et al., (2012) produced a comprehensive transcriptome of fast-twitch skeletal muscle using RNA extracted from adult and juvenile Gilthead sea bream (*Sparus aurata*) subjected to different nutritional states and temperatures. The fast-twitch skeletal muscle transcriptome for the Gilthead sea bream contained 5655 unique annotated genes and 785 full-length coding sequences including key transcription factors, signaling molecules and structural proteins involved in myogenesis and growth (Table 3.2). All of these genes and additional structural components of the muscle, including desmin and myosin (Table 3.1), may thus be related to growth differences in hybrid striped bass (HSB). The goal of this study is to use RNA-Seq for the first time to identify the important genes related to growth performance in HSB

Study Design

Discrete metabolite biomarkers and the pathways underlying the regulation of enhanced growth rate and muscle accretion were performed for two different sized groups

of HSB (good- and poor-growth) (**Chapter III**). In this chapter, a novel machine learning analysis based on gene transcript data was used to investigate expressed gene biomarkers and the pathways accountable for the growth rate differences between the same two growth groups of HSB. This approach will provide a greater context complementary to the metabolomic markers identified as representative of poor- and good-growth performance in HSB. The gene list provided in this chapter will be used in **Chapter IV** for an interrelated pathway analysis between the expressed genes and the metabolites identified in the muscle samples collected from each HSB growth-group. Such a combinatorial approach will determine the importance of the transcriptome and metabolome in predicting growth performance of the fish. For example, are gene transcripts or metabolites better indicators or biomarkers of growth performance?

Materials & Methods

Experimental animals and tissue sampling

The muscle tissue samples from 4 fish representing the poor- and good-growth groups (N=4 fish each) were held at 4°C overnight and then excess RNALater was removed and the samples frozen at -80°C until extracted of RNA (see *Gene Expression Analysis* below).

Evaluations of growth parameters

As previously mentioned in **Chapter II**.

Muscle histology analysis

As previously mentioned in **Chapter II**.

Data Analysis

Growth parameters

As previously mentioned in **Chapter II**.

Muscle histology

As previously mentioned in **Chapter II**.

Gene expression analysis

Frozen muscle tissue from 4 fish representing the poor- and good-growth groups (N=4 fish each), were transported to the NCSU Genomic Sciences Laboratory (GSL) (Raleigh, NC, USA) for RNA extraction and library preparation for RNA-Seq using the Illumina platform. The libraries were sequenced using 75 bp single read chemistry on the Illumina NexSeq 500 (samples were pooled into one-half of a lane for approximately 23 to 30 million reads per sample). Total RNA was extracted using an RNeasy Fibrous Tissue mini total RNA isolation kit and manufacturer's protocol (Qiagen, USA).

Prior to library construction, RNA integrity, purity, and concentration were assessed using an Agilent 2100 Bioanalyzer with an RNA 6000 Nano Chip (Agilent Technologies, USA). Purification of messenger RNA (mRNA) was performed using the oligo-dT beads provided in the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs, USA). Complementary DNA (cDNA) libraries for Illumina sequencing were constructed using the NEBNext Ultra Directional RNA Library Prep Kit (NEB) and NEBNext Multiplex

Oligos for Illumina (NEB) using the manufacturer-specified protocol. Briefly, the mRNA was chemically fragmented and primed with random oligos for first strand cDNA synthesis. Second strand cDNA synthesis was then carried out with dUTPs to preserve strand orientation information. The double-stranded cDNA was then purified, end repaired and A-tailed for adaptor ligation. Following ligation, the samples were selected a final library size (adapters included) of 400-550 bp using sequential AMPure XP bead isolation (Beckman Coulter, USA). Library enrichment was performed and specific indexes for each sample were added during the protocol-specified PCR amplification. The amplified library fragments were purified and checked for quality and final concentration using an Agilent 2100 Bioanalyzer with a High Sensitivity DNA chip (Agilent Technologies, USA). The final quantified libraries were pooled in equimolar amounts for sequencing on the Illumina NextSeq 500 DNA sequencer, utilizing 75 bp single read sequencing with NextSeq Reagent Kit v2 (Illumina, USA). Flow cell cluster generation for the NextSeq 500 was performed on the instrument (Illumina, USA).

Quantitative RNS-Seq analysis

RNA-Seq analysis was performed by Data2Bio (Ames, Iowa, USA). Prior to alignment, the nucleotides of each single end raw read Illumina NexSeq read were scanned for low quality. Bases with PHRED quality value <20 out of 40 (Ewing and Green 1998; Ewing et al. 1998) were removed (i.e., only those with error rates of $\leq 1\%$ were included). Each read was examined in two phases. In the first phase, reads were scanned starting at each end and nucleotides with quality values lower than the PHRED threshold were removed. The remaining nucleotides were then scanned using overlapping windows of 10

bp and sequences with average PHRED quality value less than the specified threshold beyond the last window were truncated. The trimming parameters used are in reference to the software, Lucy (Chou and Holmes, 2001; Li and Chou, 2004). Trimmed reads were aligned to the reference striped bass genome from (<https://appliedecology.cals.ncsu.edu/striped-bass-genome-project/data-downloads/>) using GSNAP (Wu and Nacu, 2010). Confidently mapped, single end reads were filtered if they were mapped uniquely (≤ 2 mismatches every 36 bp and less than 5 bp for every 75 bp as tails) and used for subsequent analyses. The coordinates of uniquely aligned reads to the reference genome were used for positional reference and read count tallies were computed for each annotated gene. Singleton reads were assigned a count of one when their aligned coordinates overlapped with an annotated gene. Normalization was conducted by DESeq2 (<http://www.bioconductor.org/packages/release/bioc/html/DESeq2.html>), which corrects for biases introduced by differences in the total numbers of uniquely mapped reads in each sample and that also partially corrects for SNP variation between genotypes (biological variation among fish). Normalized read counts were used to calculate fold-changes and statistical significance. The R package DESeq2 was used to test the null hypothesis that expression of a given gene is not different between the two growth performance groups of poor and good (Love et al., 2014). A model using the negative binomial distribution of read counts was used to test this null hypothesis. The read counts per gene obtained from all the samples were used for the differential expression analysis. Before conducting the test for differentially expressed genes, a Principal Components Analysis (PCA) was conducted with the DESeq2 package using the top 500 genes with

highest variance among all samples to assay the quality of these samples. p -values of all DESeq2 statistical tests were converted to adjusted p -values (q-values) based on false discovery rate (FDR) (Benjamini and Hochberg, 1995). An FDR of 5% (q-value) was used to account for multiple testing in this project.

Machine learning evaluation of RNA-Seq expression

Performance of four machine learning models were used to evaluate muscle gene expression in hybrid striped bass related to the good- and poor-growth groups. All analyses were performed using WEKA version 3.8 (<http://www.cs.waikato.ac.nz/ml/weka/>, University of Waikato, Hillcrest, New Zealand) (Witten et al., 2016). Multilayer perceptron artificial neural networks (MLP), sequential minimal optimization support vector machines (SMO), J48 decision trees (J48), and Random Forest models were used to predict HSB growth status (classified as either good-growth or poor-growth) based on expression of muscle gene transcripts. Briefly, data from a subset of the fish were used to train the machine learning model and then the remaining data were used to cross-validate the learned pattern. During the cross-validation, the given algorithm predicts the growth performance of each individual fish based on gene expression patterns that it has learned during the training step. Two cross-validation strategies were used to evaluate the model learning: (1) a percentage split whereby 66% of the data were randomly selected and used to train the models and the remaining 33% of the data were input as a cross-validation and (2) a 8-fold stratified hold out with $n = 8$ folds where one fold was used for cross-validation and $n - 1$ folds of the randomly reordered data set were used for training (Douros et al., 2018; Reading et al., 2013, 2018; Schilling et al., 2014, 2015; Sullivan et al., 2015). Both

classes (good- and poor-growth groups) were properly represented in the model training and cross-validation data sets. The percentages of correct assignments during cross-validation were used to evaluate the model robustness (Chapman et al., 2014).

All expressed gene transcripts (72,893) were ranked by importance during classification using respective adjusted p -values (q -values). The performance of each model was evaluated with the inclusion of the highest ranked genes based on q -value (i.e., each of the models were run using the top 10, 25, 50, 75, 100, 150, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, and 10000 most important genes based on q -value). These models were then used to identify overfitting and underfitting of the trained models such that the data dimensionality could be reduced. Briefly, the model performance for each of these models was plotted to identify the maximum and minimum number of genes required for optimal model performance. The goal of this step is to identify those genes most important to predicting the difference between fish in the good and poor growth groups and thus those relevant to this trait (i.e., analogous to statistical significance). Additionally, the top 143 genes also were used in machine learning models as the cutoff for nominal significance was $q \leq 0.05$ and these 143 genes met that criterion (i.e., they were significantly different in expression between good- and poor-growth groups).

A machine learning sensitivity analysis based on WEKA support vector machine algorithms (SVMAttributeEval) also was conducted orthogonal to the ranking of genes based on p -value to reduce the dimensionality of the data (i.e., to reduce the dataset in size by omitting the values that were not assigned a given rank). Performance of machine learning models also were used to evaluate all informative expressed transcripts in HSB

related to the good- and poor-growth groups. The MLP and SMO models were used to predict HSB growth status (i.e., classified into good-growth and poor-growth) based on the 72,893 informative expressed muscle gene transcripts. First, a SMO model was conducted using all 72,893 informative expressed gene transcripts, which were then subsequently ranked by importance during the model-learning step using “SVMAttributeEval” on the full training data set (Guyon et al., 2002; Schilling et al. 2015). The ranked gene list was then used to evaluate different MLP and SMO model performance using the top highly ranked genes as similarly described above (i.e., the top 10, 25, 50, 75, 100, 150, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, and 70000 genes were all used in SMO models; the top 10, 25, 50, 75, 100, 150, 250, 300, 400, 500, 600, 700, 800, 900, 1000 genes were all used in MLP models due to limitation in computational power for the larger gene sets). These models were used to identify overfitting and underfitting of the trained models in order to reduce data dimensionality as described above. Of these genes the top 150 were arbitrarily chosen for pathway analysis.

An orthogonal approach was taken in order to also show the importance of key genes that were ranked by support vector machines, and that are important to the growth of HSB, whereby the top 10000 genes were used as baseline for MLP and SMO model performance and then the top 10, 25, 50, 75, 100, 150, 250, 500, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9500, 9600, 9700, 9800, 9900, 9990, 9995, 9996, 9997, 9998, and 9999 were subsequently eliminated from the input list. These reduced gene lists confirmed the importance of certain genes in relation to fish growth as identified above.

All data classes were properly represented in the machine learning training and cross-validation datasets. The performance of machine learning models using either SVMAttributeEval or *p*-value ranked lists were evaluated as a percent of correctly classified instances for each cross-validation. Additionally, the Kappa statistic and area under the receiver operating characteristic curve (AUROC) for each model were reported (Schilling et al., 2015).

The performance of all machine learning models was conducted in comparison to negative controls for machine learning. One negative control was based on the Law of Probability such that randomly classifying fish into two groups (i.e., good- and poor-growth) would be achieved with a success rate of 50% based on random chance alone. The performance of models using these randomized data sets should be near 50% correct classification as determined by the Law of Probability. The second negative control was conducted by randomizing the whole data set and entering it into each model as described above. A total of 8 such iterations of training and cross-validation using different, randomly ordered data sets was conducted for each negative control set and the average performance was reported (Schilling et al., 2014, 2015). Correctly classified instances were plotted against the number of top highly ranked gene attributes. Each plot was fitted with a polynomial trendline of an order two or three to identify the optimal model performance based on the number of input attributes. Within each plot, reduction of data dimensionality was conducted to identify and reduce model overfitting and underfitting. Model overfitting is defined as too many input attributes and model underfitting is defined as too few. Analyzing the best fit of a model allowed for the identification of the optimal number of genes and the most important ones in relation to the research question (i.e., which genes

are the most important for differentiating between poor-growth and good-growth group fish).

DAVID pathway analysis

The DAVID Functional Classification Tool (Schilling et al., 2014, 2015) was used to: (1) group the top 143 and 150 important genes of the two lists as determined by either *p*-value or SVMAttributeEval rank, respectively, based on their functional similarity and (2) to identify the corresponding Gene Ontology enrichment of each functional group (Schilling et al., 2014, 2015). All the HSB gene transcripts were identified, annotated, and the related gene sequences were taken from the striped bass genome available in JBrowse format (<http://stripedbass.animalgenome.org/jbrowse>) at the North Carolina State University College of Agriculture and Life Sciences database. The hybrid striped bass genes were manually collected from the National Center for Biotechnology Information (NCBI), GeneCards, and UniProt using approved gene names for zebrafish (*Danio rerio*) as described in Schilling et al. (2014, 2015).

Results

Experimental animals, Tissue sampling, and Evaluations of growth parameters

As previously mentioned in **Chapter II**

Muscle histology analysis

As previously mentioned in **Chapter II**

Gene expression analysis

A total of 223.5 M raw of RNA-Seq reads were subjected to quality checking and trimming. Less than 0.1% of raw reads were dropped and 99.1% of base pairs remained after quality trimming (Table 3.3). Quality trimmed reads were aligned to the striped bass genome and approximately 83.9% of the trimmed reads were mapped to the genome and an average of 83.8% was uniquely aligned. A total of 88.9% of the uniquely aligned reads were mapped to the predicted gene space.

Quantitative RNS-Seq analysis

The results of the PCA test that was conducted using the top 500 genes with highest variance among all samples showed that fish from the poor- and good-growth groups generally had similar gene expression profiles within the groups with one exception (Figure 3.2). One fish sample from the poor-growth group had a gene expression profile similar to the good-growth group. Overall, the PCA1 and PCA2 explain 91% of the variance in growth (Figure 3.2). The numbers of expressed genes, informative genes and the resulting differentially expressed genes (DEGs) identified in the comparison between the good- and poor-growth group fish are presented in Table 3.3. In total, 97,191 genes were identified and measured for expression. Of these, 72,893 genes were informative and had annotations. There were 143 genes that significantly differed in expression between fish from the poor- and good- growth groups (DESeq2, $q \leq 0.05$, FDR = 0.05); these were considered DEGs. Among these DEGs, 143 DEGs were up-regulated and 9 were down-regulated in the poor-growth group fish shown in Table 3.3.

Machine learning evaluation of RNA-Seq expression

The different machine learning models used to evaluate muscle gene expression in HSB related to the good- and poor-growth groups showed different performance in predicting HSB growth status when using all informative gene transcripts (72,893) ranked by importance during classification by their respective adjusted p -value (q -values). Using the MLP artificial neural networks model with a cross-validation of 66% split, the cross-validation had a percentage of 100% correctly classified instances (Figure 3.3A), Kappa statistic of 1.00, and AUROC of 1.00. The number of genes plotted against percentage of correctly classified instances had a polynomial trendline of the second order (Figure 3.3A). Due to limitation in computational capacity for the MLP model with a 66% split cross-validation, only one thousand genes were able to be used as predictors for growth status between the good- and poor growth groups of HSB muscle. The MLP model with a cross-validation of 8-fold had a percentage of correctly classified instances that ranged from 87.50% to 100% (Figure 3.3A), Kappa statistic ranged from 0.75 to 1.00, and AUROC of 1.00. The number of genes plotted against percentage of correctly classified instances had a polynomial trendline of the second order (Figure 3.3A). Similar to the previous cross-validation strategy, and as a result of the limitation in computational capacity in the MLP model with a cross-validation of 8-fold, only one thousand genes were considered to be the optimal number for predicting the growth status between the good- and poor-growth groups of HSB muscle. Using the J48 decision tree model with a cross-validation of 66% split, the model had a percentage of correctly classified instances that ranged from 66.67% to 100% (Figure 3.3B), Kappa statistic from 0.40 to 1.00, and AUROC from 0.72 to 1.00. The

number of genes plotted against percentage of correctly classified instances had a polynomial trendline of the second order (Figure 3.3B).

For consistency in comparison to the MLP analyses, only one thousand genes were used in the J48 analysis, out of the original 10000 genes that were considered to be the optimal number for predicting the growth status between the good- and poor-growth groups of HSB. Using the J48 model with cross-validation of 8-fold, the percentage of correctly classified instances ranged from 62.50% to 100% (Figure 3.3B), Kappa statistic ranged from 0.25 to 1.00, and AUROC from 0.63 to 1.00. The number of genes plotted against percentage of correctly classified instances had a polynomial trendline of the second order (Figure 3.3B). Similar to the previous cross-validation strategy, only one thousand genes were used in the analysis, out of the original 10000 genes that were predicated to be the optimal number for predicting the growth status between the good- and poor-growth groups of hybrid striped bass.

Using the SMO of support vector machines model with a cross-validation of 66% split, the model had a percentage of correctly classified instances that ranged from 33.33% to 100% (Figure 3.3C), Kappa statistic from 0.00 to 1.00, and AUROC from 0.50 to 1.00. The number of genes plotted against percentage of correctly classified instances had a polynomial trendline of the second order (Figure 3.3C). Similar to previous model consistency measures, only one thousand genes were used in the analysis, out of the original 10000 genes. Using the SMO model with cross-validation of 8-fold, the model had a percentage of correctly classified instances that ranged from 25.00% to 100% (Figure 3.3C), Kappa statistic from -0.05 to 1.00, and AUROC from 0.25 to 1.00. The number of

genes plotted against percentage of correctly classified instances had a polynomial trendline of the second order (Figure 3.3C).

Using the Random Forest model with a cross-validation of 66% split, the cross-validation had a percentage of 100% correctly classified instances (Figure 3.3D), Kappa statistic of 1.00, and AUROC of 1.00. The number of genes plotted against percentage of correctly classified instances had a polynomial trendline of the second order (Figure 3.3D). Parallel to previous model consistency measures, only one thousand genes were used in the analysis, out of the original 10000 genes that were considered to be the optimal number for predicting the growth status between the good- and poor-growth groups of HSB. Using the Random Forest model with cross-validation of 8-fold, the model had a percentage of correctly classified instances that ranged from 87.00% to 100% (Figure 3.3D), Kappa statistic from 0.75 to 1.00, and AUROC of 1.00. The number of genes plotted against percentage of correctly classified instances had a polynomial trendline of the second order (Figure 3.3D). Like the previous cross-validation strategy, only one thousand genes were used in the analysis, out of the original 10000 genes that were considered to be the optimal number for predicting the growth status between the good- and poor-growth groups of HSB.

The top 143 most important genes that were used as the cutoff for nominal significance (i.e., $q \leq 0.05$, they were significantly different between good- and poor-growth groups) were further identified and the difference in the elevation level between the two growth groups was expressed as the average FPKM value (Fragments per Kilobase of Transcript per Million Mapped Reads) (Table 3.4). Of those top 143 important genes, 134 genes were highly expressed in the fish muscle from the poor-growth group (93.71%) with

an average FPKM of 21.54 and a range of 0.25 to 190.57; only nine genes in the fish muscle from the good-growth group (6.29%) (*myog*, *pttg1*, *cd300ld3*, *ranbp2*, *bicd2*) were highly expressed, with an average FPKM of 2.63 and a range of 1.39 to 4.26.

Similar to the above described performance, all informative gene transcripts (72,893), which were ranked by importance using the sensitivity analysis based on support vector machine “SVMAttributeEval” model, were used to evaluate the difference of expressed genes between the good- and poor-growth groups of HSB. First, using the SMO model with a cross-validation of 66% split, the model had a percentage of correctly classified instances that ranged from 66.67% to 100% (Figure 3.4A), Kappa statistic from 0.40 to 1.00, and AUROC from 0.75 to 1.00. The number of genes plotted against percentage of correctly classified instances had a polynomial trendline of the third order (Figure 3.4A). Nineteen thousand genes were considered the optimal number for predicting the growth status between the good- and poor-growth groups of HSB using this approach. Using the SMO model with cross-validation of 8-fold, the model had a percentage of correctly classified instances that ranged from 62.50% to 100% (Figure 3.4A), Kappa statistic from 0.25 to 1.00, and AUROC from 0.63 to 1.00. The number of genes plotted against percentage of correctly classified instances had a polynomial trendline of the second order (Figure 3.4A). Nineteen thousand genes were considered the optimal number for predicting the growth status between the good- and poor-growth groups of HSB using this approach.

Using the MLP model with a cross-validation of 66% split, the cross-validation had a percentage of 100% correctly classified instances (Figure 3.4B), Kappa statistic of 1.00, and AUROC of 1.00. The number of genes plotted against percentage of correctly classified instances was linear (Figure 3.4B). There was no difference in performance of these MLP

models based on the number of input genes. Due to limitation in computational capacity in the MLP model with a 66% split cross-validation, only one thousand genes were predicted to be the optimal number for predicting the growth status between the good- and poor growth groups of HSB muscle. Using the MLP model with a cross-validation of 8-fold, the model had a percentage of correctly classified instances of 100% (Figure 3.4B), Kappa statistic of 1.00, and AUROC of 1.00. The number of genes plotted against percentage of correctly classified instances was linear (Figure 3.4B). Similar to the previous cross-validation strategy, and as a result of the limitation in computational capacity in the MLP model with a cross-validation of 8-fold, only one thousand genes were predicted to be the optimal number for predicting the growth status between the good- and poor-growth groups of hybrid striped bass muscle.

To reduce the number of genes predicted to be the top most important genes in relation to muscle growth, an orthogonal approach was applied from the previous step, whereby the top 10000 genes, ranked by support vector machines “SVMAttributeEval”, were used as baseline for both model SMO and MLP, and the result had verified that the top 10000 genes were all important (Figure 3.5). Of those important 10000 genes, the top 150 were arbitrarily chosen for the pathway analysis, as this is a convenient cutoff for evaluating pathway enrichment. These 150 genes were further examined and characterized and the difference in the elevation level between the two growth groups was expressed as the average FPKM value (Fragments per Kilobase of Transcript per Million Mapped Reads) (Table 3.5). Of the 150 top important genes, thirty-six genes (24%) were highly expressed in fish muscle from the poor-growth group with an average FPKM of 4.00 that ranged from 0.11 to 42.89, whereas, 114 genes (76%) were highly expressed in fish

muscle from the good-growth group, with an average FPKM of 2.67 that ranged from 0.20 to 16.90.

Lastly, the two reduced gene lists of 143 and 150, that were ranked either by *p*-value or by Class SVMAttributeEval machine learning model, respectively, were compared and the shared genes identified from both ranking procedures are reported in Table 3.6. Eleven genes, including their spliced variants, were found to be shared between the two ranking approaches and they were all elevated in the fish muscle from the poor-growth group (*agtr2*, *cd22*, *cebpb*, *cxcr4*, *gadd45b*, *mmp17*, *msi1*, *rgl1*, *rgs13*, *rhou*, and *rundc3a*).

DAVID pathway analysis

The DAVID GO enrichment analysis could not detect any functional similarity, hence, no Gene Ontology enrichment of any functional group could be revealed among the gene list ranked by the statistical inference (i.e., top 143 transcripts ranked by *p*-value). On the other hand, the DAVID Functional Classification Tool on the gene list ranked by Class SVMAttributeEval machine learning model (top 150 transcripts), revealed three annotation clusters that represented distinct functional groups. Cluster I, composed of 16 genes with enrichment score of 0.8 and a *p*-value ranging from 0.017 to 0.75. Cluster II, composed of 6 genes with enrichment score of 0.46 and a *p*-value ranged from 0.07 to 0.099. Cluster III, composed of 10 genes with enrichment score of 0.00 and a *p*-value ranged from 0.99 to 1.00. The actual genes of each of the corresponding clusters as well as functional GO classes are shown in Table 3.7.

Discussion

Genes identified as important in this transcriptomic aspect of the study were used in an integrated analysis of gene expression and metabolite data to interrogate biological functions and processes that are differentially regulated in white muscle obtained from HSB that were considered as good-growth (top 10% of the stock which are above the market size) and poor-growth (bottom 10% of the stock which are below the market size). This analysis used the gene lists presented in Table 3.4 and Table 3.5 and is detailed in **Objective 3, Chapter IV.**

The PCA plot using the top 500 highest variance genes in the white muscle sample of the HSB showed a tendency to form clusters indicative of a diverse gene expression profile input between fish representing the different growth groups (Figure 3.2). Each cluster showed that fish from the poor- and good-growth groups generally had similar gene expression profiles within the groups with one exception, where one sample from the poor-growth group had a gene expression profile more similar to the good-growth group. The putative outlier was for a female that had the highest gonad weight and a corresponding GSI (14.97) amongst the poor-growth groups, which may indicate early onset of sexual maturity. Maturation stage varies within the same age class of fishes, whereby some individuals may reach maturation earlier than others (Trippel, 1995). Another sample, this time from the good-growth group, had a unique gene expression profile from both growth groups of fish. The sample was for the only male found among the good-growth group and had the lowest gonad weight and corresponding GSI (8.36). The reason for such an exception may be the result of gender differences, where males mature earlier than females (Saillant et al., 2001). Altogether, energy assimilation may be a key factor for these

differences, through which the energetic dedication may be switched from somatic to reproductive growth (Quince et al., 2008). Overall, the PCA was able to differentiate 91% of the variance in growth performance between the poor- and good-growth groups (Figure 3.2) and that onset of sexual maturity may play a role in HSB growth performance.

It was noted that several of the gene transcripts had multiple variants identified. Several of the variant gene transcripts had different ranking numbers, however they corresponded to the same gene and this indicates that some of these genes were identified more than once by RNA-Seq (Table 3.6). These variants could represent alternatively spliced gene transcript variants (Mercer et al., 1996) or gene orthologs in the HSB, as the fish were created by crossing striped bass with white bass (i.e., two transcript variants could represent the striped bass and white bass gene orthologs). Further research will be required to elucidate the importance of these multiple transcript copies, as the high incidence of multiple gene variants likely indicates importance of this feature of the data.

The eleven genes, including many of their spliced variants, that were shared between the two ranking approaches were all elevated in fish muscle from the poor-growth group. These genes were *agtr2*, *cd22*, *cebpb*, *cxc4*, *gadd45b*, *mmp17*, *msi1*, *rgl1*, *rgs13*, *rhou*, and *rundc3a* (Table 3.6). These genes comprise a number of imperative biological functions related to cellular development (i.e., growth, regulation of cell shape and immune response, and activation of cell proliferation and maintenance). Angiotensin II Receptor Type 2 (*agtr2*) which has shown in human to mediate programmed cell death (i.e., apoptotic function) may play an important role in developmental biology and pathophysiology (Lehtonen et al., 1999). Moreover, it has also been shown in humans to cooperate with *mtus1* to inhibit *erk2* activation and cell proliferation (Nouet et al., 2004).

Likewise, Ral Guanine Nucleotide Dissociation Stimulator Like 1 (*rgl1*) is known in mammals to play a role in cell proliferation (Sood et al., 2000), in the same way, Musashi RNA Binding Protein 1 (*msi1*) has shown in mammals to play a role in the proliferation and maintenance of stem cells (Good et al., 1998). Amongst the eleven genes are those with regulatory functions, like, Matrix Metalloproteinase 17 (*mmp17*), which has been recognized in mammals to be involved in the activation of membrane-bound precursors of growth factors or inflammatory mediators (Rozanov et al., 2004). As with *mmp17*, Regulator Of G Protein Signaling 13 (*rgs13*) is a gene with a regulatory function and a member of the RGS protein family. It is a regulator of G-protein signaling, which has been shown in humans to accelerate GTPase activity of G protein alpha-subunits, thereby driving G proteins into their inactive GDP-bound form, thus negatively regulating G protein signaling. The RGS proteins have been implicated in the fine tuning of a variety of cellular events in response to G protein-coupled receptor activation (Johnson and Druey, 2002). The biological function of this specific gene, however, is unknown. CCAAT/Enhancer Binding Protein Beta (*cebpb*) is an important transcription factor regulating the expression of genes involved in immune and inflammatory responses in zebrafish (*Danio rerio*) (Huang et al., 2016). It also plays a significant role in adipogenesis, as well as in the gluconeogenic pathway, liver regeneration, and hematopoiesis (Armesto-Jimenez et al., 2019; Tang et al., 2005). Growth Arrest And DNA Damage Inducible Beta (*gadd45b*), is another regulatory gene whose transcript levels are increased following stressful growth arrest conditions. The function of this gene in human is regulation of growth and apoptosis and it mediates activation of the stress-response. It is regulated by different mechanisms, but expression is often coordinated to inhibit cell growth (Takekawa et al., 1998). An

additional regulatory gene is Ras Homolog Family Member U (*rhou*) which is known to play a role in the regulation of cell morphology and cytoskeletal organization in zebrafish (*Danio rerio*) (Dickover et al., 2014). In catfish (*Ictalurus punctatus*), it was found to participate in immune responses after bacterial infection (Tan et al., 2017). The cluster of differentiation-22 (*cd22*) was found in tongue sole (*Cynoglossus semilaevis*) to have an immune response as well. The *cd22* gene is a member of the immunoglobulin superfamily that serves as an inhibitory receptor for B-cell receptor signaling during the B-cell responses to foreign antigens (Li et al., 2017). Like *cd22*, C-X-C Motif Chemokine Receptor 4 (*cxcr4*) was found in miiuy croaker (*Miichthys miiuy*) to act as the main receptor of CXC chemokines and it plays a vital role in the immune system (i.e., host defense) and inflammation (Xu et al., 2014). Little is known about the function of RUN domain containing protein 3A (*rundc3a*), although it is thought to interact with *rap2a*, a protein hypothesized to act as an effector gene in neuronal cells in humans (Janoueix-Lerosey et al., 1998).

Overall, eleven genes were found shared between the two ranking approaches, which confirmed their importance to growth. They were all elevated in fish muscle from the poor-growth group. These genes all could serve as biomarkers of growth status in HSB and perhaps other fishes as well. The findings represented compelling evidence of cell death, reduced cell proliferation, and an immune response to a possible stressor in fish muscle from the poor-growth group relative to the fish muscle from the good-growth fish. Such findings reveal a potential mechanism by which growth is affected in HSB, suggesting that those affected may be less healthy and suffer from apoptosis leading to muscle wasting or poor growth in the fish observed here.

Running the DAVID Functional Classification Tool on the top 150 important gene transcripts classified via SVMAttributeEval machine learning model revealed three functional groups (Table 3.7). The gene members in these groups share commonality in biological function related to growth performance, including regulatory functions such as cell signaling, and second messenger systems associated with the plasma membrane, cell proliferation, and cell differentiation. These major pathways are highlighted above as an overview of the functions of the 11 key biomarkers shared between the two ranked gene lists. Expectedly, genes involved with insulin-like growth factor were identified. Insulin-like growth factor is a hormone capable of stimulating the proliferation and differentiation of human muscle stem cells (satellite cells) (Adams, 2001). Therefore, insulin-like growth factor pathways also may be important for promoting the good-growth in the fish observed here. These findings will set the stage for future HSB research aimed to improve growth through breeding and control of certain signaling pathways. A further, more detailed analysis of these gene pathways and their shared functions with metabolite pathways described in **Chapter II** are presented in **Chapter IV**.

References

- Adams, G. R.** (2001). Insulin-like growth factor in muscle growth and its potential abuse by athletes. *Western Journal of Medicine*, 175(1), 7.
- Armesto-Jimenez, A.U., Arason, A.J., Stefasson, O.A., Gudmundsson, G., Gudjonsson, T. and Magnusson, M.K.** (2019). A MUC5B polymorphism associated with Idiopathic Pulmonary Fibrosis mediates overexpression through decreased CpG methylation and C/EBP β transcriptional activation. *bioRxiv*.
- Barbieri, E., Guescini, M., Calcabrini, C., Vallorani, L., Diaz, A.R., Fimognari, C., Canonico, B., Luchetti, F., Papa, S., Battistelli, M. and Falcieri, E.,** (2016). Creatine Prevents the Structural and Functional Damage to Mitochondria in Myogenic, Oxidatively Stressed C2C12 Cells and Restores Their Differentiation Capacity. *Oxidative Medicine and Cellular Longevity*, 2016.
- Benjamini, Y., & Hochberg, Y.** (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1), 289-300.
- Chapman, R. W., Reading, B. J., & Sullivan, C. V.** (2014). Ovary transcriptome profiling via artificial intelligence reveals a transcriptomic fingerprint predicting egg quality in striped bass, *Morone saxatilis*. *PLoS One*, 9(5).
- Chen, X., Raab, G., Deutsch, U., Zhang, J., Ezzell, R. M., & Klagsbrun, M.** (1995). Induction of heparin-binding egf-like growth factor expression during myogenesis activation of the gene by MyoD and localization of the transmembrane form of the protein on the myotube surface. *Journal of Biological Chemistry*, 270(31), 18285-18294.

- Chou, H. H., and Holmes, M. H.** (2001). DNA sequence quality trimming and vector removal. *Bioinformatics*, 17(12), 1093-1104.
- Dickover, M., Hegarty, J.M., Ly, K., Lopez, D., Yang, H., Zhang, R., Tedeschi, N., Hsiai, T.K. and Chi, N.C.** (2014). The atypical Rho GTPase, RhoU, regulates cell-adhesion molecules during cardiac morphogenesis. *Developmental biology*, 389(2), 182-191.
- Douros, J. D., Baltzegar, D. A., Reading, B. J., Seale, A. P., Lerner, D. T., Grau, E. G., & Borski, R. J.** (2018). Leptin stimulates cellular glycolysis through a STAT3 dependent mechanism in Tilapia. *Frontiers in endocrinology*, 9, 465.
- Estévez, A., Andree, K., & Johnston, I. A.** (2012). Fast skeletal muscle transcriptome of the Gilthead sea bream (*Sparus aurata*) determined by next generation sequencing. *BMC genomics*, 13(1), 1.
- Eun, J. W., Ryu, S. Y., Noh, J. H., Lee, M. J., Jang, J. J., Ryu, J. C., Jung, K.H., Kim, J.K., Bae, H.J., Xie, H. and Kim, S.Y.** (2008). Discriminating the molecular basis of hepatotoxicity using the large-scale characteristic molecular signatures of toxicants by expression profiling analysis. *Toxicology*, 249(2), 176-183.
- Ewing, B., & Green, P.** (2000). Analysis of expressed sequence tags indicates 35,000 human genes. *Nature genetics*, 25(2), 232-234.
- Ewing, R. M., Jenkins, G. I., & Langdale, J. A.** (1998). Transcripts of maize RbcS genes accumulate differentially in C 3 and C 4 tissues. *Plant molecular biology*, 36(4), 593-599.
- Fauconneau, B., Chmaitily, J., Andre, S., Cardinal, M., Cornet, J., Vallet, J.L., Dumont, J.P. and Laroche, M.** (1993). Characteristics of rainbow trout flesh. 1. Chemical

- composition and cellularity of muscle and adipose tissues. *Sciences des Aliments (France)*.
- Gibb, E. A., Brown, C. J., & Lam, W. L.** (2011). The functional role of long non-coding RNA in human carcinomas. *Molecular cancer*, 10(1), 1.
- Good, P., Yoda, A., Sakakibara, S.I., Yamamoto, A., Imai, T., Sawa, H., Ikeuchi, T., Tsuji, S., Satoh, H. and Okano, H.** (1998). The Human Musashi Homolog 1 (MSI1) Gene Encoding the Homologue of Musashi/Nrp-1, a Neural RNA-Binding Protein Putatively Expressed in CNS Stem Cells and Neural Progenitor Cells. *Genomics*, 52(3), 382-384.
- Gupta, R. C. (Ed.)**. (2014). *Biomarkers in toxicology*. Academic Press.
- Guyon, I., Weston, J., Barnhill, S., & Vapnik, V.** (2002). Gene selection for cancer classification using support vector machines. *Machine learning*, 46(1-3), 389-422.
- Harel, tamar, Elisha Nathan, Libbat Tirosh-Finkel, Hila Zigdon, Nuno Guimarães-Camboa, Sylvia M. Evans, and Eldad Tzahor** (2009) Distinct Origins and Genetic Programs of Head Muscle Satellite Cells. *Cell Stem Cell*. 16 (6): 822-832.
- Hettmer, Simone & Amy J Wagers** (2010). Muscling in: Uncovering the origins of rhabdomyosarcoma. *Nature Medicine*. 16, 171-173.
- Huang, G., Zhang, F., Ye, Q., and Wang, H.** (2016). The circadian clock regulates autophagy directly through the nuclear hormone receptor Nr1d1/Rev-erba and indirectly via Cebpb/(C/ebpβ) in zebrafish. *Autophagy*, 12(8), 1292-1309.
- Janoueix-Lerosey, I., Pasheva, E., de Tand, M. F., Tavitian, A., and de Gunzburg, J.** (1998). Identification of a specific effector of the small GTP-binding protein Rap2. *European journal of biochemistry*, 252(2), 290-298.

- Johnson, E. N., and Druey, K. M.** (2002). Functional characterization of the G protein regulator RGS13. *Journal of Biological Chemistry*, 277(19), 16768-16774.
- Johnston, I. A.** (2006). Environment and plasticity of myogenesis in teleost fish. *Journal of Experimental Biology*, 209(12), 2249-2264.
- Johnston, I. A., Bower, N. I., & Macqueen, D. J.** (2011). Growth and the regulation of myotomal muscle mass in teleost fish. *Journal of Experimental Biology*, 214(10), 1617-1628.
- Lehtonen, J.Y., Daviet, L., Nahmias, C., Horiuchi, M. and Dzau, V.J.** (1999). Analysis of functional domains of angiotensin II type 2 receptor involved in apoptosis. *Molecular Endocrinology*, 13(7), 1051-1060.
- Li, S., & Chou, H. H.** (2004). LUCY2: an interactive DNA sequence quality trimming and vector removal tool. *Bioinformatics*, 20(16), 2865-2866.
- Li, Y. Q., Zhang, J., Li, J., and Sun, L.** (2017). First characterization of fish CD22: An inhibitory role in the activation of peripheral blood leukocytes. *Veterinary immunology and immunopathology*, 190, 39-44.
- Love, M. I., Huber, W., & Anders, S.** (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, 15(12), 550.
- Mazurais, D., Darias, M., Zambonino-Infante, J. L., & Cahu, C. L.** (2011). Transcriptomics for understanding marine fish larval development, This review is part of a virtual symposium on current topics in aquaculture of marine fish and shellfish. *Canadian Journal of Zoology*, 89(7), 599-611.

- McLean, T. I.** (2013). "Eco-omics": A Review of the Application of Genomics, Transcriptomics, and Proteomics for the Study of the Ecology of Harmful Algae. *Microbial ecology*, 65(4), 901-915.
- Mercer, J. G., Hoggard, N., Williams, L. M., Lawrence, C. B., Hannah, L. T., and Trayhurn, P.** (1996). Localization of leptin receptor mRNA and the long form splice variant (Ob-Rb) in mouse hypothalamus and adjacent brain regions by in situ hybridization. *FEBS letters*, 387(2-3), 113-116.
- Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B.** (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods*, 5(7), 621-628.
- Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., & Snyder, M.** (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science*, 320(5881), 1344-1349.
- Nouet, S., Amzallag, N., Li, J.M., Louis, S., Seitz, I., Cui, T.X., Alleaume, A.M., Di Benedetto, M., Boden, C., Masson, M. and Strosberg, A.D.** (2004). Trans-inactivation of receptor tyrosine kinases by novel angiotensin II AT2 receptor-interacting protein, ATIP. *Journal of Biological Chemistry*, 279(28), 28989-28997.
- Palstra, A.P., Beltran, S., Burgerhout, E., Brittijin, S.A., Magnoni, L.J., Henkel, C.V., Jansen, H.J., van den Thillart, G.E., Spaink, H.P. and Planas, J.V.** (2013). Deep RNA sequencing of the skeletal muscle transcriptome in swimming fish. *PloS one*, 8(1), e53171.
- Quince, C., Abrams, P. A., Shuter, B. J., and Lester, N. P.** (2008). Biphase growth in fish I: theoretical foundations. *Journal of Theoretical Biology*, 254(2), 197-206.

- Reading, B.J., McGinty, A.S., Clark, R.W., Hopper, M.S., Woods III, L.C., and Baltzegar, D.A.** (2018). Genomic Enablement of Temperate Bass Aquaculture (Family Moronidae), *In* Breeding and Culture of Perch and Bass (Wang, H., and Liang, X., Eds.). Science China Press (Chinese Academy of Sciences).
- Reading, B. J., Williams, V. N., Chapman, R. W., Williams, T. I., & Sullivan, C. V.** (2013). Dynamics of the striped bass (*Morone saxatilis*) ovary proteome reveal a complex network of the translasome. *Journal of proteome research*, 12(4), 1691-1699.
- Rozanov, D. V., Hahn-Dantona, E., Strickland, D. K., and Strongin, A. Y.** (2004). The low density lipoprotein receptor-related protein LRP is regulated by membrane type-1 matrix metalloproteinase (MT1-MMP) proteolysis in malignant cells. *Journal of Biological Chemistry*, 279(6), 4260-4268.
- Saillant, E., Fostier, A., Menu, B., Haffray, P., and Chatain, B.** (2001). Sexual growth dimorphism in sea bass *Dicentrarchus labrax*. *Aquaculture*, 202(3-4), 371-387.
- Sandvik, A. K., Alsberg, B. K., Nørsett, K. G., Yadetie, F., Waldum, H. L., & Laegreid, A.** (2006). Gene expression analysis and clinical diagnosis. *Clinica chimica acta*, 363(1), 157-164.
- Schilling, J., Loziuk, P. L., Muddiman, D. C., Daniels, H. V., & Reading, B. J.** (2015). Mechanisms of egg yolk formation and implications on early life history of white perch (*Morone americana*). *PloS one*, 10(11).
- Schilling, J., Nepomuceno, A., Schaff, J. E., Muddiman, D. C., Daniels, H. V., & Reading, B. J.** (2014). Compartment proteomics analysis of white perch (*Morone americana*) ovary using support vector machines. *Journal of proteome research*, 13(3), 1515-1526.

- Sood, R., Makalowska, I., Carpten, J.D., Robbins, C.M., Stephan, D.A., Connors, T.D., Morgenbesser, S.D., Su, K., Pinkett, H.W., Graham, C.L. and Quesenberry, M.I.** (2000). The human RGL (RalGDS-like) gene: cloning, expression analysis and genomic organization. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, 1491(1-3), 285-288.
- Sullivan C.V., Chapman R.W., Reading B.J., & Anderson P.E.** (2015) Transcriptomics of egg quality in teleost fish: New developments and future directions. *General and Comparative Endocrinology* 221, 23-30.
- Takekawa, M. and Saito, H.** (1998). A family of stress-inducible GADD45-like proteins mediate activation of the stress-responsive MTK1/MEKK4 MAPKKK. *Cell*, 95(4), 521-530.
- Tang, Q.Q., Grønborg, M., Huang, H., Kim, J.W., Otto, T.C., Pandey, A. and Lane, M.D.** (2005). Sequential phosphorylation of CCAAT enhancer-binding protein β by MAPK and glycogen synthase kinase 3 β is required for adipogenesis. *Proceedings of the National Academy of Sciences*, 102(28), 9766-9771.
- Tan, S., Yao, J., Zhou, T., Liu, S., Yuan, Z., Tian, C., Li, Q. and Liu, Z.** (2017). Identification, annotation and expression analysis of 29 Rho GTPase genes from channel catfish (*Ictalurus punctatus*) after bacterial infections. *Developmental & Comparative Immunology*, 67, 445-451.
- Ting, H. J., Yeh, S., Nishimura, K., & Chang, C.** (2002). Supervillin associates with androgen receptor and modulates its transcriptional activity. *Proceedings of the National Academy of Sciences*, 99(2), 661-666.

- Trippel, E. A.** (1995). Age at maturity as a stress indicator in fisheries. *Bioscience*, 45(11), 759-771.
- Vélez, E. J., Lutfi, E., Azizi, S., Perelló, M., Salmerón, C., Riera-Codina, M., Ibarz, A., Fernández-Borràs, J., Blasco, J., Capilla, E. and Navarro, I** (2016). Understanding fish muscle growth regulation to optimize aquaculture production. *Aquaculture*.
- Wang, Z., Gerstein, M., & Snyder, M.** (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews genetics*, 10(1), 57-63.
- Witten, I. H., Frank, E., & Mark, A. Hall, and Christopher J Pal.** (2016). Data Mining: Practical machine learning tools and techniques.
- Wu, T. D., & Nacu, S.** (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics*, 26(7), 873-881.
- Xu, T., Zhu, Z., Sun, Y., Ren, L., and Wang, R.** (2014). Characterization and expression of the CXCR1 and CXCR4 in miiuy croaker and evolutionary analysis shows the strong positive selection pressures imposed in mammal CXCR1. *Developmental & Comparative Immunology*, 44(1), 133-144.

Table 3.1. List of selected rainbow trout skeletal muscle genes grouped by functional categories. An RNA-seq of the transcriptome of skeletal muscle identified 1,085 gene sequences in red muscle (R) and 1,228 gene sequences in white muscle (W) (for more information about the other genes see Palstra et al., 2013).

No	Putative name and function	Muscle	Species
<i>Growth and Myogenic Factors</i>			
1	fibroblast growth factor 1 (acidic)	R	Zebra Fish (<i>Danio rerio</i>)
2	follistatin-like 1b	W	Zebra Fish (<i>Danio rerio</i>)
3	growth arrest-specific 7	R	African clawed frog (<i>Xenopus laevis</i>)
4	heparin-binding EGF-like growth factor	R, W	Zebra Fish (<i>Danio rerio</i>)
5	insulin-like growth factor binding protein 5a	R	Zebra Fish (<i>Danio rerio</i>)
6	insulin-like growth factor binding protein 6b	W	Zebra Fish (<i>Danio rerio</i>)
7	insulin-like growth factor-binding protein 3	W	Cattle (<i>Bos Taurus</i>)
8	myocyte enhancer factor 2ca	R, W	Zebra Fish (<i>Danio rerio</i>)
9	myocyte-specific enhancer factor 2C	R	Zebra Fish (<i>Danio rerio</i>)
10	myogenic factor 6	R, W	Zebra Fish (<i>Danio rerio</i>)
11	protein Wnt-2	R	Zebra Fish (<i>Danio rerio</i>)
12	Angiopoietin-1	R, W	Zebra Fish (<i>Danio rerio</i>)
<i>Receptors</i>			
13	androgen receptor	W	Zebra Fish (<i>Danio rerio</i>)
14	cation-independent mannose-6-phosphate receptor	W	Zebra Fish (<i>Danio rerio</i>)
15	TGF-beta receptor type-2	R	Zebra Fish (<i>Danio rerio</i>)
16	vascular endothelial growth factor receptor 1	R	Zebra Fish (<i>Danio rerio</i>)
17	frizzled homolog 3-like	R	Zebra Fish (<i>Danio rerio</i>)
18	bone morphogenetic protein receptor, type 1a	R	Zebra Fish (<i>Danio rerio</i>)
19	leukemia inhibitory factor receptor alpha	W	Zebra Fish (<i>Danio rerio</i>)

Table 3.1. (continued)

No	Putative name and function	Muscle	Species
20	ryanodine receptor 1b (skeletal)	R, W	Zebra Fish (<i>Danio rerio</i>)
21	ryanodine receptor 3	R, W	Red jungle fowl (<i>Gallus gallus</i>)
22	acetylcholine receptor subunit alpha	R, W	Zebra Fish (<i>Danio rerio</i>)
23	acetylcholine receptor subunit beta	W	House mouse (<i>Mus musculus</i>)
24	acetylcholine receptor subunit delta	R, W	Zebra Fish (<i>Danio rerio</i>)
<i>Structural and Cytoskeletal Elements</i>			
25	myosin heavy chain, cardiac muscle isoform	R	Red jungle fowl (<i>Gallus gallus</i>)
26	myosin-binding protein C, slow-type	R	Zebra Fish (<i>Danio rerio</i>)
27	myosin-Va	R	Zebra Fish (<i>Danio rerio</i>)
28	myosin-VI	R	Zebra Fish (<i>Danio rerio</i>)
29	Dystrophin	R	Zebra Fish (<i>Danio rerio</i>)
30	actin-binding protein IPP	W	Zebra Fish (<i>Danio rerio</i>)
31	actin-binding Rho-activating protein	W	Zebra Fish (<i>Danio rerio</i>)
32	actin-related protein 3B	W	Zebra Fish (<i>Danio rerio</i>)
33	Alpha-actinin-2	R	Zebra Fish (<i>Danio rerio</i>)
34	ankyrin 1, erythrocytic	R,W	Zebra Fish (<i>Danio rerio</i>)
35	ankyrin 3, epithelial isoform 1	W	Brown Rat (<i>Rattus norvegicus</i>)
36	Huntingtin	R	Zebra Fish (<i>Danio rerio</i>)
37	Supervillin	R, W	Zebra Fish (<i>Danio rerio</i>)
38	synaptopodin-2	R	Brown Rat (<i>Rattus norvegicus</i>)
39	Synemin	R, W	Zebra Fish (<i>Danio rerio</i>)
40	Tensin 1	W	Brown Rat (<i>Rattus norvegicus</i>)
41	troponin I, skeletal, slow like	R	Zebra Fish (<i>Danio rerio</i>)
42	troponin T type 1 (skeletal, slow)	R	African clawed frog (<i>Xenopus laevis</i>)
43	troponin T2a, cardiac	R	Zebra Fish (<i>Danio rerio</i>)

Table 3.2. List of selected paralogous genes identified in the Gilthead sea bream skeletal muscle transcriptome that are mainly related to fast skeletal muscle activities (Estévez et al., 2012).

No	Paralogue Gene Name	Gene Function
1	Calpain small subunit 1	Calcium-regulated thiol-protease involved in cytoskeletal remodeling.
2	Dysferlin interacting protein 1	Sarcolemma repair mechanism of both skeletal muscle and cardiomyocytes.
3	Glioblastoma amplified sequence	Widely expressed. Most abundant in heart and skeletal muscle
4	Myomesin 185 kDa	Major component of the vertebrate myofibrillar M band
5	Serine threonine-protein phosphatase	Essential for cell division, and participates in muscle contractility and protein synthesis

Table 3.3. Summary of hybrid striped bass gene expression data. The number of expressed genes that had an average read count of ≥ 5 , informative genes that passed DESeq2 filtering, and differentially expressed genes (DEGs) are provided. There were 143 up-regulated and 9 down-regulated genes in fish from poor-growth group. These DEGs were annotated via the top 5 hits of BLASTN, against non-redundant nucleotide of NCBI, with a e-value cutoff of 1×10^{-10} .

Expressed Genes	Informative Genes	DEGs (Up/Down)
97,191	72,893	143 (134/9)

Table 3.4. List of the top 143 most important genes in the white skeletal muscle of hybrid striped bass (143 is the number of genes that significantly differed in expression between fish from the poor- and good- growth groups). The genes were ranked based on their significant *p*-value level ($q \leq 0.05$, FDR = 0.05). The difference in the elevation level between the two growth groups (i.e., poor- and good-growth) is expressed as the average FPKM value (Fragments per Kilobase of Transcript per Million Mapped Reads). The black box in the table represents the upregulated genes between the two growth groups. The variant designation indicates genes that were identified more than once, such that different *de novo* transcript assemblies corresponding to the same gene were identified (e.g., spliced variants).

No.	Gene Name	Gene Code	Average FPKM	
			Poor-Growth	Good-Growth
1	mRNA decay activator protein ZFP36 (Variant 1)	<i>zfp36</i>	14.100	5.750
2	mRNA decay activator protein ZFP36 (Variant 2)	<i>zfp36</i>	23.486	10.207
3	Regulator of G Protein Signaling 13	<i>rgs13</i>	3.866	0.579
4	CCAAT/Enhancer Binding Protein Beta (Variant 1)	<i>cebpb</i>	3.408	0.884
5	CCAAT/Enhancer Binding Protein Beta (Variant 2)	<i>cebpb</i>	3.408	0.884
6	CCAAT/Enhancer Binding Protein Beta (Variant 3)	<i>cebpb</i>	2.940	0.813
7	C-X-C Motif Chemokine Receptor 4 (Variant 1)	<i>cxcr4</i>	4.881	1.609
8	CCAAT/Enhancer Binding Protein Beta (Variant 4)	<i>cebpb</i>	5.282	1.396
9	C-X-C Motif Chemokine Receptor 4 (Variant 2)	<i>cxcr4</i>	2.834	0.949
10	Sphingosine-1-Phosphate Receptor 1 (Variant 1)	<i>s1pr1</i>	1.912	0.563
11	mRNA decay activator protein ZFP36 (Variant 3)	<i>zfp36</i>	17.528	8.409
12	Ral Guanine Nucleotide Dissociation Stimulator Like 1 (Variant 1)	<i>rgl1</i>	3.043	1.313
13	Angiotensin II Receptor Type 2 (Variant 1)	<i>agtr2</i>	1.260	0.614
14	CCAAT/Enhancer Binding Protein Beta (Variant 5)	<i>cebpb</i>	5.877	1.521
15	Suppressor Of Cytokine Signaling 3 (Variant 1)	<i>socs3</i>	2.244	0.932
16	Ral Guanine Nucleotide Dissociation Stimulator Like 1 (Variant 2)	<i>rgl1</i>	1.674	0.723
17	Insulin Like Growth Factor Binding Protein 1 (Variant 1)	<i>igfbp1</i>	2.012	0.385
18	mitochondrial uncoupling protein 2 (Variant 1)	<i>ucp2</i>	190.569	31.935

Table 3.4. (continued).

No.	Gene Name	Gene Code	Average FPKM	
			Poor-Growth	Good-Growth
19	Uncoupling Protein 3	<i>ucp3</i>	70.482	11.856
20	Myogenin	<i>myog</i>	0.858	4.258
21	mitochondrial uncoupling protein 2 (Variant 2)	<i>ucp2</i>	37.850	6.504
22	Insulin Like Growth Factor Binding Protein 1 (Variant 2)	<i>igfbp1</i>	6.012	1.378
23	Sphingosine-1-Phosphate Receptor 1 (Variant 2)	<i>s1pr1</i>	1.941	0.556
24	mitochondrial uncoupling protein 2 (Variant 3)	<i>ucp2</i>	45.623	8.143
25	mitochondrial uncoupling protein 2 (Variant 4)	<i>ucp2</i>	45.623	8.143
26	Rho Family GTPase 3 (Variant 1)	<i>rnd3</i>	3.942	1.469
27	Rho Family GTPase 3 (Variant 2)	<i>rnd3</i>	3.942	1.469
28	Insulin Like Growth Factor Binding Protein 1 (Variant 3)	<i>igfbp1</i>	4.628	1.117
29	Regulator Of G Protein Signaling 21	<i>rgs21</i>	1.628	0.446
30	Nudix Hydrolase 16	<i>nudt16</i>	0.963	0.248
31	Ral Guanine Nucleotide Dissociation Stimulator Like 1 (Variant 3)	<i>rgl1</i>	6.481	3.842
32	Ral Guanine Nucleotide Dissociation Stimulator Like 1 (Variant 4)	<i>rgl1</i>	3.196	1.365
33	CD22 Molecule (Variant 1)	<i>cd22</i>	0.321	0.129
34	Rho Family GTPase 3 (Variant 3)	<i>rnd3</i>	4.235	1.587
35	High Mobility Group Box 1 (Variant 1)	<i>hmgb1</i>	33.082	23.954
36	High Mobility Group Box 1 (Variant 2)	<i>hmgb1</i>	32.859	23.792
37	High Mobility Group Box 1 (Variant 3)	<i>hmgb1</i>	32.859	23.792
38	Apolipoprotein A1 (Variant 1)	<i>apoa1</i>	18.656	7.704
39	Apolipoprotein A1 (Variant 2)	<i>apoa1</i>	13.971	5.786
40	Glutamate-Ammonia Ligase (Variant 1)	<i>glul</i>	10.029	3.659
41	Glutamate-Ammonia Ligase (Variant 2)	<i>glul</i>	7.402	2.580
42	Serpin Family A Member 1 (Variant 1)	<i>serpina1</i>	1.793	0.667

Table 3.4. (continued).

No.	Gene Name	Gene Code	Average FPKM	
			Poor-Growth	Good-Growth
43	C-X-C Motif Chemokine Ligand 8	<i>cxcl8</i>	1.533	0.644
44	Angiotensinogen	<i>agt</i>	1.394	0.519
45	Thrombospondin 1 (Variant 1)	<i>thbs1</i>	32.892	14.735
46	Thrombospondin 1 (Variant 2)	<i>thbs1</i>	12.900	5.619
47	Glutamate-Ammonia Ligase (Variant 3)	<i>glul</i>	11.832	4.138
48	Glutamate-Ammonia Ligase (Variant 4)	<i>glul</i>	11.832	4.138
49	Serpin Family A Member 1 (Variant 2)	<i>serpina1</i>	2.457	0.926
50	Serpin Family A Member 1 (Variant 3)	<i>serpina1</i>	2.251	0.845
51	Arrestin Domain Containing 2 (Variant 1)	<i>arrdc2</i>	84.926	18.076
52	Arrestin Domain Containing 2 (Variant 2)	<i>arrdc2</i>	84.926	18.076
53	Glutamate-Ammonia Ligase (Variant 5)	<i>glul</i>	10.028	3.700
54	Musashi RNA Binding Protein 1	<i>msi1</i>	4.717	2.792
55	Diablo IAP-Binding Mitochondrial Protein (Variant 1)	<i>diablo</i>	1.653	0.172
56	RUN Domain Containing 3A	<i>rundc3a</i>	5.779	3.751
57	Arrestin Domain Containing 2 (Variant 3)	<i>arrdc2</i>	121.441	25.657
58	Arrestin Domain Containing 2 (Variant 4)	<i>arrdc2</i>	121.441	25.657
59	Growth Arrest And DNA Damage Inducible Beta (Variant 1)	<i>gadd45b</i>	1.084	0.449
60	Aldolase, Fructose-Bisphosphate B (Variant 1)	<i>aldob</i>	5.415	3.235
61	Delta Like Canonical Notch Ligand 4 (Variant 1)	<i>dll4</i>	16.781	3.509
62	Glutamate-Ammonia Ligase (Variant 6)	<i>glul</i>	10.638	3.510
63	PDZ And LIM Domain 5	<i>pdlim5</i>	0.669	0.181
64	Aldolase, Fructose-Bisphosphate B (Variant 2)	<i>aldob</i>	4.081	2.373
65	Aldolase, Fructose-Bisphosphate B (Variant 3)	<i>aldob</i>	4.062	2.350
66	Aldolase, Fructose-Bisphosphate B (Variant 4)	<i>aldob</i>	3.265	1.898

Table 3.4. (continued).

No.	Gene Name	Gene Code	Average FPKM	
			Poor-Growth	Good-Growth
67	CD22 Molecule (Variant 2)	<i>cd22</i>	1.138	0.511
68	Arrestin Domain Containing 2 (Variant 5)	<i>arrdc2</i>	100.600	21.330
69	Arrestin Domain Containing 2 (Variant 6)	<i>arrdc2</i>	100.600	21.330
70	Arrestin Domain Containing 2 (Variant 7)	<i>arrdc2</i>	100.232	21.274
71	Glutamate-Ammonia Ligase (Variant 7)	<i>glul</i>	7.229	2.470
72	Aldolase, Fructose-Bisphosphate B (Variant 5)	<i>aldob</i>	3.735	2.169
73	Aldolase, Fructose-Bisphosphate B (Variant 6)	<i>aldob</i>	3.537	2.056
74	Aldolase, Fructose-Bisphosphate B (Variant 7)	<i>aldob</i>	3.471	2.018
75	Pituitary Tumor-Transforming 1 (Variant 1)	<i>pttg1</i>	0.424	2.876
76	Pituitary Tumor-Transforming 1 (Variant 2)	<i>pttg1</i>	0.415	2.820
77	Pituitary Tumor-Transforming 1 (Variant 3)	<i>pttg1</i>	0.313	2.125
78	Pituitary Tumor-Transforming 1 (Variant 4)	<i>pttg1</i>	0.309	2.094
79	Angiotensin II Receptor Type 2 (Variant 2)	<i>agtr2</i>	1.294	0.690
80	Serpin Family A Member 1 (Variant 4)	<i>serpina1</i>	1.319	0.490
81	Serpin Family A Member 1 (Variant 5)	<i>serpina1</i>	1.301	0.484
82	Serpin Family A Member 1 (Variant 6)	<i>serpina1</i>	1.238	0.460
83	Serpin Family A Member 1 (Variant 7)	<i>serpina1</i>	1.235	0.459
84	Arrestin Domain Containing 2 (Variant 8)	<i>arrdc2</i>	97.687	20.735
85	PRELI Domain Containing 3B	<i>prelid3b</i>	8.606	4.103
86	Fibrinogen Beta Chain (Variant 1)	<i>fgb</i>	4.256	1.451
87	Aldolase, Fructose-Bisphosphate B (Variant 8)	<i>aldob</i>	4.493	2.645
88	Regulator Of G Protein Signaling 5	<i>rgs5</i>	0.254	0.073
89	Phosphoenolpyruvate Carboxykinase 1	<i>pck1</i>	0.470	0.094
90	Arrestin Domain Containing 2 (Variant 9)	<i>arrdc2</i>	50.844	11.077

Table 3.4. (continued).

No.	Gene Name	Gene Code	Average FPKM	
			Poor-Growth	Good-Growth
91	Arrestin Domain Containing 2 (Variant 10)	<i>arrdc2</i>	50.844	11.077
92	TSC22 Domain Family Member 3 (Variant 1)	<i>tsc22d3</i>	32.676	14.439
93	Apolipoprotein A1 (Variant 3)	<i>apoa1</i>	7.028	3.097
94	Early Growth Response 1	<i>egr1</i>	1.751	0.780
95	Matrix Metalloproteinase 17	<i>mmp17</i>	2.056	0.885
96	Delta Like Canonical Notch Ligand 4 (Variant 2)	<i>dll4</i>	89.949	15.985
97	Growth Arrest And DNA Damage Inducible Alpha (Variant 1)	<i>gadd45a</i>	1.547	0.608
98	Growth Arrest And DNA Damage Inducible Alpha (Variant 2)	<i>gadd45a</i>	1.541	0.606
99	Ral Guanine Nucleotide Dissociation Stimulator Like 1 (Variant 5)	<i>rgl1</i>	7.190	4.522
100	TSC22 Domain Family Member 3 (Variant 2)	<i>tsc22d3</i>	38.958	17.163
101	CCAAT/Enhancer Binding Protein Beta (Variant 6)	<i>cebpb</i>	2.718	1.003
102	Fatty Acid Binding Protein 1	<i>fabp1</i>	4.577	0.854
103	Pituitary Tumor-Transforming 1 (Variant 5)	<i>pttg1</i>	0.244	1.393
104	Arrestin Domain Containing 2 (Variant 11)	<i>arrdc2</i>	40.232	9.866
105	Growth Arrest And DNA Damage Inducible Gamma (Variant 1)	<i>gadd45g</i>	66.373	27.044
106	Arrestin Domain Containing 2 (Variant 12)	<i>arrdc2</i>	31.104	7.642
107	Growth Arrest And DNA Damage Inducible Beta (Variant 2)	<i>gadd45b</i>	2.413	0.975
108	Growth Arrest And DNA Damage Inducible Beta (Variant 3)	<i>gadd45b</i>	2.302	0.930
109	Solute Carrier Family 3 Member 2 (Variant 1)	<i>slc3a2</i>	60.590	18.032
110	Solute Carrier Family 3 Member 2 (Variant 2)	<i>slc3a2</i>	48.054	14.302
111	Growth Arrest And DNA Damage Inducible Gamma (Variant 2)	<i>gadd45g</i>	53.649	21.904
112	Growth Arrest And DNA Damage Inducible Gamma (Variant 3)	<i>gadd45g</i>	53.649	21.904
113	Low Density Lipoprotein Receptor (Variant 1)	<i>ldlr</i>	2.192	0.933
114	Flavin Containing Monooxygenase 5	<i>fmo5</i>	0.485	0.082

Table 3.4. (continued).

No.	Gene Name	Gene Code	Average FPKM	
			Poor-Growth	Good-Growth
115	Arginine Vasopressin Induced 1	<i>avpi1</i>	5.048	2.330
116	Actin Binding Rho Activating Protein	<i>abra</i>	12.611	5.415
117	CMRF35-like molecule 3	<i>cd300ld3</i>	1.251	3.896
118	Fibrinogen Beta Chain (Variant 2)	<i>fgb</i>	3.402	1.235
119	Solute Carrier Family 3 Member 2 (Variant 3)	<i>slc3a2</i>	58.096	17.731
120	Progesterin And AdipoQ Receptor Family Member 6	<i>paqr6</i>	3.574	1.957
121	Growth Arrest And DNA Damage Inducible Gamma (Variant 4)	<i>gadd45g</i>	81.323	33.078
122	Complement C7 (Variant 1)	<i>c7</i>	2.348	0.575
123	Ras Homolog Family Member U (Variant 1)	<i>rhou</i>	2.949	1.104
124	Ras Homolog Family Member U (Variant 2)	<i>rhou</i>	2.949	1.104
125	Solute Carrier Family 3 Member 2 (Variant 4)	<i>slc3a2</i>	59.921	18.510
126	Solute Carrier Family 3 Member 2 (Variant 5)	<i>slc3a2</i>	51.338	15.973
127	Diablo IAP-Binding Mitochondrial Protein (Variant 2)	<i>diablo</i>	3.344	0.408
128	Complement C7 (Variant 2)	<i>c7</i>	4.633	1.125
129	RAN Binding Protein 2	<i>ranbp2</i>	0.685	1.603
130	Suppressor Of Cytokine Signaling 3 (Variant 2)	<i>socs3</i>	2.238	0.795
131	Suppressor Of Cytokine Signaling 3 (Variant 3)	<i>socs3</i>	2.206	0.784
132	Solute Carrier Family 3 Member 2 (Variant 6)	<i>slc3a2</i>	60.648	18.595
133	Suppressor Of Cytokine Signaling 3 (Variant 4)	<i>socs3</i>	1.438	0.605
134	Suppressor Of Cytokine Signaling 3 (Variant 5)	<i>socs3</i>	1.438	0.605
135	Solute Carrier Family 3 Member 2 (Variant 7)	<i>slc3a2</i>	62.066	18.524
136	Myosin Regulatory Light Chain Interacting Protein	<i>mylip</i>	1.148	0.567
137	Cysteine And Serine Rich Nuclear Protein 1	<i>csrnp1</i>	2.323	1.305

Table 3.4. (continued).

No.	Gene Name	Gene Code	Average FPKM	
			Poor-Growth	Poor-Growth
138	BICD Cargo Adaptor 2	<i>bicd2</i>	1.428	2.568
139	Insulin Like Growth Factor 1	<i>igf1</i>	35.674	10.798
140	Myelin Associated Glycoprotein	<i>mag</i>	0.865	0.470
141	Growth Arrest And DNA Damage Inducible Gamma (Variant 5)	<i>gadd45g</i>	90.043	37.788
142	Low Density Lipoprotein Receptor (Variant 2)	<i>ldlr</i>	0.358	0.107
143	Fibrinogen Alpha Chain	<i>fga</i>	3.035	1.167

Table 3.5. List of the top 150 most important genes in the white skeletal muscle of hybrid striped bass (150 is the number of genes that were arbitrarily chosen for the pathway analysis). The genes were ranked based on the Machine learning SVMAttributeEval model. The difference in the elevation level between the two growth groups (i.e., poor- and good-growth) is expressed as the Average FPKM value (Fragments per Kilobase of transcript per Million mapped reads). The black box in the table represents the upregulated genes between the two growth groups (i.e., poor- and good growth). The variant designation indicates genes that were identified more than once, such that different *de novo* transcript assemblies corresponding to the same gene were identified (e.g., spliced variants).

No	Gene Name	Gene Code	Average FPKM	
			Poor-Growth	Good-Growth
1	RUN Domain Containing 3A	<i>rundc3a</i>	5.779	3.751
2	Cyclic AMP-responsive element-binding protein 3-like protein 4 (Variant 1)	<i>creb3l4</i>	0.494	1.221
3	Anti-Silencing Function 1B Histone Chaperone (Variant 1)	<i>asf1b</i>	0.320	0.773
4	Matrix Metallopeptidase 17 (Variant 1)	<i>mmp17</i>	2.056	0.885
5	Fragile X Mental Retardation 1 (Variant 1)	<i>fmr1</i>	1.936	2.565
6	Anti-Silencing Function 1B Histone Chaperone (Variant 2)	<i>asf1b</i>	0.147	0.485
7	Proliferation-Associated 2G4 (Variant 1)	<i>pa2g4</i>	11.019	16.897
8	TRNA-YW Synthesizing Protein 3 Homolog (Variant 1)	<i>tyw3</i>	1.599	2.439
9	Cyclic AMP-responsive element-binding protein 3-like protein 4 (Variant 2)	<i>creb3l4</i>	0.322	0.807
10	Fragile X Mental Retardation 1 (Variant 2)	<i>fmr1</i>	1.067	1.420
11	ADAM Metallopeptidase Domain 10	<i>adam10</i>	0.697	0.790
12	PWWP Domain Containing 2B	<i>pwwp2b</i>	1.878	2.450
13	Coiled-Coil Domain Containing 88A	<i>ccdc88a</i>	0.111	0.073
14	Slit Guidance Ligand 2	<i>slit2</i>	0.209	0.402
15	Clustered Mitochondria Homolog (Variant 1)	<i>cluh</i>	3.435	5.123
16	STE20-Related Kinase Adaptor Alpha (Variant 1)	<i>strada</i>	0.776	1.108
17	Anti-Silencing Function 1B Histone Chaperone (Variant 3)	<i>asf1b</i>	0.147	0.485
18	TRNA-YW Synthesizing Protein 3 Homolog (Variant 2)	<i>tyw3</i>	0.513	0.783
19	Zinc Finger Protein 675	<i>znf675</i>	0.495	0.783

Table 3.5. (continued).

No	Gene Name	Gene Code	Average FPKM	
			Poor-Growth	Poor-Growth
20	STE20-Related Kinase Adaptor Alpha (Variant 2)	<i>strada</i>	0.776	1.108
21	Angiotensin II Receptor Type 2 (Variant 1)	<i>agtr2</i>	1.260	0.614
22	Clustered Mitochondria Homolog (Variant 2)	<i>cluh</i>	3.435	5.123
23	SNF-related serine/threonine-protein kinase-like (Variant 1)	<i>snrk</i>	0.853	0.455
24	NudE Neurodevelopment Protein 1 (Variant 1)	<i>nde1</i>	3.397	4.348
25	Regulator Of G Protein Signaling 13	<i>rgs13</i>	3.866	0.579
26	U2 Small Nuclear RNA Auxiliary Factor 2	<i>u2af2</i>	2.433	3.890
27	REM2 And RAB Like Small GTPase 1	<i>rsg1</i>	0.401	0.562
28	BCL2 Like 1	<i>bcl2l1</i>	0.610	0.929
29	Splicing Factor 1 (Variant 1)	<i>sf1</i>	2.907	4.681
30	Cysteine Rich Transmembrane BMP Regulator 1	<i>crim1</i>	2.284	1.566
31	Proliferation-Associated 2G4 (Variant 2)	<i>pa2g4</i>	9.802	15.013
32	NudE Neurodevelopment Protein 1 (Variant 2)	<i>nde1</i>	1.411	1.736
33	Dual Specificity Tyrosine Phosphorylation Regulated Kinase 2 (Variant 1)	<i>dyrk2</i>	7.261	11.763
34	Solute Carrier Family 25 Member 32	<i>slc25a32</i>	1.236	1.655
35	Dual Specificity Tyrosine Phosphorylation Regulated Kinase 2 (Variant 2)	<i>dyrk2</i>	7.261	11.763
36	Hepatitis A Virus Cellular Receptor 2	<i>havcr2</i>	0.211	1.404
37	Splicing Factor 1 (Variant 2)	<i>sf1</i>	2.936	4.728
38	Kelch Like Family Member 20	<i>klhl20</i>	0.219	0.362
39	Conserved Helix-Loop-Helix Ubiquitous Kinase (Variant 1)	<i>chuk</i>	0.388	0.693
40	Inter-Alpha-Trypsin Inhibitor Heavy Chain Family Member 6	<i>itih6</i>	42.895	28.646
41	PDS5 Cohesin Associated Factor A (Variant 1)	<i>pda5a</i>	1.547	2.540
42	Metastasis Associated 1 Family Member 2 (Variant 1)	<i>mta2</i>	1.115	1.737
43	Clathrin Interactor 1	<i>clint1</i>	1.754	2.354
44	SEC23 Interacting Protein	<i>sec23ip</i>	0.337	0.541
45	C-X-C Motif Chemokine Receptor 4 (Variant 1)	<i>cxcr4</i>	4.881	1.609

Table 3.5. (continued).

No	Gene Name	Gene Code	Average FPKM	
			Poor-Growth	Poor-Growth
46	Metastasis Associated 1 Family Member 2 (Variant 2)	<i>mta2</i>	1.082	1.679
47	Janus Kinase 2	<i>jak2</i>	0.609	0.497
48	Carbamoyl-Phosphate Synthetase 2, Aspartate Transcarbamylase, And Dihydroorotase (Variant 1)	<i>cad</i>	0.164	0.369
49	Solute Carrier Family 37 Member 2	<i>slc37a2</i>	0.332	0.218
50	PDS5 Cohesin Associated Factor A (Variant 2)	<i>pda5a</i>	0.919	1.508
51	Metastasis Associated 1 Family Member 2 (Variant 3)	<i>mta2</i>	1.222	1.941
52	Conserved Helix-Loop-Helix Ubiquitous Kinase (Variant 2)	<i>chuk</i>	0.207	0.356
53	Anti-Silencing Function 1B Histone Chaperone (Variant 4)	<i>asf1b</i>	0.526	1.191
54	Metastasis Associated 1 Family Member 2 (Variant 4)	<i>mta2</i>	1.069	1.697
55	WD Repeat Domain 46 (Variant 1)	<i>wdr46</i>	0.862	1.252
56	Microtubule Affinity Regulating Kinase 3	<i>mark3</i>	1.920	2.645
57	Collagen Type XX Alpha 1 Chain	<i>col20a1</i>	0.465	0.258
58	Myocyte Enhancer Factor 2D	<i>mef2d</i>	0.744	1.048
59	Serine/Arginine Repetitive Matrix 2	<i>srrm2</i>	1.681	2.498
60	CD22 Molecule	<i>cd22</i>	1.138	0.511
61	RAN Binding Protein 10	<i>ranbp10</i>	1.715	2.459
62	F-Box And Leucine Rich Repeat Protein 7	<i>fbxl7</i>	0.261	0.181
63	U-Box Domain Containing 5 (Variant 1)	<i>ubox5</i>	0.442	0.807
64	C-X-C Motif Chemokine Receptor 4 (Variant 2)	<i>cxcr4</i>	2.834	0.949
65	GDP Dissociation Inhibitor 1	<i>gdi1</i>	3.495	2.451
66	Carbamoyl-Phosphate Synthetase 2, Aspartate Transcarbamylase, And Dihydroorotase (Variant 2)	<i>cad</i>	0.099	0.201
67	Nuclear Apoptosis Inducing Factor 1	<i>naif1</i>	0.231	0.408
68	WD Repeat Domain 46 (Variant 2)	<i>wdr46</i>	0.876	1.272
69	CRK Like Proto-Oncogene, Adaptor Protein	<i>crkl</i>	0.251	0.399

Table 3.5. (continued).

No	Gene Name	Gene Code	Average FPKM	
			Poor-Growth	Poor-Growth
70	Growth Arrest And DNA Damage Inducible Beta (Variant 1)	<i>gadd45b</i>	2.413	0.975
71	DEAH-Box Helicase 8	<i>dhx8</i>	1.327	3.003
72	G Protein-Coupled Receptor 68 (Variant 1)	<i>gpr68</i>	1.476	2.092
73	ATPase Family, AAA Domain Containing 2B (Variant 1)	<i>atad2b</i>	0.245	0.362
74	Vacuolar Protein Sorting 13 Homolog B	<i>vps13b</i>	0.168	0.290
75	Methylenetetrahydrofolate Reductase (Variant 1)	<i>mthfr</i>	9.215	11.999
76	Carbamoyl-Phosphate Synthetase 2, Aspartate Transcarbamylase, And Dihydroorotase (Variant 3)	<i>cad</i>	0.109	0.226
77	Ras Homolog Family Member U	<i>rhou</i>	2.914	1.222
78	Cell Division Cycle 73 (Variant 1)	<i>cdc73</i>	1.690	2.454
79	U-Box Domain Containing 5 (Variant 2)	<i>ubox5</i>	0.447	0.816
80	Ral Guanine Nucleotide Dissociation Stimulator Like 1 (Variant 1)	<i>rgl1</i>	3.043	1.313
81	SMG1, Nonsense Mediated mRNA Decay Associated PI3K Related Kinase (Variant 1)	<i>smg1</i>	0.375	0.568
82	Zinc Finger Protein 644	<i>znf644</i>	0.248	0.409
83	U-Box Domain Containing 5 (Variant 3)	<i>ubox5</i>	0.393	0.701
84	Methylenetetrahydrofolate Reductase (Variant 2)	<i>mthfr</i>	5.192	6.804
85	Mitochondrial Ribosomal Protein L42	<i>mrpl42</i>	2.354	3.147
86	Musashi RNA Binding Protein 1 (Variant 1)	<i>msi1</i>	21.721	12.383
87	Zinc Finger Protein 318	<i>znf318</i>	0.564	0.721
88	Ral Guanine Nucleotide Dissociation Stimulator Like 1 (Variant 2)	<i>rgl1</i>	1.674	0.723
89	U-Box Domain Containing 5 (Variant 4)	<i>ubox5</i>	0.396	0.706
90	Cysteine Rich Angiogenic Inducer 61	<i>cyr61</i>	2.262	2.856
91	ECSIT Signalling Integrator (Variant 1)	<i>ecsit</i>	1.575	2.150
92	G Protein-Coupled Receptor 68 (Variant 2)	<i>gpr68</i>	1.040	1.474
93	Clustered Mitochondria Homolog (Variant 3)	<i>cluh</i>	2.591	3.090

Table 3.5. (continued).

No	Gene Name	Gene Code	Average FPKM	
			Poor-Growth	Poor-Growth
94	H2.0 Like Homeobox	<i>hlx</i>	0.363	0.931
95	Enhancer Of mRNA Decapping 3 (Variant 1)	<i>edc3</i>	0.149	0.334
96	Deoxyhypusine Hydroxylase (Variant 1)	<i>dohh</i>	2.891	3.750
97	Progesterin And AdipoQ Receptor Family Member 7	<i>paqr7</i>	0.321	0.163
98	MAP7 Domain Containing 1 (Variant 1)	<i>map7d1</i>	4.683	2.706
99	Deoxyhypusine Hydroxylase (Variant 2)	<i>dohh</i>	2.891	3.750
100	Chromosome 16 Open Reading Frame 72 (Variant 1)	<i>c16orf72</i>	4.749	6.108
101	MAP7 Domain Containing 1 (Variant 2)	<i>map7d1</i>	4.683	2.706
102	Musashi RNA Binding Protein 1 (Variant 2)	<i>msi1</i>	4.717	2.792
103	Transmembrane Protein 131	<i>tmem131</i>	4.146	5.501
104	Growth Arrest And DNA Damage Inducible Beta (Variant 2)	<i>gadd45b</i>	2.302	0.930
105	Hook Microtubule Tethering Protein 2 (Variant 1)	<i>hook2</i>	2.727	3.431
106	SNF-related serine/threonine-protein kinase-like (Variant 2)	<i>snrk</i>	0.477	0.268
107	Enhancer Of MRNA Decapping 3 (Variant 2)	<i>edc3</i>	0.149	0.334
108	Transmembrane 9 Superfamily Member 1	<i>tm9sf1</i>	0.287	0.497
109	CREB Regulated Transcription Coactivator 1	<i>crtc1</i>	0.947	1.626
110	ARP8 Actin Related Protein 8 Homolog	<i>actr8</i>	0.243	0.445
111	Serine/Threonine Kinase 17a	<i>stk17a</i>	0.568	0.415
112	Chromosome 12 Open Reading Frame 49	<i>cunh12orf49</i>	0.203	0.296
113	TRNA-YW Synthesizing Protein 3 Homolog (Variant 3)	<i>tyw3</i>	0.449	0.630
114	Metastasis Associated 1 Family Member 2 (Variant 5)	<i>mta2</i>	1.357	2.442
115	Serine And Arginine Repetitive Matrix 1	<i>srrm1</i>	2.998	4.651
116	Cell Division Cycle 73 (Variant 2)	<i>cdc73</i>	1.690	2.454
117	CCAAT/Enhancer Binding Protein Beta (Variant 1)	<i>cebpb</i>	3.408	0.884
118	Metastasis Associated 1 Family Member 2 (Variant 6)	<i>mta2</i>	1.387	2.496
119	SDS3 Homolog, SIN3A Corepressor Complex Component	<i>suds3</i>	1.831	2.946

Table 3.5. (continued).

No	Gene Name	Gene Code	Average FPKM	
			Poor-Growth	Poor-Growth
120	CCAAT/Enhancer Binding Protein Beta (Variant 2)	<i>cebpb</i>	3.408	0.884
121	ATPase Family, AAA Domain Containing 2B (Variant 2)	<i>atad2b</i>	0.170	0.237
122	ECSIT Signalling Integrator (Variant 2)	<i>ecsit</i>	2.321	3.148
123	WD Repeat Domain 46 (Variant 3)	<i>wdr46</i>	0.823	1.223
124	Chromosome 16 Open Reading Frame 72 (Variant 2)	<i>c16orf72</i>	4.749	6.108
125	WD Repeat Domain 46 (Variant 4)	<i>wdr46</i>	0.778	1.155
126	Phosphoenolpyruvate Carboxykinase 1	<i>pck1</i>	0.470	0.094
127	Spindlin 1 (Variant 1)	<i>spin1</i>	4.926	8.031
128	SMG1, Nonsense Mediated mRNA Decay Associated PI3K Related Kinase (Variant 2)	<i>smg1</i>	0.643	0.962
129	Ankyrin Repeat And Zinc Finger Domain Containing 1	<i>ankzf1</i>	0.964	1.447
130	Zinc Finger And BTB Domain Containing 33	<i>zbtb33</i>	0.234	0.340
131	Metastasis Associated 1 Family Member 2 (Variant 7)	<i>mta2</i>	0.818	1.472
132	Angiotensin II Receptor Type 2 (Variant 2)	<i>agtr2</i>	1.294	0.690
133	Hook Microtubule Tethering Protein 2 (Variant 2)	<i>hook2</i>	2.300	3.143
134	CDK5 Regulatory Subunit Associated Protein 2	<i>cdk5rap2</i>	4.763	2.954
135	Snf2 Related CREBBP Activator Protein	<i>srcap</i>	1.974	3.071
136	Matrix Metalloproteinase 17 (Variant 2)	<i>mmp17</i>	0.298	0.144
137	Catenin Delta 1	<i>ctnnd1</i>	0.774	0.871
138	Phosphatidylglycerophosphate Synthase 1	<i>pgs1</i>	2.144	3.108
139	Transmembrane Protein 183A	<i>tmem183a</i>	1.164	1.490
140	Glucosidase Alpha, Acid	<i>gaa</i>	7.034	5.655
141	Aggrecan	<i>acan</i>	0.565	0.320
142	Spindlin 1 (Variant 2)	<i>spin1</i>	4.926	8.031
143	NLR Family CARD Domain Containing 3	<i>nlrc3</i>	0.546	0.700
144	Exocyst Complex Component 2	<i>exoc2</i>	0.362	0.587

Table 3.5. (continued).

No	Gene Name	Gene Code	Average FPKM	
			Poor-Growth	Poor-Growth
145	Kelch Like Family Member 40	<i>klhl40</i>	6.481	8.428
146	Metastasis Associated 1 Family Member 2 (Variant 8)	<i>mta2</i>	1.631	2.562
147	Leucine Rich Repeats And Calponin Homology Domain Containing 1	<i>lrch1</i>	14.501	16.012
148	Beta-1,3-Glucuronyltransferase 1	<i>b3gat1</i>	0.522	0.743
149	Serine And Arginine Rich Splicing Factor 10	<i>srsf10</i>	3.309	5.113
150	Thyroid Hormone Receptor Interactor 11	<i>trip11</i>	1.164	2.162

Table 3.6. List of alphabetically arranged hybrid striped bass muscle genes that were shared between the two lists of the top 143 and 150 (143 is the number of genes that significantly differed in expression between fish from the poor- and good-growth groups by p -value ($q \leq 0.05$, FDR = 0.05) and 150 is the number of genes that were chosen for the pathway analysis based on machine learning rankings by SVMAttributeEval). The table shows the ranking of each gene in each classifying approach and the number of variants for each of the genes also is provided; more than one ranking number corresponds to the different variant designations. The arrows correspond to the gene elevation level which was based on the value of the average FPKM (Fragments per Kilobase of transcript per Million mapped reads) between the two growth groups.

No	Gene Code	Ranking		Variants		Elevation Level	
		SVMAttribute Eval	p -value	SVMAttribute Eval	p -value	SVMAttribute Eval	p -value
1	<i>agtr2</i>	21, 132	13, 79	2	2	↑ Poor-Growth	↑ Poor-Growth
2	<i>cd22</i>	60	33, 67	1	2	↑ Poor-Growth	↑ Poor-Growth
3	<i>cebpb</i>	117, 120	4, 5, 6, 8, 14, 101	2	6	↑ Poor-Growth	↑ Poor-Growth
4	<i>cxcr4</i>	45, 64	7, 9	2	2	↑ Poor-Growth	↑ Poor-Growth
5	<i>gadd45b</i>	70, 104	59, 107, 108	2	3	↑ Poor-Growth	↑ Poor-Growth
6	<i>mmp17</i>	4, 136	95	2	1	↑ Poor-Growth	↑ Poor-Growth
7	<i>msi1</i>	86, 102	54	2	1	↑ Poor-Growth	↑ Poor-Growth
8	<i>rgl1</i>	80, 88	12, 16, 31, 32, 99	2	5	↑ Poor-Growth	↑ Poor-Growth
9	<i>rgs13</i>	25	3	1	1	↑ Poor-Growth	↑ Poor-Growth
10	<i>rhou</i>	77	123, 124	1	2	↑ Poor-Growth	↑ Poor-Growth
11	<i>rundc3a</i>	1	56	1	1	↑ Poor-Growth	↑ Poor-Growth

Table 3.7. DAVID Functional Annotation Clustering report based on the top 150 most important expressed gene transcripts to hybrid striped bass growth as ranked using Support Vector Machine (SVMAttributeEval) model. Each Annotation cluster shares gene members with similar biological function. Gene Ontology (GO) class provides the common biological characteristics for all the Gene Members in the corresponding cluster. The Enrichment Score and the *p*-value of each cluster also are provided.

Annotation	GO Class	Gene Members	<i>p</i>-value	Enrichment Score
Cluster 1	ATP binding, Kinase activity, Phosphorylation, Transferase activity, Nucleotide binding	<i>msi1, pgs1, srcap,</i> <i>atad2b, cad, rsg13,</i> <i>tyw3, dhx8, actr8,</i> <i>nlrc3, pck1, strada,</i> <i>smg1, dyrk2, chuk,</i> <i>stk17a</i>	(0.015 – 0.57)	0.80
Cluster 2	Insulin-like growth factor binding protein, N-terminal extracellular region, Signal	<i>tmem131, cd22,</i> <i>nlrc3, cyr61, crim1,</i> <i>slits</i>	(0.07 – 0.099)	0.46
Cluster 3	Integral component of membrane, Membrane transmembrane helix, Transmembrane	<i>tmem183a,</i> <i>tm9sf1, itih6,</i> <i>bcl2l1, cd22,</i> <i>tmem131, slc37a2,</i> <i>crim1, havcr2,</i> <i>agtr2</i>	(0.99 – 1.00)	0.00

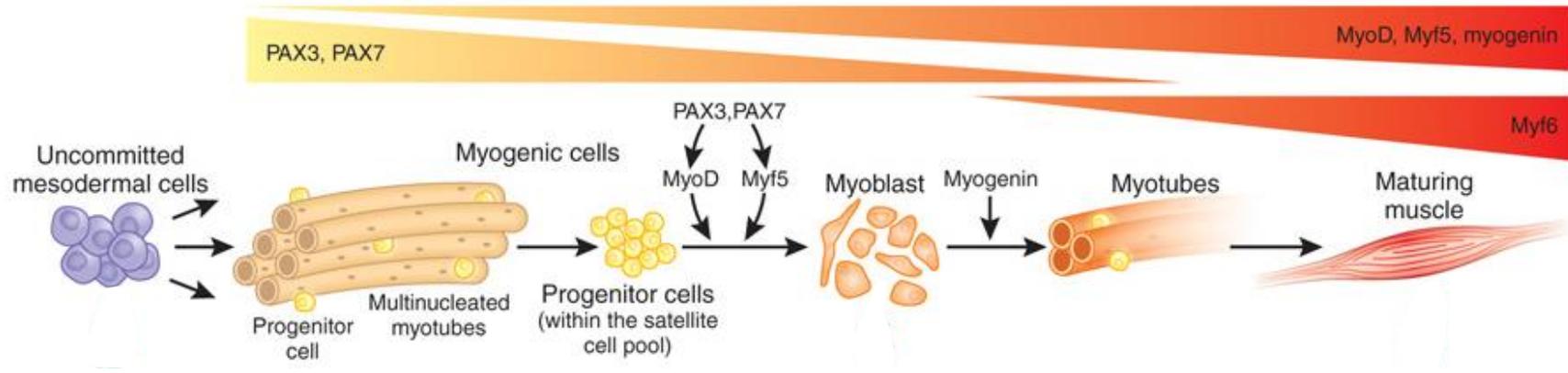


Figure 3.1. Muscle specification from the mesoderm depends on myogenic transcription factors (PAX3, PAX7, MyoD, Myf5, myogenin and Myf6). Postnatal muscle maintenance and regeneration invokes quiescent muscle precursor cells within the satellite cell pool that proliferate, terminally differentiate and fuse to generate multinucleated myotubes (From Hattmer et al., 2010).

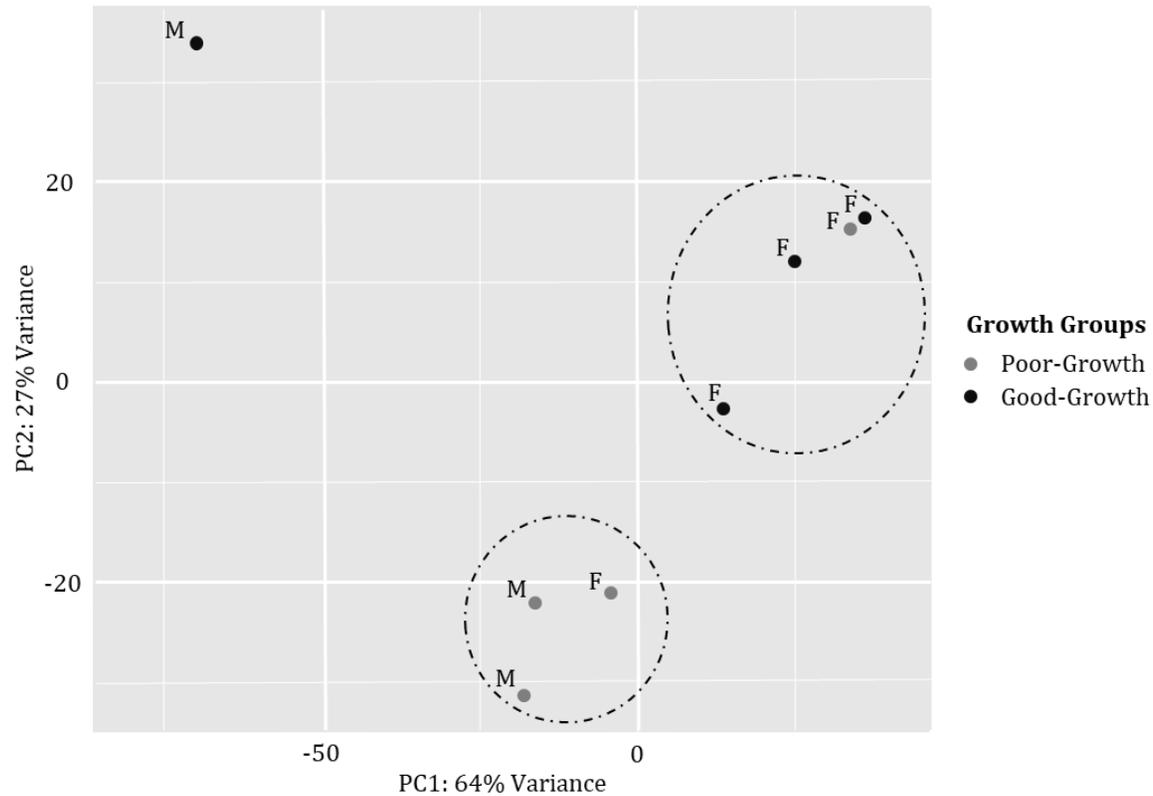
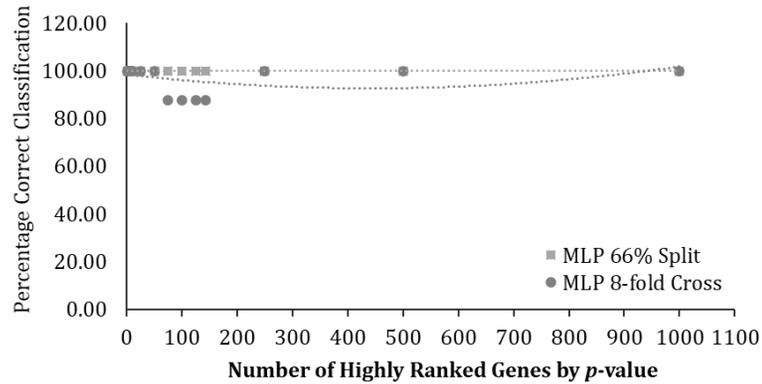


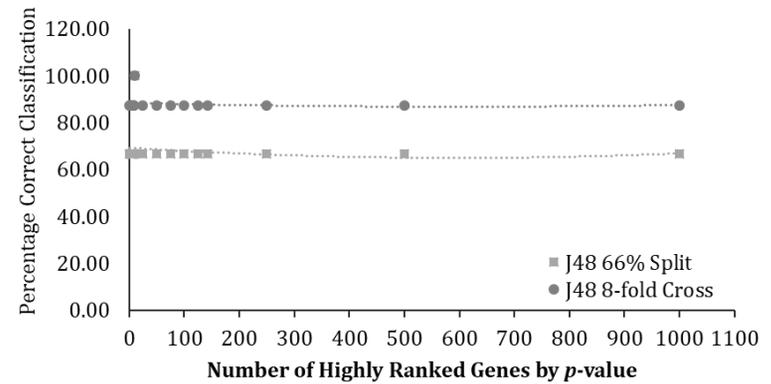
Figure 3.2. Principal Component Analysis using top 500 highest variance genes in the white muscle samples of hybrid striped bass. The data were clustered together based on growth performance of the white muscle in the hybrid striped bass, where the light grey color points show the poor-growth group, and the dark grey points show the good-growth group. PC1 and PC2, are the parameters that comprise these 500 genes and together they explained 91% of the variation in growth performance

Figure 3.3. Performance of machine learning models to evaluate all informative gene transcripts (72,893) in hybrid striped bass white muscle (i.e., related to the good- and poor-growth groups) that were ranked by importance during classification using the statistical inference tests ($p \leq 0.05$). Hybrid striped bass growth status was predicted (classified into good-growth and poor-growth groups) based on expression of muscle gene transcripts. (a) The Multilayer Perceptron Artificial Neural Networks (MLP) model, (b) The J48 Decision Tree (J48), (c) The Sequential Minimal Optimization Support Vector Machines (SMO) model, and (d) Random Forest model were used. Data points were fit with a polynomial line of the second order. The graphs show model robustness performance of the cross-validation as percentage of correctly classified instances when using the top highly ranked genes (10, 25, 50, 75, 100, 150, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, and 10000). Two cross-validation strategies were used to evaluate each model: (1) a percentage split whereby 66% of the data were randomly selected and used to train the models and the remaining 33% of the data were input as a cross-validation and (2) a 8-fold stratified hold out with $n = 8$ folds where one fold was used for cross-validation and $n - 1$ folds of the randomly reordered data set were used for training. Both classes (good- and poor-growth groups) were properly represented in the model training and cross-validation data sets.

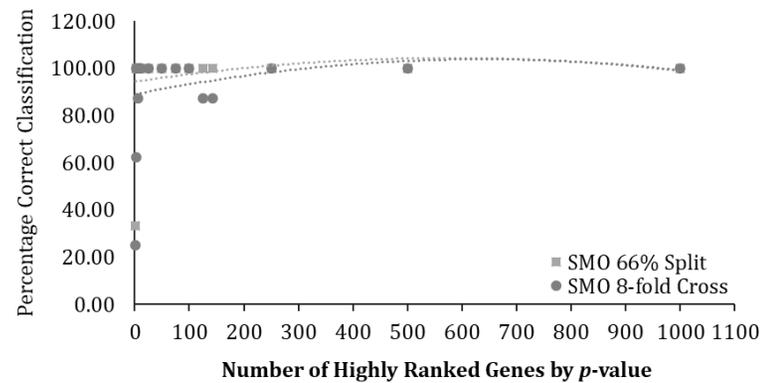
(a) Percentage of correctly classified instances using the Multilayer Perceptron Artificial Neural Networks (MLP) classifier



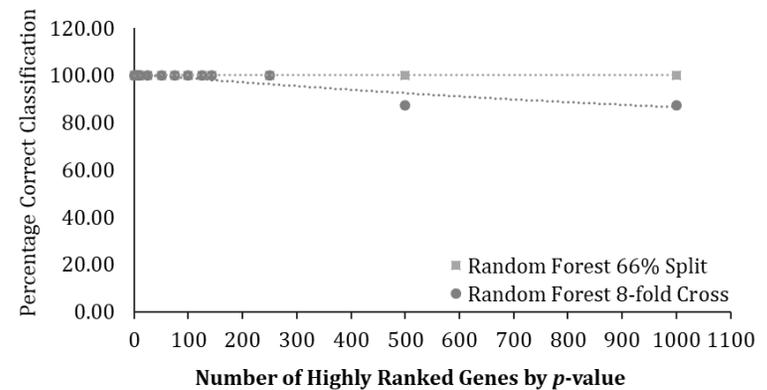
(b) Percentage of correctly classified instances using the J48 Decision Trees (J48) classifier



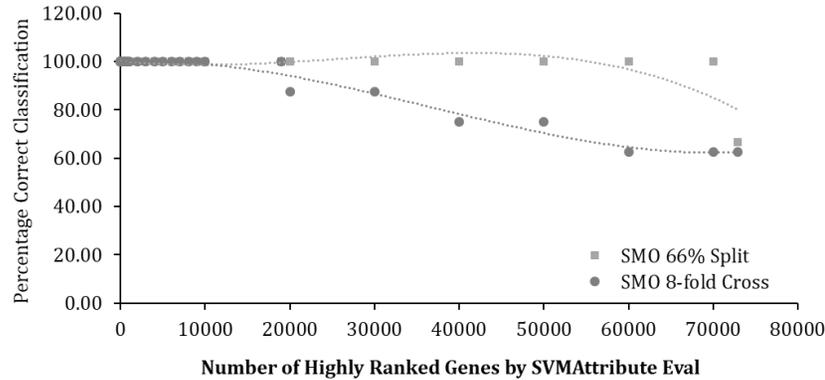
(c) Percentage of correctly classified instances using the Sequential Minimal Optimization Support Vector Machines (SMO) classifier



(d) Percentage of correctly classified instances using the Random Forest classifier



(a) Percentage of correctly classified instances using the Sequential Minimal Optimization Support Vector Machines (SMO) classifier



(b) Percentage of correctly classified instances using the Multilayer Perceptron Artificial Neural Networks (MLP) classifier

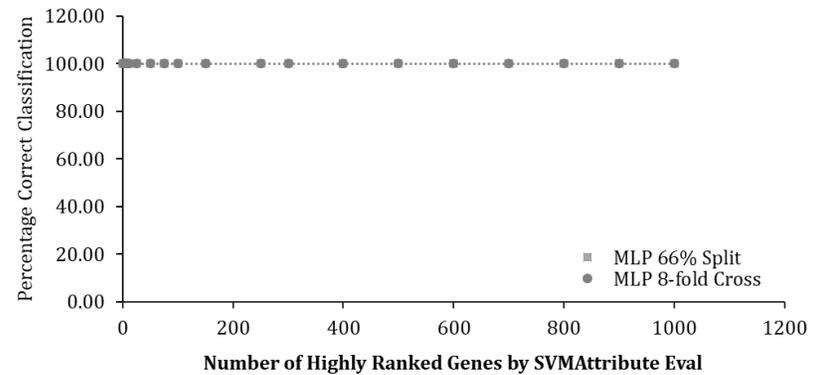


Figure 3.4. Performance of machine learning models to evaluate all informative gene transcripts (72,893) in hybrid striped bass white muscle (i.e., related to the good- and poor-growth groups) that were ranked by importance during classification using the Machine Learning SVMAttributeEval model. Hybrid striped bass muscle growth status was predicted (classified into good-growth and poor-growth groups) based on expression of muscle gene transcripts. (a) The Sequential Minimal Optimization Support Vector Machines (SMO) model and (b) The Multilayer Perceptron Artificial Neural Networks (MLP) model were used. Data points were fit with a polynomial line of the third order. The graphs show model robustness performance of the cross-validation as percentage of correctly classified instances when using the top highly ranked genes (10, 25, 50, 75, 100, 150, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, and 70000). The MLP model could not be performed with 1100 or more genes due to limitations in computational capacity. Two cross-validation strategies were used to evaluate the model learning: (1) a percentage split whereby 66% of the data were randomly selected and used to train the models and the remaining 33% of the data were input as a cross-validation and (2) a 8-fold stratified hold out with $n = 8$ folds where one fold was used for cross-validation and $n - 1$ folds of the randomly reordered data set were used for training. Both classes (good- and poor-growth groups) were properly represented in the model training and cross-validation data sets.

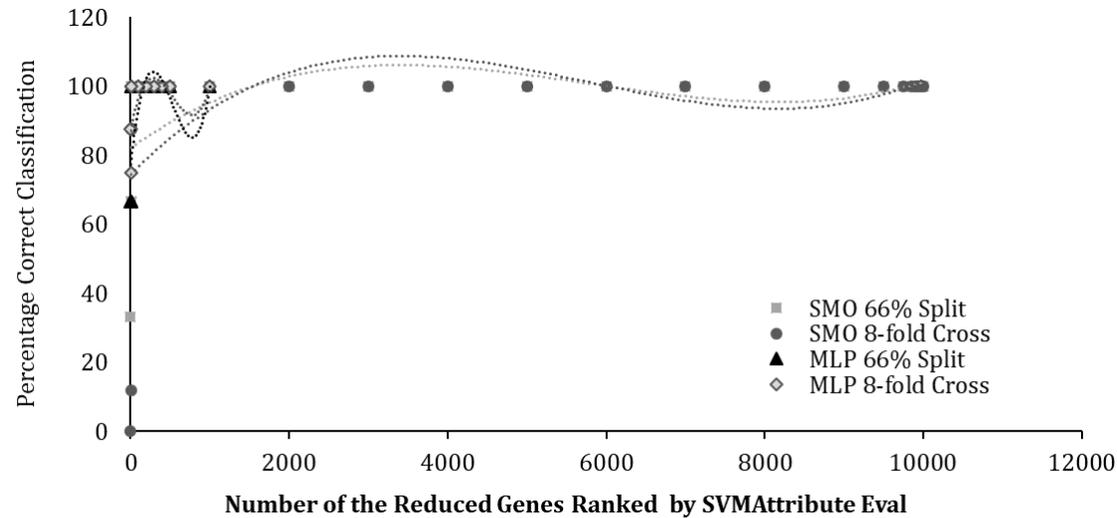


Figure 3.5. Performance of machine learning models to evaluate all informative gene transcripts (72,893) in hybrid striped bass white muscle (i.e., related to the good- and poor-growth groups) that were ranked by importance during classification using the Machine Learning SVMAttributeEval model. Hybrid striped bass growth status (classified into good-growth and poor-growth groups) was predicted based on expression of muscle gene transcripts using the Sequential Minimal Optimization Support Vector Machines (SMO) model and the Multilayer Perceptron Artificial Neural Networks (MLP) model. Data points were fit with a polynomial line of the third order. The graph shows the top 10,000 genes that were used in the SMO and MLP, if applicable, model and subsequently the top 10, 25, 50, 75, 100, 150, 250, and 500 genes that were eliminated from the list. The curves are the SMO and MLP model performance evaluation in what is called "reduction of data dimensionality". The MLP model could not be performed with 1100 or more genes due to limitations in computational capacity. Two cross-validation strategies were used to evaluate the model learning: (1) a percentage split whereby 66% of the data were randomly selected and used to train the models and the remaining 33% of the data were input as a cross-validation and (2) a 8-fold stratified hold out with $n = 8$ folds where one fold was used for cross-validation and $n - 1$ folds of the randomly reordered data set were used for training. Both classes (good- and poor-growth groups) were properly represented in the model training and cross-validation data sets.

Chapter IV: A Correlation Between Gene Expression And Metabolites

Abstract

A combination of gene expression and metabolite data was used to interrogate biological functions and processes that are differentially regulated in white muscle obtained from hybrid striped bass (striped bass x white bass) that were considered to have good-growth (top 10% of the stock, which are above the market size) and poor-growth (bottom 10% of the stock, which are below the market size). To ascertain the biological processes that may determine the size differences in hybrid striped bass (HSB), differential expression of genes and metabolites found in the muscles of these two fish groups was performed. Two gene lists of log₂ fold change were generated and annotated based on comparisons of good-growth vs. poor-growth groups: 1) statistical inference (DEseq2), 143 genes with a *p*-value of differential expression of <0.05 and 2) Support Vector Machine Learning (SVMAttributeEval)-ranked gene list of the top 150 genes (**Objective 2, Chapter III**). In addition to these gene lists, 30 out of 469 metabolites were identified from the global quantitative metabolomics panel analysis by importance of mean decrease accuracy during Random Forest. These 30 metabolites were identified as the top, most important predictors when comparing HSB muscle tissue from the good- and poor- growth groups (**Objective 1, Chapter II**). These gene and metabolite lists were used in Ingenuity Pathway Analysis (IPA). The shared and curated biological findings of the SVMAttributeEval machine learning ranked list collectively suggested a promotion of cell proliferation and perhaps cell differentiation generally in fish muscle from the good-growth group; the list ranked by statistical approaches similarly identified cell proliferation and cell differentiation. Inflammation, however, particularly of the liver, was inferred as an

important pathway that was predominantly activated based on the data from muscle of the fish from the good-growth group. The implications of these findings could provide a potential mechanism by which growth is affected in the fish muscle from the poor-growth group, which appear to be less healthy and suffer from increased apoptosis.

Introduction

Integrated pathways of metabolomics and transcriptomics data

The metabolome and the transcriptome are tightly coordinated and they are sometimes considered as compartmentalized with metabolism being represented as the actual biochemical processes that occur within the cell and the gene transcripts being the messages derived from the genome that direct these metabolic activities. For a cell to move, change shape, repair itself, grow, divide, and respond to external changes, metabolism must be coordinated by the genome and it does this through the mRNA and other non-coding RNAs (i.e., the transcriptome). Transcriptomic and metabolomic studies allow for the simultaneous analysis of the gene transcripts and metabolites and how they drive complex biological processes, such as muscle growth. Furthermore, a combinatorial analysis of the transcriptome and metabolome could be used to identify important cross-over between these important processes. For example, shared pathways that are shown to differ between larger and smaller hybrid striped bass (HSB) may be represented both by metabolites (end products of the cellular activities) and gene transcripts (instructional materials that direct the cellular activities). This is an important reminder that, if one desires a comprehensive understanding of gene function and/or expression, it is important to monitor the metabolic processes to which they frequently converge (Gut and Verdin, 2013).

This approach provides combined evidence of regulation and permits the identification of biological pathways that would otherwise not have been distinguished by independent single functional analyses (Carrillo et al., 2016). As an example, there are many levels of post-transcriptional regulation that may result in mRNA being degraded or stored as opposed to being translated and that these gene transcripts then would not elicit an overt effect on the metabolome. On the other hand, other gene transcripts may be translated into proteins, such as enzymes, that would actively alter the metabolome. Concurrent analysis of both datasets will allow for evaluation of these differences and similarities and thus the metabolomics data would reinforce conclusions drawn from the transcriptomic data in regard to accrual biological impact or function.

Specific overview on integrated metabolomics and transcriptomics

A deep insight into the potential integrated pathways between the identified metabolome and the transcriptome of HSB white muscle will offer a better understanding of the relationship between the transcriptome and the metabolome involved in white muscle growth pathways and gene regulation underlying growth performance differences. Gene pathways important to growth that were identified in the transcriptomics portion of the study (**Objective 2, Chapter III**) will likely be reflected by metabolites representing the downstream processes of those same pathways (**Objective 1, Chapter II**). Not all of these pathways are anticipated to overlap, but it is expected that several will. Very little information is available about parallel transcriptomic and metabolomic studies in other animals and no studies have been published about HSB growth in this manner. Furthermore, there is still a lack of investigation into the relationships between white

muscle primary metabolites, which are synthesized by the muscle cells, and are crucial for morphometric muscle growth (e.g., amino acids, fatty acids, carbohydrates, and nucleotides) (Harris, 2014; Holmes, 2008) and the transcriptome. A similar approach was conducted by Carrillo et al. (2016) using gene expression and metabolomics analysis of blood and muscle sampled from grass-fed and grain-fed steers (Wye Angus herd, *Longissimus dorsi*). The muscle results included 241 differentially expressed genes and the metabolome examined revealed 163 altered compounds. Accordingly, the alteration in glucose metabolism, divergence in free fatty acids and carnitine conjugated lipid levels, and altered beta-oxidation were observed. The anti-inflammatory n3-polyunsaturated fatty acids (PUFAs) were enriched in grass-fed beef, while higher levels of n6-PUFAs in grain-fed animals may promote inflammation and oxidative stress. Most importantly, blood cortisol levels strongly indicate that grass-fed animals may experience less stress than the grain-fed individuals (Zhou et al, 2012). These findings provided important insights into how rearing strategies for wye angus can be adjusted to support improved growth of angus, and are further supported by the findings that blood cortisol levels strongly indicate that grass-fed animals may experience less stress than the grain-fed individuals (Zhou et al. 2012). Although these differences were noted, direct pathway comparisons between the measured gene transcripts and metabolites were not investigated. In the proposed study, different metabolite levels may correspond to gene expression levels of related pathways in the white muscle of larger and smaller sized HSB groups. For example, up regulation of a gene transcript that encodes an enzyme may be correlated to change in a metabolite from a related pathway if the action of that enzyme results in synthesis or degradation/conversion of that particular metabolite. Additionally, and in agreement with Zhou et al, (2012),

elevated cortisol levels in the muscle tissue of beef may also be evident in the muscle of HSB from the smaller sized group in comparison to the larger group, which may indicate that the smaller fish were somehow stressed during rearing. Other metabolites related to the stress response in fishes will need to be evaluated in addition to muscle cortisol in order to conclude that the stress response is indeed related to reduced growth rates in HSB.

Differences in gene expression, which in turn will alter the protein and/or enzyme activity, muscle growth (i.e., hyperplasia versus hypertrophy), and muscle metabolites in HSB will be evaluated. Since gene expression and metabolite levels in each size group is expected to differ in response to the different body size (i.e., muscle growth), this will demonstrate the cause(s) underlying the growth rate differences in hybrid striped bass groups. Specifically applying these two tools with a novel machine learning-based analysis using Support Vector Machines (SVMAttributeEval) we will provide an informative list of the most important metabolite and gene indicators and their likely integrated pathways involved in HSB growth performance. Based on the previous citations, none have yet to analyze metabolomics and the transcriptomics data using such an approach.

After identifying the discrete molecular and genomic biomarkers and the pathways underlying the regulation of enhanced growth rate and muscle accretion in the two different sized groups (larger and smaller) of HSB from **Objectives 1 and 2, Chapter II and III**, a novel machine learning analysis based on both metabolomics and transcriptomics data was used to investigate the potential interrelated pathways between the expressed genes and the metabolites identified in the muscle samples that were collected from each HSB group. Furthermore, the investigation also will provide sensitivity

analysis which will be used to determine the importance of the transcriptome and metabolome in predicting growth performance of the fish. For example, are gene transcripts or metabolites better indicators or biomarkers of growth performance?

Material & Methods

To ascertain the biological processes that may determine the size differences between the two different HSB growth groups, differential expression of genes and metabolites found in the muscle tissues of these two groups was performed. Two gene lists of log₂ fold change with human homology were generated and annotated based on comparisons of good-growth versus poor-growth groups: 1) statistical inference (DEseq2), 143 genes with a *p*-value of differential expression of <0.05 cutoff and 2) Support Vector Machine SVMAttributeEval-ranked gene list of the top 150 genes from (**Objective 2, Chapter III**). In addition to these gene lists, 30 out of 469 metabolites were identified from the global quantitative metabolomics panel analysis by importance of mean decrease accuracy during Random Forest. These 30 metabolites were identified as the top, most important predictors when comparing HSB muscle tissue from the good- and poor- growth groups (**Objective 1, Chapter II**) and were used for Ingenuity Pathway Analysis (IPA) as well. Finally, BioProfiler was used to make novel discoveries by providing the ability to filter the fine-grained relationships between molecules (genes, RNAs, proteins, and chemicals) and diseases or functions.

Comparison Datasets

To interpret the differential expression of muscle genes and metabolites in the datasets described, the gene signatures were loaded into the Qiagen Ingenuity Pathway Analysis (IPA) software for further analysis. Before combining the gene and metabolite lists, each list was individually assessed in IPA to understand the biological meaning of the datasets alone. These lists were then combined to determine if additional information could be collected by joining the gene and metabolite lists for Core Analysis. The following combinations of muscle datasets were provided to IPA:

- 1) **GeneListA:** 143 identified genes (p -value of <0.05 ; \log_2 fold change good-growth vs. poor-growth) (**Objective 2, Chapter III**, Table 3.4).
- 2) **GeneListB:** 150 identified genes (SVMAttributeEval) ranked; \log_2 fold change good-growth vs. poor-growth) (**Objective 2, Chapter III**, Table 3.5).
- 3) **MetaboliteList:** 469 identified metabolites (\log_2 fold change good-growth vs. poor-growth; with p -values) (**Objective 1, Chapter II**, Table 2.6).
- 4) **GeneMetaboliteListA:** 143 genes from list 1 and 30 top metabolites from list 3 (a total of 173 IDs were identified by IPA database).
- 5) **GeneMetaboliteListB:** 150 genes from list 2 and 30 top metabolites from list 3 based on p -value (a total of 180 IDs were identified by IPA database).

Ingenuity Pathway Analysis (IPA) Upload

Genes were mapped to human ortholog gene IDs for use in IPA. Metabolites were mapped against the IPA KnowledgeBase (Qiagen, Germantown MD, USA) using three

identifiers (KEGG, HMDB, PUBCHEM ID). Comparison results were marked for expression log₂ fold change and *p*-value columns to be used in downstream analysis.

- 1) **GeneListA:** 142 genes were identified by IPA database from the original 143 genes (*p*-value of <0.05) that were previously identified in **Objective 2, Chapter III**.
- 2) **GeneListB:** 145 genes were identified by IPA database from the original 150 genes (SVMAttributeEval) that were previously identified in **Objective 2, Chapter III**.
- 3) **MetaboliteList:** 358 metabolites were identified by IPA database from the original 469 metabolites (by increasing importance and mean decrease accuracy during Random Forest analysis) that were previously identified in **Objective 1, Chapter II**.
- 4) **GeneMetaboliteListA :**142 IDs were identified from combining 173 genes and 358 metabolite IDs in IPA database.
- 5) **GeneMetaboliteListB:** 166 IDs were identified from combining 180 genes and 358 metabolite IDs in IPA database.

The above-mentioned datasets were used for running IPA core and comparison analysis, the overview of which along with the cut-offs used are defined in the sections below.

Ingenuity Pathway Analysis (IPA) Core Analysis

Ingenuity maintains databases on interactions between genes and other molecules, categorized by pathways and functions. In Qiagen IPA core analyses, the enrichment or depletion of a pathway or function is identified by the number of significantly differentially expressed genes within a pathway or function:

- 1) **Canonical Pathways Analysis:** Qiagen Ingenuity maintains and updates Canonical Pathways which consist of curated metabolic and signaling pathways. Canonical Pathway analysis predicts the changes in pathways based on the observed expression changes of genes within each pathway.
- 2) **Upstream Regulators Analysis:** Causal effects of upstream regulator proteins (e.g., miRNA, chemicals/drugs) and their downstream genes derived from scientific literature are compiled in the Qiagen Ingenuity Knowledge Base. Upstream regulator analysis uses the observed expression changes of the downstream genes in the dataset and predicts the changes of the upstream regulators.
- 3) **Diseases and Functions Analysis:** Causal effects between genes and functions (i.e., downstream effects, such as cellular processes and biological functions) derived from literature are compiled in the Qiagen Ingenuity KnowledgeBase. Diseases-and-functions analysis predicts the downstream functions that are expected to change given the observed gene expression changes in the dataset.

IPA Core Analysis Settings

The mapped datasets were loaded into IPA and each used independently to perform a Core Analysis in which an assessment of the signaling and metabolic pathways, Upstream Regulators, molecular interaction networks, and Disease and Biological Functions that are most likely to be perturbed based on the changes of expression of genes and metabolites in the datasets provided. To perform the Core Analysis, a further built-in p -value cutoff was applied by IPA database to each of the uploaded data sets. Accounting for the p -value cutoff, the following number of “analysis-ready” molecules were used to perform a Core Analysis:

- 1) **GeneListA:** Applying a p -value cutoff of $p < 0.05$; 57 molecules were identified in the IPA database of which 53 molecules were downregulated and 4 molecules were upregulated in the white muscle of the good-growth group relative to the poor-growth group.
- 2) **GeneListB:** Without applying a p -value cutoff; 96 molecules were identified in IPA database of which 30 molecules were downregulated and 66 molecules upregulated in the white muscle of the good-growth group relative to the poor-growth group.
- 3) **MetaboliteList:** Applying a p -value cutoff of $p < 0.05$; 39 molecules were identified in IPA database of which 19 molecules were downregulated and 20 molecules were upregulated in the white muscle of the good-growth group relative to the poor-growth group.
- 4) **GeneMetaboliteListA:** Without applying a p -value cutoff; 70 molecules were identified in IPA database of which 56 molecules were downregulated and 14 molecules upregulated in the white muscle of the good-growth group relative to the poor-growth group.
- 5) **GeneMetaboliteListB:** Without applying a p -value cutoff; 117 molecules were identified in IPA database of which 41 molecules were downregulated and 76 molecules upregulated in the white muscle of the good-growth group relative to the poor-growth group.

IPA Core Analysis Statistics

Core Analysis was performed using the settings above to identify Canonical Pathways, Upstream Regulators and Diseases and Functions that are significantly enriched

in these datasets. The analyses were performed with the combinations above to investigate the individual contribution of the gene and metabolite functions as well as together (genes and metabolites) to determine if additional interpretations can be garnered by analyzing the combined datasets. The Fisher's Exact Test was used to calculate a statistical significance of overlap (sharing) of the dataset molecules with various pathways. The test looks at the number of molecules that are in the reference and those:

- a. That match between a pathway and analysis-ready molecules (i.e., those that passed cutoffs and any filters applied).
- b. That are associated with the annotation, but did not match the analysis-ready molecules.
- c. That are analysis-ready, but did not match the annotation molecules.
- d. That were possible to assay in the experiment (were detectable on the platform used), but were not in the annotation or analysis ready; this is the reference set with the other sets subtracted.

In the right-tailed Fisher's Exact Test, only over-represented annotations (i.e., those that have more analysis-ready molecules than expected by chance) are significant. Under-represented annotations ("left-tailed" p -values), which have significantly fewer molecules than expected by chance, are not shown (i.e., not considered enriched or over-represented).

The p -values were calculated using the Reference Set that was defined in the settings within the Create Analysis page. Known identifiers were used in the input file and a cutoff applied to indicate the molecules of interest for Core Analysis. The unfiltered dataset was used as the reference set whereby the resulting p -values indicating "how significant"

the molecule overlap was with each annotation and considering both the molecules assayed (reference set) and the input molecules that met the cutoff (analysis-ready molecules).

For the RNA-Seq experiment, gene IDs were only generated for genes determined to be differentially expressed (GeneListA) or ranked by SVMAttributeEval (GeneListB), so the Reference Set was a universal set of genes in the Ingenuity KnowledgeBase. For the MetaboliteList of 469 compounds, a p -value cutoff of 0.05 was used for the input list of metabolites. For GeneMetaboliteListA and GeneMetaboliteListB, since the gene lists provided had p -values > 0.05 , to avoid filtering these genes out (these genes were ranked by SVMAttributeEval and p -value was not appropriate to use), thirty analysis-ready metabolites were combined with either the 143 genes ranked by p -value or 150 genes ranked by SVMAttributeEval and analyzed against the IPA KnowledgeBase reference set.

IPA Comparison Analysis

To determine the overlap among the five core analyses performed in IPA, a comparison analysis was performed, in which the activation/inhibition status of Canonical Pathways, Upstream Regulators and Diseases and Functions were compared. A comparison analysis was performed to determine the overlap of these functional enrichments identified in the ranked gene lists and the metabolite measurements derived from good-growth versus poor-growth fish. In this analysis, the data from individual core analyses was combined to generate heatmaps comparing predicted activation or inhibition of key functional processes.

IPA BioProfiler Analysis

To determine the connections between GeneMetaboliteListA and GeneMetaboliteListB to specific diseases and functions, BioProfiler analysis was run on the two data lists independently and filtered to phenotypes of interest, including growth of muscle and response of liver.

Results and Discussion

To identify the molecular mechanisms that govern the quality of growth between the good- and poor- growth white muscle of HSB, comparisons were generated at the gene and metabolite level. Core Analyses were performed on these datasets to assess the biological processes that were significantly different in the size-dependent grouping of the fish. The results of these Core Analyses represent the enrichment of curated Canonical Pathways, Upstream Regulators, and Functions that are not only enriched in the dataset, but are also predictive of phenotype (i.e., growth performance).

Canonical Pathways

The Canonical Pathway Analysis identified Qiagen IPA pathways that were significantly enriched in the dataset. The significance of the association between the dataset and a Canonical Pathway was determined from a *p*-value of overlap calculated using a right-tailed Fisher's Exact Test. In addition, pathways were predicted to be activated or inhibited when possible based on the entire dataset provided to IPA. The outputs from Canonical Pathways analysis are shown in Figure 4.1 and Figure 4.2.

Of the five datasets, GeneListA and GeneMetaboliteListA (those lists generated based upon statistics) had very similar canonical pathway results. Namely, the most significant pathways were acute phase response signaling and LXR/RXR activation. In both datasets, there was predicted inhibition of these two pathways in the good-growth group white muscle. The acute phase response signaling pathway provides a rapid inflammatory response that provides protection against microorganisms using non-specific defense mechanisms (Gabay and Kushner, 1999; Kyriakis and Avruch, 2001; Ahmad et al., 2011), while LXR/RXR is involved in the regulation of lipid metabolism, inflammation, and cholesterol to bile acid catabolism (Rasmussen and Wolfe, 1999; Choy and Panayi, 2001; Joseph et al., 2003). Furthermore, the senescence pathway was also predicted to be significantly inhibited in fish muscle from the good-growth group (Figure 4.1). Further details of the pathways are shown in Figure 4.3, where the downregulation of *cebpbeta*, *cxcl8* and *gadd45a*, *gadd45b* and *gadd45g* were observed in the fish from the good-growth group muscle. This is further predicted to lead to increased proliferation of cells (Formichi et al., 2006; Fridlyanskaya et al., 2015) and reduced cellular senescence (Salminen et al., 2012) indicating better growth of cells in fish from the good-growth muscle group. Full details on the legend of Figure 4.3 (and Figures 4.4, 4.5, 4.6, 4.9, 4.10, 4.11, and 4.12) are provided at:

[http://qiagen.force.com/KnowledgeBase/articles/Basic Technical Q A/Legend](http://qiagen.force.com/KnowledgeBase/articles/Basic_Technical_Q_A/Legend).

Meanwhile, GeneListB and GeneMetaboliteListB (those that were generated by the machine learning model) had a significant enrichment of cancer related pathways, including pathways related to chronic myeloid leukemia, and lung and prostate cancer (Figure 4.2). However, there was not enough information for IPA to accurately predict

whether these pathways were activated or inhibited. One pathway that was significantly affected (predicted activation in the muscle from in the good-growth group) based on the lists used, was B cell receptor signaling (Figure 4.2). When Molecule Activity Predictor was used for the pathway (Figure 4.4), it was shown that upregulation of *bcl-x*, *creb* and downregulation of *cd22* predictively leads to activation of transcription and inhibition of apoptosis. B cell activation is crucial for cellular development and is involved in the activation of many downstream pathways, from inhibition of apoptosis, to accumulation of protein (Daigle et al., 2002; Malissein et al., 2006). These results suggest an increase in the activation of B cell receptor signaling in the muscle of fish from the good-growth group, correlating positively with the phenotype of better muscle growth and reduced cell death in the good-growth group HSB.

Canonical Pathway Analysis of the MetaboliteList led to an enrichment of mainly nucleotide and amino acid synthesis pathways. Since the input list was only 30 metabolites, it was not enough to show significant prediction of inhibition or activation of the pathways. Overall, it suggested activation of gene transcription leading to amino acids synthesis in the muscle of fish in the good-growth group, supporting active anabolism and growth (Figure 4.2 and Figure 4.4).

Upstream Regulators

For GeneListA and GeneMetaboliteListA, a common set of Upstream Regulators predicted to be inhibited in the muscle from fish in the good-growth group were cytokines and regulators that are related to inflammation (including *stat3*, *pparg*, *il1b*, and *il6*) (Shibuya et al., 2002; Kano et al., 2003; Khoury et al., 2008; Liang et al., 2009; Park et al.,

2010; Wang et al., 2011; Biet et al., 2017; Nobs et al., 2017). Each of these regulators were predicted to be inhibited based on the input differential expression lists provided (Figure 4.5). These results also point to key inflammation pathways that would be most likely be inhibited in the good-growth group fish muscle, since *il1b* and *il6* are downregulated. These genes are known to regulate such pathways including the acute phase response signaling pathway discussed previously (Figure 4.1).

For GeneListB and GeneMetaboliteListB, a few Upstream Regulators were found to be significantly affected (*egfr*, *myc*, *notch1*, and *sirt1*) (Figure 4.6). These genes all have crucial roles in development and growth (Balint et al., 2005; Weng et al., 2006; Hershey, 2010; Pedersen et al., 2010; Nimmagadda et al., 2013; Elangovan et al., 2011; Sibia et al., 2016). Indeed, these regulators are all predicted to be activated in fish muscle from the good-growth group (Figure 4.6). This leads to the conclusion that upregulation of these four Upstream Regulators and possibly others as denoted in Figure 4.7 leads to the Canonical Pathway and Functional profile seen in the fish from the good-growth muscle group, including the upregulation of B cell receptor signaling pathway discussed previously (Figure 4.2).

For the MetaboliteList, though Upstream Regulators were found to be significantly enriched, there were none with a clear activation or inhibition prediction based on the dataset provided as shown in Figure 4.7.

Diseases and Functions

Consistent with the Canonical Pathway and Upstream Regulator Analyses for GeneListA and GeneMetaboliteListA, a number of significant functions were identified

using the Diseases and Functions analysis of IPA. Specifically, a couple of functions such as inflammation of body cavity, inflammation of liver, and nephritis were predicted to be activated in fish muscle from the good-growth group (Figure 4.8). This indicates that the fishes in this group may be suffering from liver or possibly another disease (Glass et al., 2003; Mbow and Sarisky, 2004; Dalakas et al., 2005; Tachibana et al., 2007; Dominguez et al., 2009). This may be in agreement with the metabolite profile, which as discussed, consistently points to lipid metabolism being affected in fish from the good-growth group, although the role of disease in this regard is unclear. Furthermore, apoptosis and necrosis were predicted to be inhibited in the fish muscle from the good-growth group pointing to possibly lower cell death and hence better muscle growth phenotype.

There were 12 genes and one metabolite that were shown to be most influential to growth performance of the fish (Figure 4.9 and Table 4.1). This summary figure was generated in the present study and also other studies on muscle growth (Van Buul-Offers et al., 2000; Chaves et al., 2002; Shao et al., 2003; Erlandsson Harris, 2004; Rankinen et al., 2006; Watson et al., 2006; Cyrus et al. 2007; Hu et al., 2012; Gupta et al., 2013). These genes and metabolite were represented on the network map presented in Figure 4.5, and found to be consistent with the predictions determined through the Upstream Regulator Analysis (Figure 4.7).

A number of functions were also found to be consistent with the analyses for GeneListB and GeneMetaboliteListB, including cell cycle progression and growth of connective tissue (Figure 4.8). Similar to the processes from GeneListB, and consistent with the Upstream Regulator network in Figure 4.6, a correlation is seen with the activation of Upstream Regulators (*egfr*, *myc*, *notch1* and *sirt1*) in the fish muscle from the good-growth

group. The activation of functions for these genes are also related to body size, cell cycle control, proliferation of fibroblast cell lines and hyperplasia of cell lines (Figure 4.6).

Comparison Analysis

Two gene lists were provided for the core analyses: 1) genes designated as significant based on statistical inference tests ($p < 0.05$, $n=143$) and 2) genes ranked by Support Vector Machines SVMAttributeEval (top 150 ranked genes). The overlap between these two gene lists was minimal (11 genes found in both datasets, **Objective II, Chapter III**). To test the overlap in Canonical Pathways, Upstream Regulators, and Diseases and Functions, a comparison analysis was performed, in which the predicted activation and inhibition of processes across datasets was compared.

Not surprisingly, there is little overlap of Core Analysis among the two types of gene lists. While there were no Canonical Pathways that were commonly predicted to be affected among all datasets used for Core Analysis, there were many Upstream Regulators that showed similar trends (Figure 4.7). There were a number of regulators, that showed a predicted inhibition, including *tgfb1*, *p38 mapk* and *akt* in the fish muscle from the good-growth group, while one regulator, *sirt1*, was predicted to be activated in these fish based on analysis of both gene lists. Even though the trends were similar, the intensity of inhibition predicted varied between GeneListA and GeneListB.

Another common theme among the Disease and Biological Function Core Analysis of inference-based genes and SVMAttributeEval ranked genes is that apoptosis/cell death is predicted to influence muscle growth (Figure 4.8). Necrosis and cell death are all commonly predicted to be downregulated in fish muscle from the good-growth group.

These results suggest that fish from the poor-growth group could have smaller muscle due to increased cell death relative to the fish from the good-growth group. Interestingly, the differences in GeneListA and GeneListB was overwhelmingly related to activation of inflammation (liver, body cavity, etc.), nephritis, and insulin sensitivity as seen in Figure 4.8. This indicates that individual genes detected by statistical inference-based and machine learning based methods may slightly differ; however, the overall pathways represented are similar.

BioProfiler Analysis

To determine the connections between GeneMetaboliteListA and GeneMetaboliteListB to specific Diseases and Functions, BioProfiler analysis was run on the two data lists independently. Running the BioProfiler within IPA on these datasets tabulates all the information in the Qiagen KnowledgeBase, where all the molecules in the list provided are linked to all possible Diseases and Functions for which any findings have been curated in the KnowledgeBase. Based on the interest in muscle phenotype and impact seen on liver from both gene lists as well as metabolites from the Core Analysis. The lists were further filtered down to only show data specific to growth of muscle and response of liver. The filtered molecules and their Disease or Function associations were displayed as independent networks to not only better visualize the analysis, but also to overlay activity predictions and make better inferences coming from the specific lines of evidence from the KnowledgeBase. Only gene *agtr2* and the metabolite xanthine were common between the two combined gene and metabolite datasets that were linked in the BioProfiler Analysis for

possible Diseases and Functions (Figures 4.10 and 4.11). This is not surprising, since the initial input gene lists A and B had minimal overlap as well as discussed previously.

The network for GeneMetaboliteListA as represented in Figure 4.10 indicated that the proliferation of myofibroblasts was predicted to be activated in fish muscle from the good-growth group, but the majority of the other functions related to growth or proliferation of the muscle cells were generally inhibited. Therefore, some of the identified pathways may support or contrast the phenotype observed in fish from the good-growth group, where muscle tissue samples were observed to have higher hyperplastic activities (**Objective 2, Chapter III**, Figure 2.4B). Inflammation of liver also was significantly activated in fish muscle from the good-growth group, which may suggest that these fish were suffering from some form of liver disease. Most of the pathways were represented by the connected lines in Figure 4.10, where orange and blue connected lines between the molecules and the Diseases and Function indicated agreement in the findings, whereas the yellow line indicated disagreement with the KnowledgeBase (Nam et al., 2000; Rossi et al., 2002; Takeda et al., 2004; Jung et al., 2011; Wang et al., 2012; Liu et al., 2013; McGrath et al., 2014; Quan et al., 2014).

In contrast to the above inferences made using statistical ranking methods, the network from BioProfiler Analysis for GeneMetaboliteListB generated using machine learning-based ranking as represented in Figure 4.11 predicted highly activated pathways in proliferation of myofibroblasts, proliferation of smooth muscle cells and development of muscle cell lines and only proliferation of vascular smooth muscle cells were predicted to be inhibited in fish from the good-growth group. Most of the pathways from the dataset are completely consistent with the known and curated findings found in IPA KnowledgeBase

(Du et al., 2008; Lu et al., 2008; Golden-Mason et al., 2009; Miyake et al., 2011; Cattaruzza et al., 2012; Park et al., 2012; Bian et al., 2013; Shi et al., 2017). Additionally, and unlike the Core Analysis and the Comparison Analysis observations, the BioProfiler Analysis predicted increased Inflammation of liver. This result overall agrees with the phenotype observed in fish muscle from the good-growth group, which further strengthens the posited hypothesis that the machine learning SVMAttributeEval-based inference may be a superior approach for this particular type of pathway analysis compared to using statistics and *p*-value-based ranking. When the network was represented in a different view referred to as Organic View as shown in Figure 4.12 it can be seen that GeneMetaboliteListB molecules were spread out into five major independent networks. This illustration specifies two verdicts, one that SVMAttributeEval-based method prioritized genes with independent vector connections. The other verdict was that BioProfiler can help to infer the importance even with the slightest variation in gene expression observations, that could not be detected by the other analysis. For example, the inflammation of liver that was predicted from fish muscle data of the good-growth group fish and that was identified and discussed above. This highlights the power of pathway approach, as the physiology of the liver was inferred based on that data collected from the muscle.

In conclusion, the shared and curated biological findings between the expressed gene lists and the metabolite lists from the white muscle tissue of HSB, identified potential candidates and pathways that influence growth performance for further experimental testing. The SVMAttributeEval machine learning ranked list collectively suggested a promotion of cell proliferation and perhaps cell differentiation generally in fish muscle from the good-growth group fish, which would be with an agreement with the observation

of the superior growth and muscle hyperplasia. Overall, the gene lists provided more meaningful biological information regarding muscle growth, compared to the metabolite list. The gene list ranked by statistical approaches provided a slightly different pathway analysis compared to the machine learning-ranked approach. In this case, cell proliferation and cell differentiation were similarly identified as activated in the fish muscle from the good-growth group. Inflammation, however, particularly of the liver, was identified as an important pathway that was predominantly activated in fish muscle from the good-growth group. Candidate genes and metabolites representing this pathway were not reliably identified as promoting inflammation based on the SVMAttributeEval machine learning-based ranking. Therefore, the role of inflammation in growth performance of HSB is unclear. Nevertheless, inflammation pathways may be involved in this process and chronic inflammation is known to impair growth in humans (Cirillo et al., 2017). Although the underlying mechanism of inflammatory suppression of growth is not understood in fish from this study, suboptimal nutrition conditions might be one of the factors influencing this and cortisol is an anti-inflammatory hormone (Cirillo et al., 2017). There also was compelling evidence that cell death was increased in fish muscle from the poor-growth group relative to fish muscle from the good-growth group possibly related to sub-optimal growth performance.

The implications of these findings could provide a potential mechanism by which growth is affected in the fish muscle from the poor-growth group, which appear to be less healthy and suffer from increased apoptosis. Based on the SVMAttributeEval gene ranking, a number of key developmental factors were identified to be putatively activated in the fish muscle from the good-growth group. In particular, *sirt1* was commonly predicted to be

activated in fish muscle from the good-growth group across both the datasets ranked by statistics and machine learning. Coupled with the observations of increased apoptosis and inflammation biomarkers in the fish from the poor-growth group, activation of *sirt1* may play a role in protecting against inflammation or cell death leading to superior growth.

References

- Ahmad, G., Sial, G. Z. K., Ramadori, P., Dudas, J., Batusic, D. S., & Ramadori, G. (2011).** Changes of hepatic lactoferrin gene expression in two mouse models of the acute phase reaction. *The international journal of biochemistry & cell biology*. 43(12): 1822-1832.
- Balint, K., Xiao, M., Pinnix, C.C., Soma, A., Veres, I., Juhasz, I., Brown, E.J., Capobianco, A.J., Herlyn, M. and Liu, Z.J. (2005).** Activation of Notch1 signaling is required for β -catenin-mediated human primary melanoma progression. *The Journal of clinical investigation*. 115(11): 3166-3176.
- Bian, Z., Peng, Y., You, Z., Wang, Q., Miao, Q., Liu, Y., Han, X., Qiu, D., Li, Z. and Ma, X., (2013).** CCN1 expression in hepatocytes contributes to macrophage infiltration in nonalcoholic fatty liver disease in mice. *Journal of lipid research*. 54(1): 44-54.
- Biet, F., Locht, C., and Kremer, L. (2002).** Immunoregulatory functions of interleukin 18 and its role in defense against bacterial pathogens. *Journal of molecular medicine*. 80(3): 147-162.
- Carrillo, J. A., He, Y., Li, Y., Liu, J., Erdman, R. A., Sonstegard, T. S., & Song, J. (2016).** Integrated metabolomic and transcriptome analyses reveal finishing forage affects metabolic pathways related to beef quality and animal welfare. *Scientific reports*, 6.
- Cattaruzza, M., Nogoy, N., Wojtowicz, A., and Hecker, M. (2012).** Zinc finger motif-1 antagonizes PDGF-BB-induced growth and dedifferentiation of vascular smooth muscle cells. *The FASEB Journal*. 26(12): 4864-4875.

- Chaves, F. J., Giner, V., Corella, D., Pascual, J., Marin, P., Armengod, M. E., and Redon, J.** (2002). Body weight changes and the A-6G polymorphism of the angiotensinogen gene. *International journal of obesity*. 26(9): 1173-1178.
- Choy, E. H., and Panayi, G. S.** (2001). Cytokine pathways and joint inflammation in rheumatoid arthritis. *New England Journal of Medicine*. 344(12): 907-916.
- Cirillo, F., Lazzeroni, P., Sartori, C., and Street, M. E.** (2017). Inflammatory Diseases and Growth: Effects on the GH-IGF Axis and on Growth Plate. *International journal of molecular sciences*. 18(9): 1878.
- Cyrus, T., Yao, Y., Ding, T., Dogné, J. M., and Pratico, D.** (2007). Thromboxane receptor blockade improves the antiatherogenic effect of thromboxane A2 suppression in LDLR KO mice. *Blood*. 109(8): 3291-3296.
- Daigle, I., Yousefi, S., Colonna, M., Green, D. R., and Simon, H. U.** (2002). Death receptors bind SHP-1 and block cytokine-induced anti-apoptotic signaling in neutrophils. *Nature medicine*. 8(1): 61-67.
- Dalakas, E., Newsome, P. N., Harrison, D. J., and Plevris, J. N.** (2005). Hematopoietic stem cell trafficking in liver injury. *The FASEB Journal*. 19(10): 1225-1231.
- Dominguez, M., Miquel, R., Colmenero, J., Moreno, M., García-Pagán, J.C., Bosch, J., Arroyo, V., Ginès, P., Caballería, J. and Bataller, R.** (2009). Hepatic expression of CXC chemokines predicts portal hypertension and survival in patients with alcoholic hepatitis. *Gastroenterology*. 136(5): 1639-1650.
- Du, M., Perry, R.L., Nowacki, N.B., Gordon, J.W., Salma, J., Zhao, J., Aziz, A., Chan, J., Siu, K.M. and McDermott, J.C.** (2008). Protein kinase A represses skeletal myogenesis

by targeting myocyte enhancer factor 2D. *Molecular and cellular biology*, 28(9), 2952-2970.

Elangovan, S., Ramachandran, S., Venkatesan, N., Ananth, S., Gnana-Prakasam, J.P., Martin, P.M., Browning, D.D., Schoenlein, P.V., Prasad, P.D., Ganapathy, V. and Thangaraju, M. (2011). SIRT1 is essential for oncogenic signaling by estrogen/estrogen receptor α in breast cancer. *Cancer research*. 71(21): 6654-6664.

Erlandsson Harris, H., and Andersson, U. (2004). Mini-review: the nuclear protein HMGB1 as a proinflammatory mediator. *European journal of immunology*. 34(6), 1503-1512.

Formichi, P., Radi, E., Battisti, C., Tarquini, E., Leonini, A., Di Stefano, A., and Federico, A. (2006). Human fibroblasts undergo oxidative stress-induced apoptosis without internucleosomal DNA fragmentation. *Journal of cellular physiology*. 208(2): 289-297.

Fridlyanskaya, I., Alekseenko, L., and Nikolsky, N. (2015). Senescence as a general cellular response to stress: a mini-review. *Experimental gerontology*, 72: 124-128.

Gabay, C., and Kushner, I. (1999). Acute-phase proteins and other systemic responses to inflammation. *New England journal of medicine*. 340(6): 448-454.

Glass, W. G., Rosenberg, H. F., and Murphy, P. M. (2003). Chemokine regulation of inflammation during acute viral infection. *Current opinion in allergy and clinical immunology*. 3(6): 467-473.

Golden-Mason, L., Palmer, B.E., Kassam, N., Townshend-Bulson, L., Livingston, S., McMahon, B.J., Castelblanco, N., Kuchroo, V., Gretch, D.R. and Rosen, H.R. (2009). Negative immune regulator Tim-3 is overexpressed on T cells in hepatitis C

- virus infection and its blockade rescues dysfunctional CD4+ and CD8+ T cells. *Journal of virology*. 83(18): 9122-9130.
- Gupta, V., Vinay, D.G., Sovio, U., Rafiq, S., Kumar, M.V.K., Janipalli, C.S., Evans, D., Mani, K.R., Sandeep, M.N., Taylor, A. and Kinra, S.** (2013). Association study of 25 type 2 diabetes related Loci with measures of obesity in Indian sib pairs. *PLoS One*. 8(1).
- Gut, P., and Verdin, E.** (2013). The nexus of chromatin regulation and intermediary metabolism. *Nature*. 502(7472): 489-498.
- Harris, E.** (2014). Biochemical facts behind the definition and properties of METABOLITES.
- Hershey, J. W.** (2010). Regulation of protein synthesis and the role of eIF3 in cancer. *Brazilian Journal of Medical and Biological Research*. 43(10): 920-930.
- Holmes, E., Wilson, I. D., and Nicholson, J. K.** (2008). Metabolic phenotyping in health and disease. *Cell*. 134(5): 714-717.
- Hu, M., Chu, W.C.W., Yamashita, S., Yeung, D.K.W., Shi, L., Wang, D., Masuda, D., Yang, Y. and Tomlinson, B.** (2012). Liver fat reduction with niacin is influenced by DGAT-2 polymorphisms in hypertriglyceridemic patients. *Journal of lipid research*. 53(4): 802-809.
- Joseph, S. B., Castrillo, A., Laffitte, B. A., Mangelsdorf, D. J., & Tontonoz, P.** (2003). Reciprocal regulation of inflammation and lipid metabolism by liver X receptors. *Nature medicine*. 9(2): 213-219.
- Jung, J. H., Park, J. H., Jee, M. H., Keum, S. J., Cho, M. S., Yoon, S. K., and Jang, S. K.** (2011). Hepatitis C virus infection is blocked by HMGB1 released from virus-infected cells. *Journal of virology*. 85(18): 9359-9368.

- Kano, A., Wolfgang, M.J., Gao, Q., Jacoby, J., Chai, G.X., Hansen, W., Iwamoto, Y., Pober, J.S., Flavell, R.A. and Fu, X.Y.** (2003). Endothelial Cells Require STAT3 for Protection against Endotoxin-induced Inflammation. *The Journal of experimental medicine*. 198(10): 1517-1525.
- Khoury, M., Escriou, V., Courties, G., Galy, A., Yao, R., Largeau, C., Scherman, D., Jorgensen, C. and Apparailly, F.** (2008). Efficient suppression of murine arthritis by combined anticytokine small interfering RNA lipoplexes. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*. 58(8): 2356-2367.
- Kyriakis, J. M., and Avruch, J.** (2001). Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation. *Physiological reviews*. 81(2): 807-869.
- Liang, B., Song, Z., Wu, B., Gardner, D., Shealy, D., Song, X. Y., and Wooley, P. H.** (2009). Evaluation of anti-IL-6 monoclonal antibody therapy using murine type II collagen-induced arthritis. *Journal of Inflammation*. 6(1): 10.
- Liu, Q. F., Yu, H. W., You, L., Liu, M. X., Li, K. Y., and Tao, G. Z.** (2013). Apelin-13-induced proliferation and migration induced of rat vascular smooth muscle cells is mediated by the upregulation of Egr-1. *Biochemical and biophysical research communications*. 439(2): 235-240.
- Lu, X., Kovalev, G.I., Chang, H., Kallin, E., Knudsen, G., Xia, L., Mishra, N., Ruiz, P., Li, E., Su, L. and Zhang, Y.** (2008). Inactivation of NuRD component Mta2 causes abnormal T cell activation and lupus-like autoimmune disease in mice. *Journal of Biological Chemistry*. 283(20): 13825-13833.

- Malissein, E., Verdier, M., Ratinaud, M. H., and Troutaud, D.** (2006). Activation of Bad trafficking is involved in the BCR-mediated apoptosis of immature B cells. *Apoptosis*. 11(6): 1003-1012.
- Mbow, M. L., and Sarisky, R. T.** (2004). What is disrupting IFN- α 's antiviral activity?. *Trends in biotechnology*. 22(8): 395-399.
- McGrath, K.C., Li, X.H., Whitworth, P.T., Kasz, R., Tan, J.T., McLennan, S.V., Celermajer, D.S., Barter, P.J., Rye, K.A. and Heather, A.K.** (2014). High density lipoproteins improve insulin sensitivity in high-fat diet-fed mice by suppressing hepatic inflammation. *Journal of lipid research*. 55(3), 421-430.
- Miyake, H., Maeda, K., Asai, N., Shibata, R., Ichimiya, H., Isotani-Sakakibara, M., Yamamura, Y., Kato, K., Enomoto, A., Takahashi, M. and Murohara, T.** (2011). The actin-binding protein Girdin and its Akt-mediated phosphorylation regulate neointima formation after vascular injury. *Circulation research*. 108(10): 1170-1179.
- Nam, T. J., Busby Jr, W. H., Rees, C., and Clemmons, D. R.** (2000). Thrombospondin and osteopontin bind to insulin-like growth factor (IGF)-binding protein-5 leading to an alteration in IGF-I-stimulated cell growth. *Endocrinology*. 141(3): 1100-1106.
- Nimmagadda, V.K., Bever, C.T., Vattikunta, N.R., Talat, S., Ahmad, V., Nagalla, N.K., Trisler, D., Judge, S.I., Royal, W., Chandrasekaran, K. and Russell, J.W.** (2013). Overexpression of SIRT1 protein in neurons protects against experimental autoimmune encephalomyelitis through activation of multiple SIRT1 targets. *The Journal of Immunology*. 190(9): 4595-4607.
- Nobs, S.P., Natali, S., Pohlmeier, L., Okreglicka, K., Schneider, C., Kurrer, M., Sallusto, F. and Kopf, M.** (2017). PPAR γ in dendritic cells and T cells drives pathogenic type-2

- effector responses in lung inflammation. *Journal of Experimental Medicine*. 214(10): 3015-3035.
- Park, E.J., Lee, J.H., Yu, G.Y., He, G., Ali, S.R., Holzer, R.G., Österreicher, C.H., Takahashi, H. and Karin, M.** (2010). Dietary and genetic obesity promote liver inflammation and tumorigenesis by enhancing IL-6 and TNF expression. *Cell*. 140(2): 197-208.
- Park, J., Kang, W., Ryu, S.W., Kim, W.I., Chang, D.Y., Lee, D.H., Park, D.Y., Choi, Y.H., Choi, K., Shin, E.C. and Choi, C.** (2012). Hepatitis C virus infection enhances TNF α -induced cell death via suppression of NF- κ B. *Hepatology*. 56(3): 831-840.
- Pedersen, M. W., Jacobsen, H. J., Koefoed, K., Hey, A., Pyke, C., Haurum, J. S., and Kragh, M.** (2010). Sym004: a novel synergistic anti-epidermal growth factor receptor antibody mixture with superior anticancer efficacy. *Cancer research*. 70(2), 588-597.
- Quan, J., Liu, J., Gao, X., Liu, J., Yang, H., Chen, W., Li, W., Li, Y., Yang, W. and Wang, B.** (2014). Palmitate induces interleukin-8 expression in human aortic vascular smooth muscle cells via T oll-like receptor 4/nuclear factor- κ B pathway (TLR4/NF- κ B 通路介导了软脂酸诱导人主动脉血管平滑肌细胞白细胞介素-8 的基因表达). *Journal of diabetes*. 6(1): 33-41.
- Rankinen, T., Zuberi, A., Chagnon, Y.C., Weisnagel, S.J., Argyropoulos, G., Walts, B., Pérusse, L. and Bouchard, C.** (2006). The human obesity gene map: the 2005 update. *Obesity*. 14(4): 529-644.
- Rasmussen, B. B., and Wolfe, R. R.** (1999). Regulation of fatty acid oxidation in skeletal muscle. *Annual review of nutrition*. 19(1): 463-484.

- Rossi, F., Ferraresi, A., Romagni, P., Silvestroni, L., and Santiemma, V. (2002).**
Angiotensin II stimulates contraction and growth of testicular peritubular myoid cells in vitro. *Endocrinology*. 143(8): 3096-3104.
- Salminen, A., Kauppinen, A., and Kaarniranta, K. (2012).** Emerging role of NF- κ B signaling in the induction of senescence-associated secretory phenotype (SASP). *Cellular signaling*. 24(4): 835-845.
- Shao, J. S., Cheng, S. L., Charlton-Kachigian, N., Loewy, A. P., and Towler, D. A. (2003).**
Teriparatide (human parathyroid hormone (1-34)) inhibits osteogenic vascular calcification in diabetic low density lipoprotein receptor-deficient mice. *Journal of Biological Chemistry*. 278(50): 50195-50202.
- Shibuya, A., Wada, K., Nakajima, A., Saeki, M., Katayama, K., Mayumi, T., Kadowaki, T., Niwa, H. and Kamisaki, Y. (2002).** Nitration of PPAR γ inhibits ligand-dependent translocation into the nucleus in a macrophage-like cell line, RAW 264. *FEBS letters*. 525(1-3): 43-47.
- Shi, S.Y., Luk, C.T., Schroer, S.A., Kim, M.J., Dodington, D.W., Sivasubramaniam, T., Lin, L., Cai, E.P., Lu, S.Y., Wagner, K.U. and Bazinet, R.P. (2017).** Janus Kinase 2 (JAK2) dissociates hepatosteatosis from hepatocellular carcinoma in mice. *Journal of Biological Chemistry*. 292(9): 3789-3799.
- Sibilia, M., Wagner, B., Hoebertz, A., Elliott, C., Marino, S., Jochum, W., and Wagner, E. F. (2016).** Correction: Mice humanised for the EGF receptor display hypomorphic phenotypes in skin, bone and heart. *Development*. 143(24): 4755-4755.

- Tachibana, Y., Nakamoto, Y., Mukaida, N., and Kaneko, S.** (2007). Intrahepatic interleukin-8 production during disease progression of chronic hepatitis C. *Cancer letters*. 251(1): 36-42.
- Takeda, H., Katagata, Y., Hozumi, Y., and Kondo, S.** (2004). Effects of angiotensin II receptor signaling during skin wound healing. *The American journal of pathology*. 165(5): 1653-1662.
- Van Buul-Offers, S.C., Van Kleffens, M., Koster, J.G., Lindenberg-Kortleve, D.J., Gresnigt, M.G., Drop, S.L.S., Hoogerbrugge, C.M., Bloemen, R.J., Koedam, J.A. and Van Neck, J.W.** (2000). Human insulin-like growth factor (IGF) binding protein-1 inhibits IGF-I-stimulated body growth but stimulates growth of the kidney in snell dwarf mice. *Endocrinology*. 141(4): 1493-1499.
- Wang, H., Lafdil, F., Kong, X., and Gao, B.** (2011). Signal transducer and activator of transcription 3 in liver diseases: a novel therapeutic target. *International journal of biological sciences*. 7(5): 536.
- Wang, J., Razuvaev, A., Folkersen, L., Hedin, E., Roy, J., Brismar, K., and Hedin, U.** (2012). The expression of IGFs and IGF binding proteins in human carotid atherosclerosis, and the possible role of IGF binding protein-1 in the regulation of smooth muscle cell proliferation. *Atherosclerosis*. 220(1): 102-109.
- Watson, C. S., Bialek, P., Anzo, M., Khosravi, J., Yee, S. P., and Han, V. K.** (2006). Elevated circulating insulin-like growth factor binding protein-1 is sufficient to cause fetal growth restriction. *Endocrinology*. 147(3): 1175-1186.
- Weng, A.P., Millholland, J.M., Yashiro-Ohtani, Y., Arcangeli, M.L., Lau, A., Wai, C., del Bianco, C., Rodriguez, C.G., Sai, H., Tobias, J. and Li, Y.** (2006). c-Myc is an

important direct target of Notch1 in T-cell acute lymphoblastic leukemia/lymphoma. *Genes & development*. 20(15): 2096-2109.

Zhou, X., Ding, Y., and Wang, Y. (2012). Proteomics: present and future in fish, shellfish and seafood. *Reviews in Aquaculture*. 4(1): 11-20.

Table 4.1. Qiagen IPA Diseases and Function Analysis of muscle GeneListA (ranked by *p*-value) and GeneMetaboliteListA (ranked by *p*-value) identified some molecules that were predicted to either activate or inhibit key pathways that were related to the growth performance of hybrid striped bass muscle from the good-growth group based on the gene and metabolite signatures provided. These genes and metabolites were represented in the network map presented in Figure 4.9. The table shows the category of each molecule, their elevation level, and their direct effect on the growth performance.

No.	Molecule	Category	Elevation level	Relationship Type	Relationship Leads to
1	<i>cxcr4</i>	G-Protein Coupled Receptor	Decreased	Causation	Inhibition of Growth
2	<i>hmgb1</i>	Transcription Regulator	Decreased	Causation	Activation of Growth
3	<i>igfbp1</i>	Other - Unspecified	Decreased	Causation	Activation of Growth
4	<i>pttg1</i>	Transcription Regulator	Increased	Causation	Activation of Growth
5	<i>ucp3</i>	Transporter	Decreased	Causation	Activation of Growth
6	Nicotinic acid	Chemical Compound	Decreased	Causation	Activation of Growth

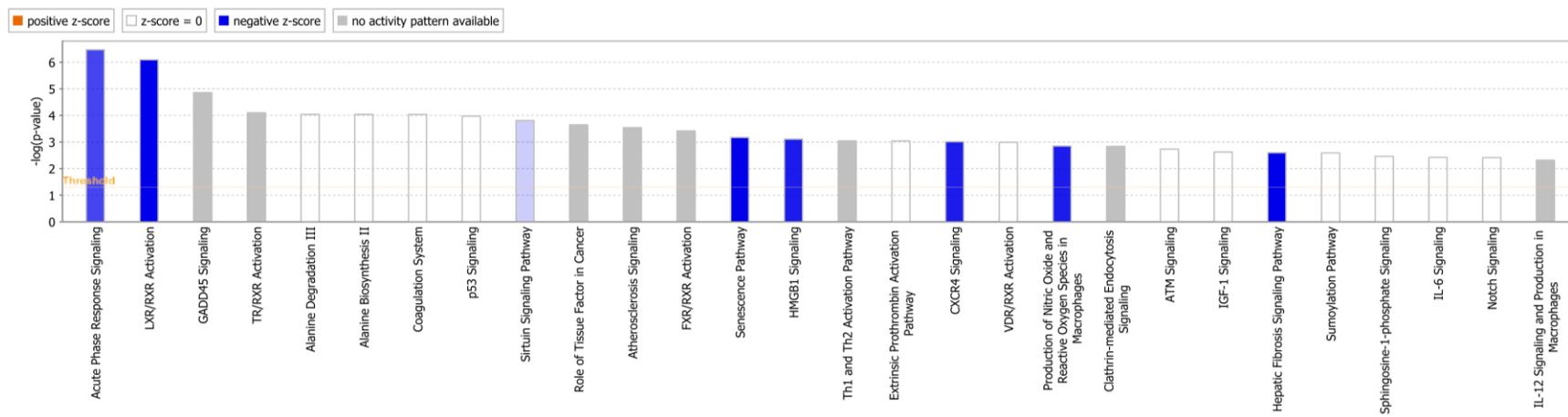


Figure 4.1. Bar charts of significantly-enriched Qiagen IPA Canonical Pathways ($p \leq 0.05$) from Core Analysis of GeneMetaboliteListA (ranked by p -value) for hybrid striped bass muscle from the good-growth group. Each bar represents a canonical pathway's p -value on a negative logarithmic scale, such that the taller bars are more significant than the shorter bars. Bars are shaded according to their z-score activity predictions. Orange: pathways with positive z-scores (upregulated), blue: pathways with negative z-scores (down regulated), white: pathways that have z-score of 0, indicating that the differential gene expression data did not allow for a clear determination of the activity prediction; in other words, the weight of the evidence for a prediction of activation is equal to that of inhibition, and grey: pathways for which no activity predictions can currently be made due to a lack of information in the Ingenuity Knowledge Base.

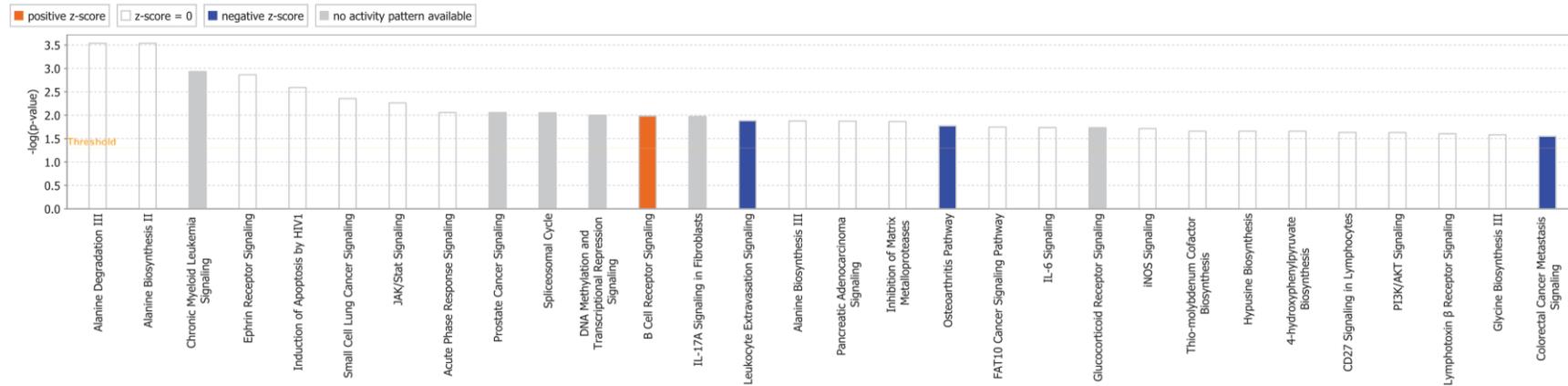


Figure 4.2. Bar charts of significantly-enriched Qiagen IPA Canonical Pathways ($p \leq 0.05$) from core analysis of GeneMetaboliteListB (ranked by the Machine Learning-based SVMAttributeEval) for hybrid striped bass muscle from the good-growth group. Each bar represents a canonical pathway's p -value on a negative logarithmic scale, such that the taller bars are more significant than the shorter bars. Bars are shaded according to their z-score activity predictions. Orange: pathways with positive z-scores (upregulated), blue: pathways with negative z-scores (downregulated), white: pathways that have z-score of 0, indicating that the differential gene expression data did not allow for a clear determination of the activity prediction; in other words, the weight of the evidence for a prediction of activation is equal to that of inhibition, and grey: pathways for which no activity predictions can currently be made due to a lack of information in the Ingenuity Knowledge Base.

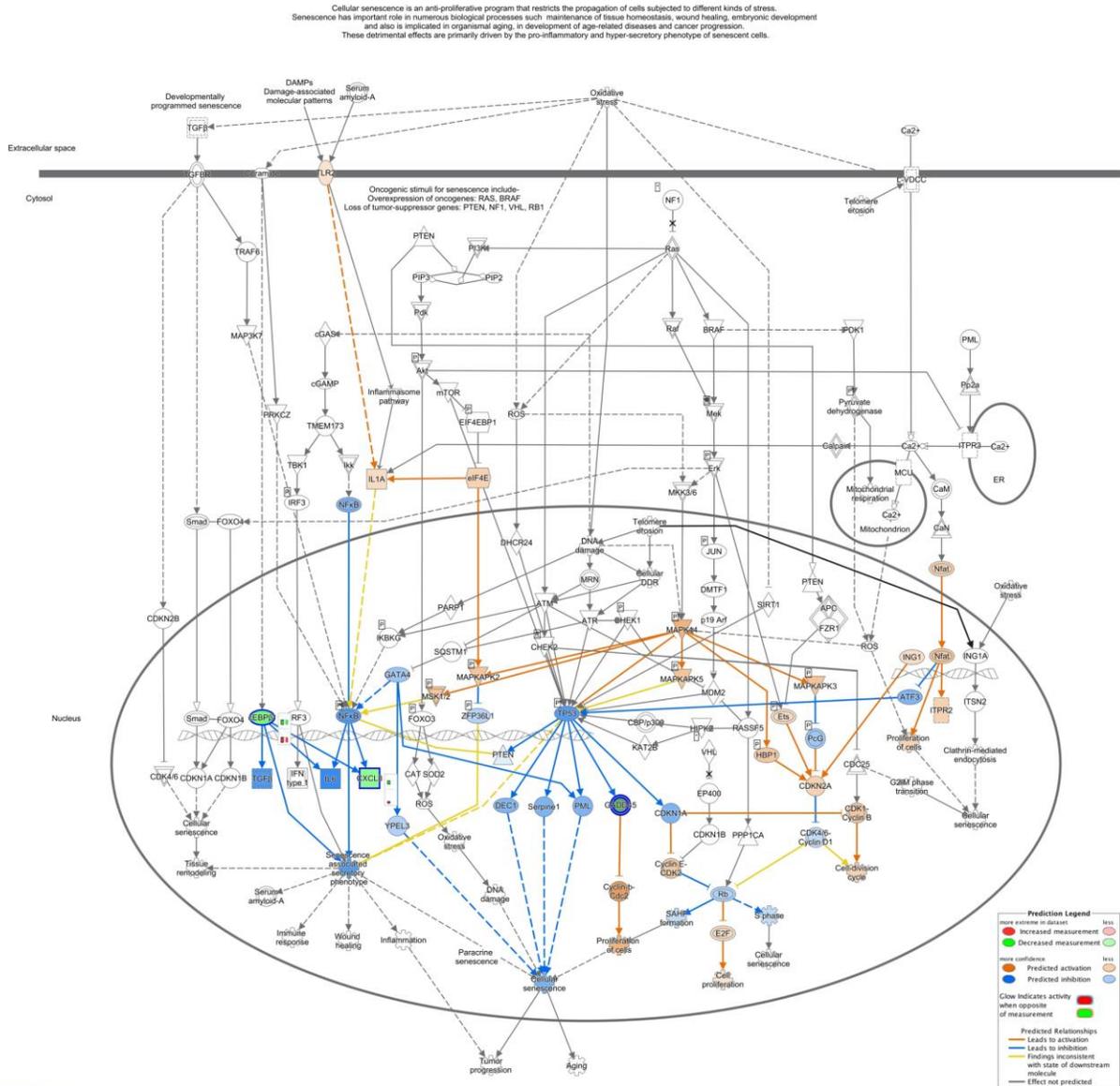


Figure 4.3. Cellular senescence pathway enriched in GeneMetaboliteListA (ranked by *p*-value) for hybrid striped bass muscle from the different growth groups. The overlapping molecules from the gene list (green for downregulated and red for upregulated) were used to predict the activity pattern of the pathway using Molecule Activity Predictor (MAP) in IPA. It predicted gene expression which probably leads to inhibition of cellular senescence (blue) and activation of proliferation of cells (orange). Downregulation (green) refers to lower expression in fish muscle from the good-growth group relative to fish muscle from the poor-growth group, while upregulation (red) refers to increased expression in fish muscle from the good-growth group relative to fish muscle from the poor-growth group.

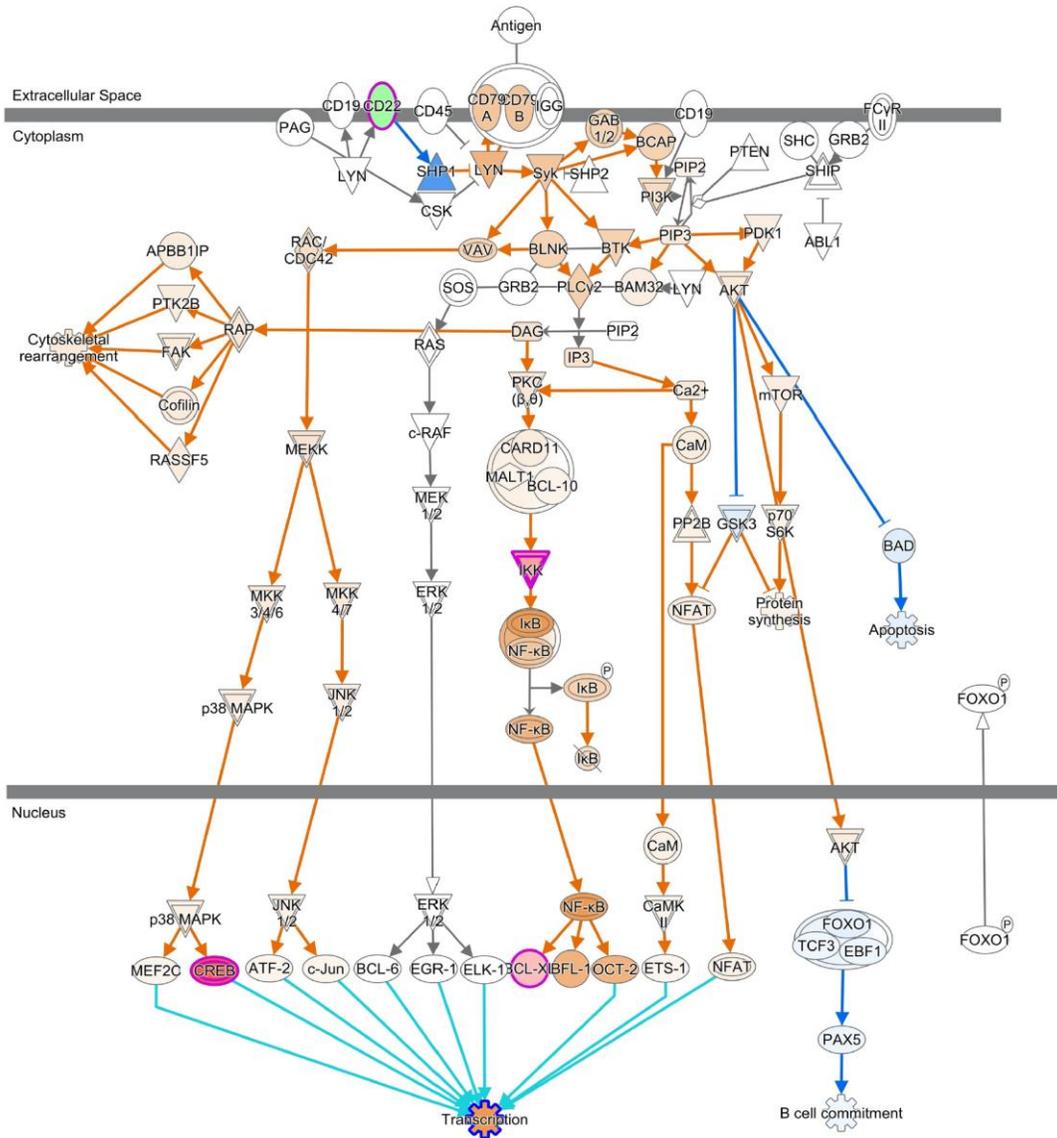


Figure 4.4. B Cell receptor signaling pathway enriched in GeneMetaboliteListB (ranked by the Machine Learning-based SVMAttributeEval) for hybrid striped bass muscle from the different growth groups. The overlapping molecules from the gene list (green for downregulated and red for upregulated) were used to predict the activity pattern of the pathway using Molecule Activity Predictor (MAP) in IPA. It predicted gene expression which probably leads to inhibition of apoptosis (blue) and activation of transcription (orange). Downregulation (green) refers to lower expression in fish muscle from the good-growth group relative to fish muscle from the poor-growth group, while upregulation (red) refers to increased expression in fish muscle from the good-growth group relative to fish muscle from the poor-growth group.

IL1B,IL6,PPARG,STAT3 7

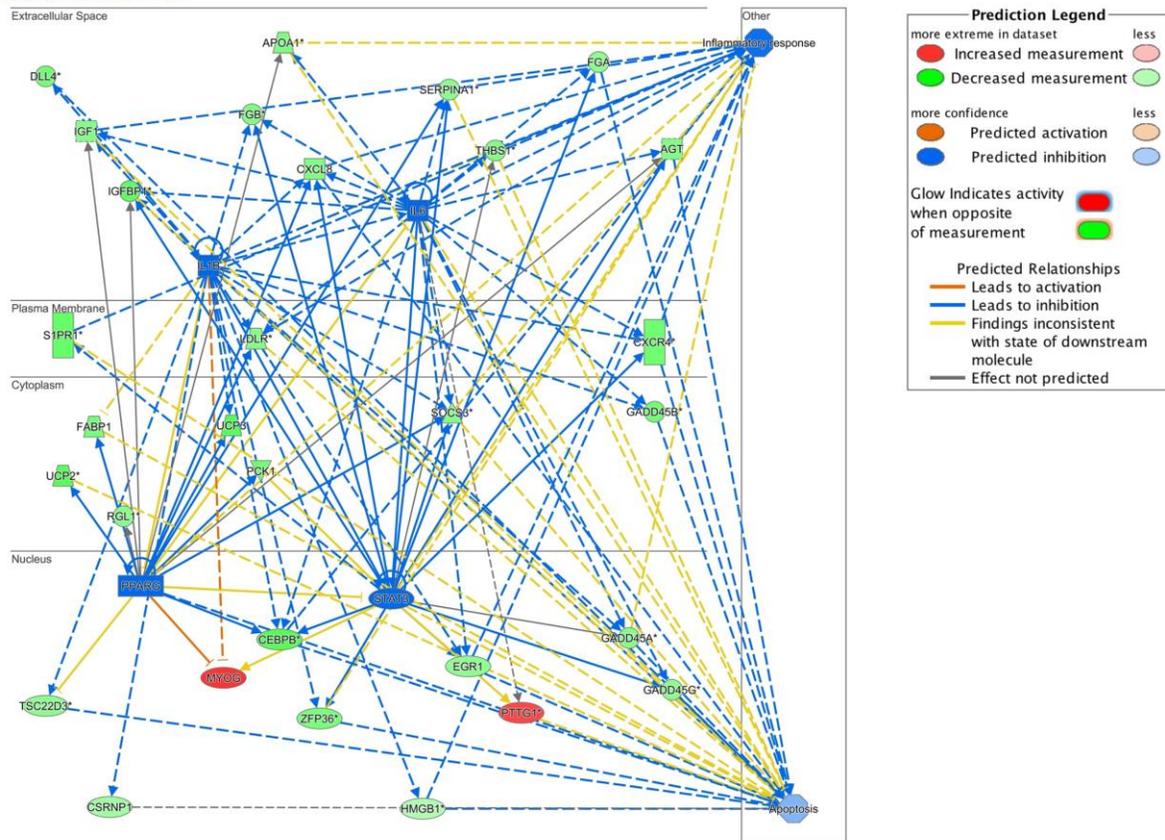


Figure 4.5. Upstream Regulator analysis of GeneMetaboliteListA (ranked by *p*-value) identified a number of factors (*il6*, *il1b*, *stat3*, and *pparg*) that were predicted to be inhibited based on gene expression in muscle of hybrid striped bass. The network indicates direct relationships from five inhibited factors (colored blue due to predicted inhibition in fish from the good-growth group). Connections are drawn between molecules that have been found to have relationships in the literature. Arrows indicate activation and perpendicular lines indicate inhibition. Genes and metabolites from the input dataset with a relationship to the Upstream Regulators are colored by fold change (red: upregulated in fish muscle from the good-growth group; green; downregulated in fish muscle from the good-growth group). The pathway also includes curated functions (mass of organism, apoptosis, and inflammatory response) that are associated with the molecules. Based on the signature of input genes, apoptosis and inflammatory response were predicted to be inhibited (blue) in fish muscle from the good-growth group relative to fish from the poor-growth group. Conversely, mass of organism was predicted to be activated (orange).

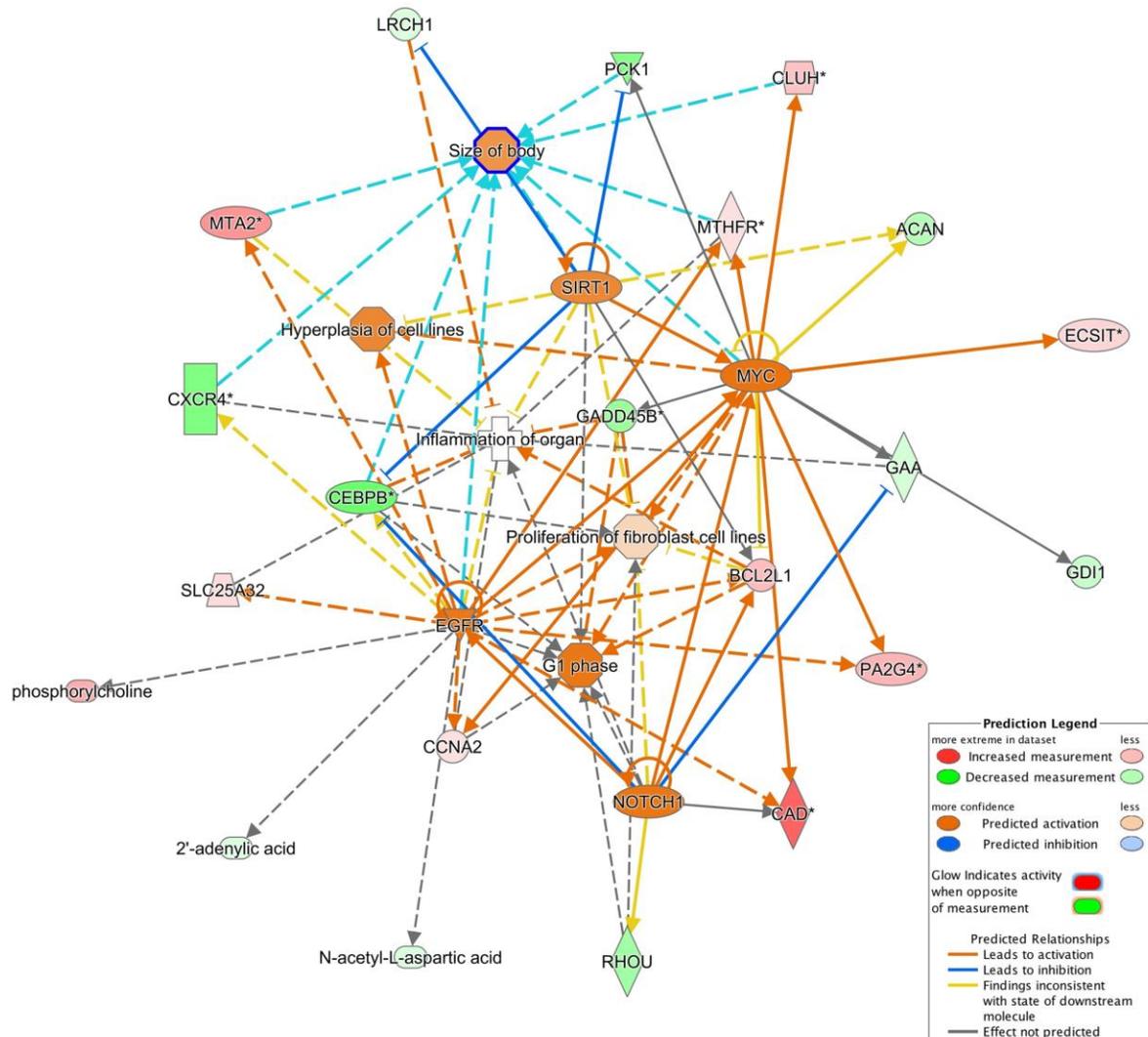


Figure 4.6. Upstream Regulator Analysis of GeneMetaboliteListB (ranked by the Machine Learning-based SVMAttributeEval) identified a number of developmental regulators (*myc*, *sirt1*, *egfr*, *notch1*) that were predicted to be activated based on the gene signature provided from hybrid striped bass muscle. Network map indicates the direct relationships of these factors (colored orange due to predicted activation in fish from the good growth group). Genes and metabolites from the input dataset with a relationship to the Upstream Regulators are colored by fold change (red: upregulated in fish muscle from the good-growth group; green; downregulated in fish muscle from the good-growth group). The pathway includes curated functions (hyperplasia of cell lines; proliferation of fibroblasts, size of body) and that are associated with the molecules. Based on the signature of input genes and the relationships with the upstream regulators, each of these functions were predicted to be activated (orange).

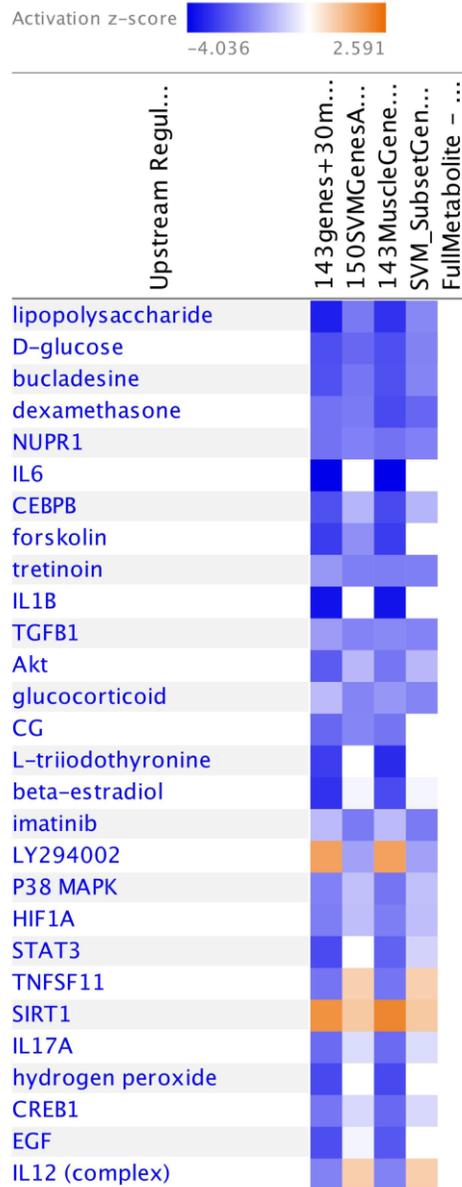


Figure 4.7. Comparison analysis of IPA Upstream regulators of muscle from hybrid striped bass in the good-growth group. Heatmaps were generated to show the key Upstream Regulators that were identified in Core Analyses from each gene list (order of columns from left to right: Upstream Regulators, GeneMetaboliteListA, GeneListA, GeneMetaboliteListB, GeneListB, MetaboliteList). Blue squares indicate predicted inhibition of the upstream regulator or function, while orange indicates the component is predicted to be activated based on the dataset provided. Intensity of color is scaled by z-score, with legend provided above heatmap.

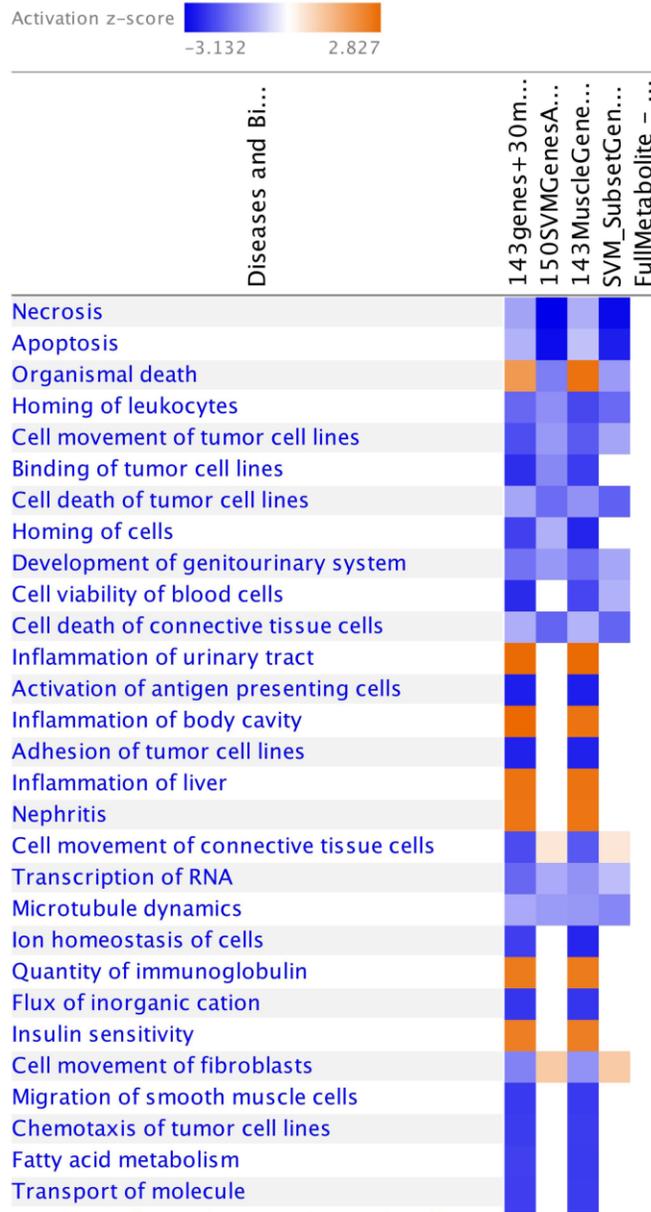


Figure 4.8. Comparison Analysis of IPA Diseases and Biological Functions of muscle from hybrid stiped bass in the good-growth group. Heatmaps were generated to show the key Diseases and Functions that were identified in Core Analyses from each gene list (order of columns from left to right: Disease and Biological Functions, GeneMetaboliteListA, GeneListA, GeneMetaboliteListB, GeneListB, MetaboliteList). Blue squares indicate predicted inhibition of the upstream regulator or function, while orange indicates the component is predicted to be activated based on the dataset provided. Intensity of color is scaled by z-score, with legend provided above heatmap.

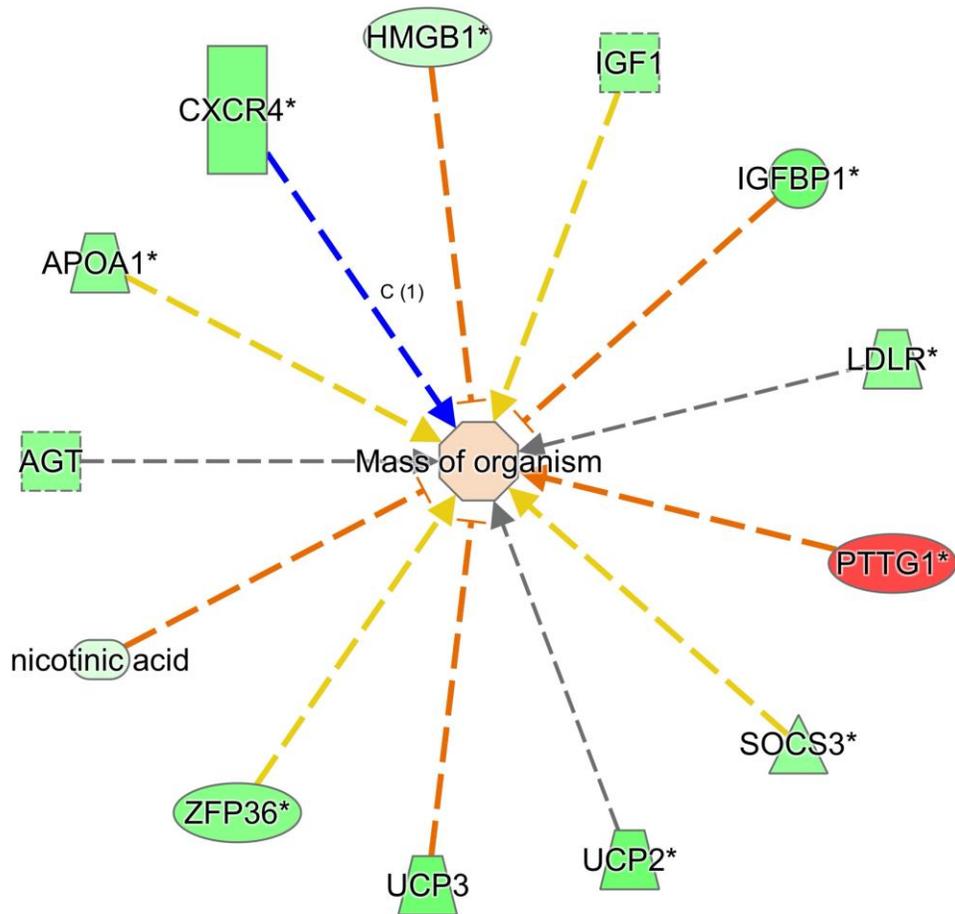


Figure 4.9. One metabolite and 12 genes shown to most influence the growth performance hybrid striped bass (mass of organism) in GeneMetaboliteListA (ranked by *p*-value). Table 4.1 lists all molecules that had a direct effect on the growth performance in fish muscle from the good-growth group. Network map indicates the direct relationships of these factors (colored orange due to predicted activation in fish from the good growth group). Genes and metabolites from the input dataset with a relationship to the Upstream Regulators are colored by fold change (red: upregulated in fish muscle from the good-growth group; green; downregulated in fish muscle from the good-growth group). The filtered functions based on the signature of input genes and metabolites were predicted to be inhibited (blue) or activated (orange) in fish muscle from the good-growth group relative to fish muscle from the poor-growth group. Connections are drawn between molecules that have been found to have relationships in the literature. Arrows indicate activation, and perpendicular lines indicate inhibition. Orange and blue lines indicate agreement and yellow line indicates disagreement with previously published literature.

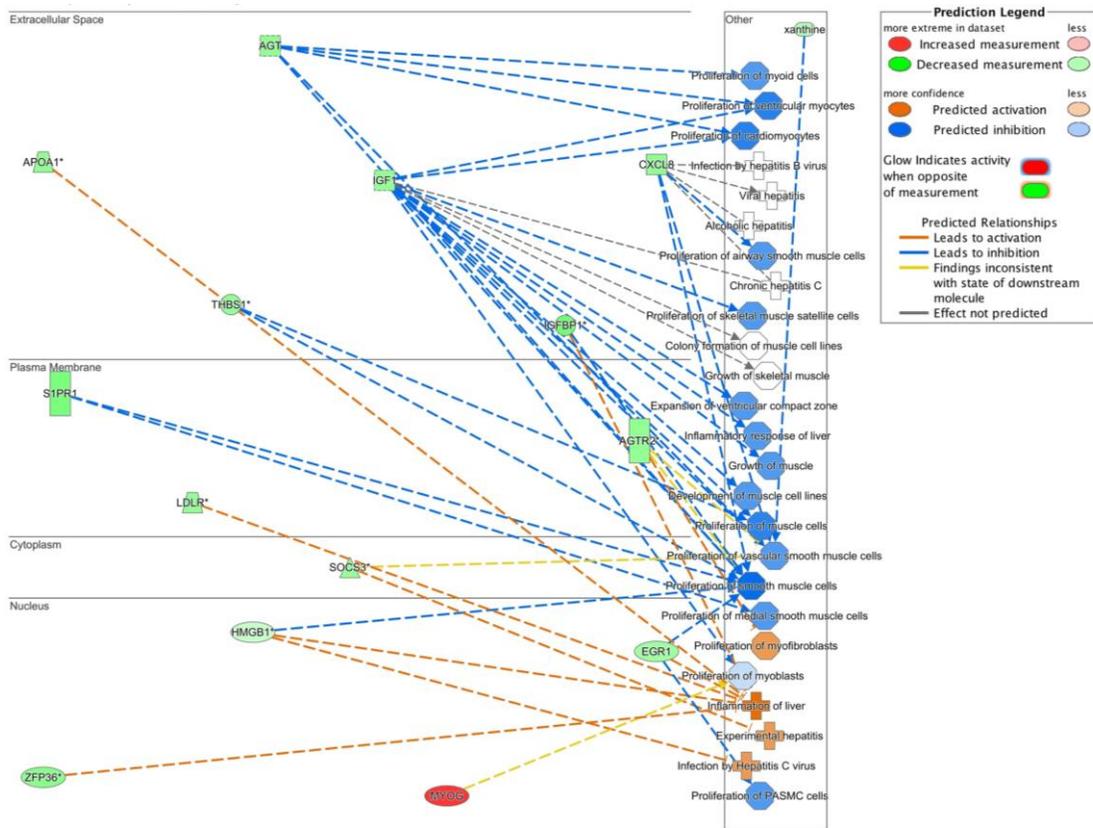


Figure 4.10. BioProfiler analysis of GeneMetaboliteListA (ranked by p -value) of hybrid striped bass. Network was generated from filtered results for muscle and liver interactions. Genes and metabolites from the input dataset are colored by fold change (fish muscle from the good-growth group/fish muscle from the poor-growth group; red: upregulated in fish muscle from the good-growth group; green: downregulated in fish muscle from the good-growth group). The filtered functions based on the signature of input genes were predicted to be inhibited (blue) or activated (orange) in fish muscle from the good-growth group relative to fish muscle from the poor-growth group. Connections are drawn between molecules that have been found to have relationships in the literature. Arrows indicate activation, and perpendicular lines indicate inhibition. Legend is provided indicating what each color represents. Orange and blue lines indicate agreement and yellow line indicates disagreement.

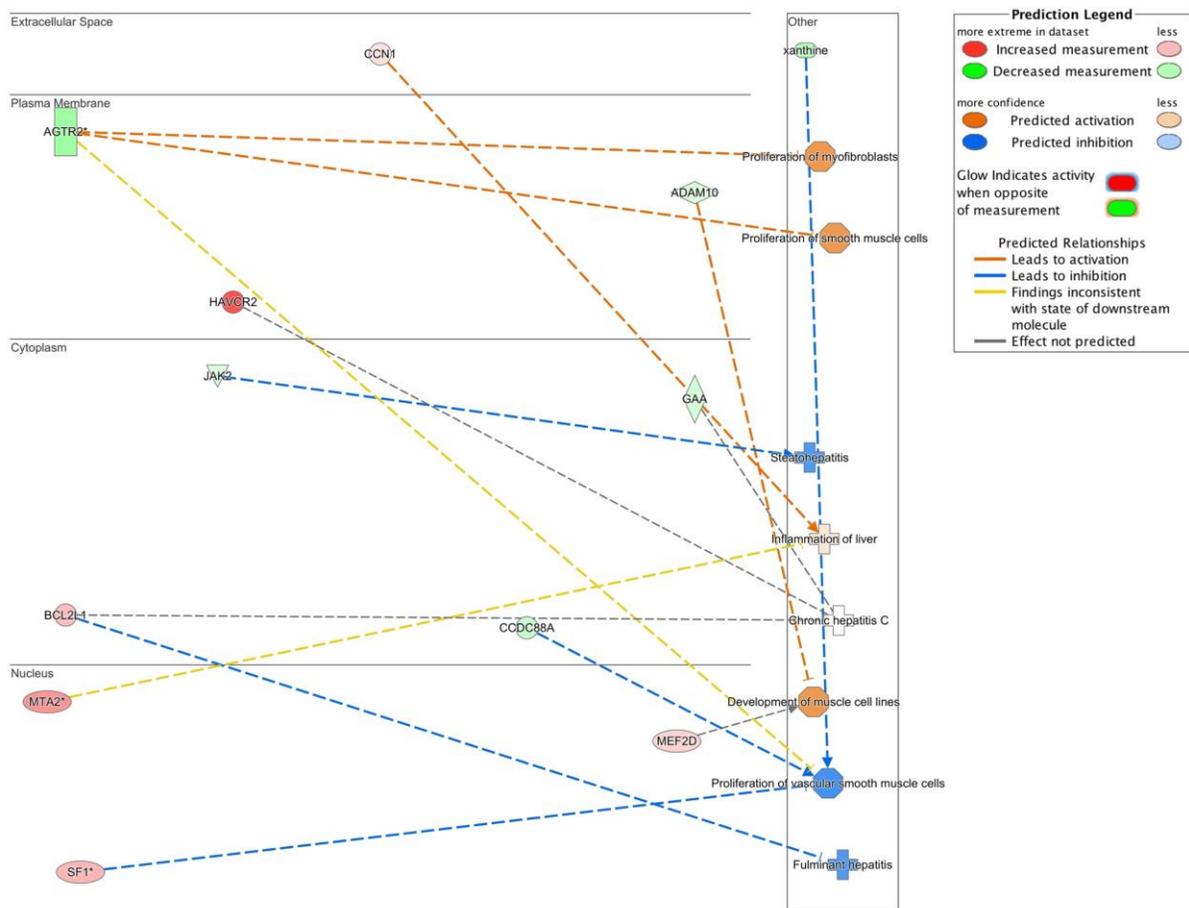


Figure 4.11. BioProfiler analysis of GeneMetaboliteListB (ranked by the Machine Learning-based SVMAttributeEval) from hybrid striped bass. Network was generated from filtered results for muscle and liver interactions. Genes and metabolites from the input dataset are colored by fold change (fish muscle from the good-growth group/fish muscle from the poor-growth group; red: upregulated in fish muscle from the good-growth group; green: downregulated in fish muscle from the good-growth group). The filtered functions based on the signature of input genes were predicted to be inhibited (blue) or activated (orange) in fish muscle from the good-growth group relative to the fish muscle from the poor-growth group. Connections are drawn between molecules that have been found to have relationships in the literature. Arrows indicate activation, and perpendicular lines indicate inhibition. Legend is provided indicating what each color represents. Orange and blue lines indicate agreement and yellow line indicates disagreement.

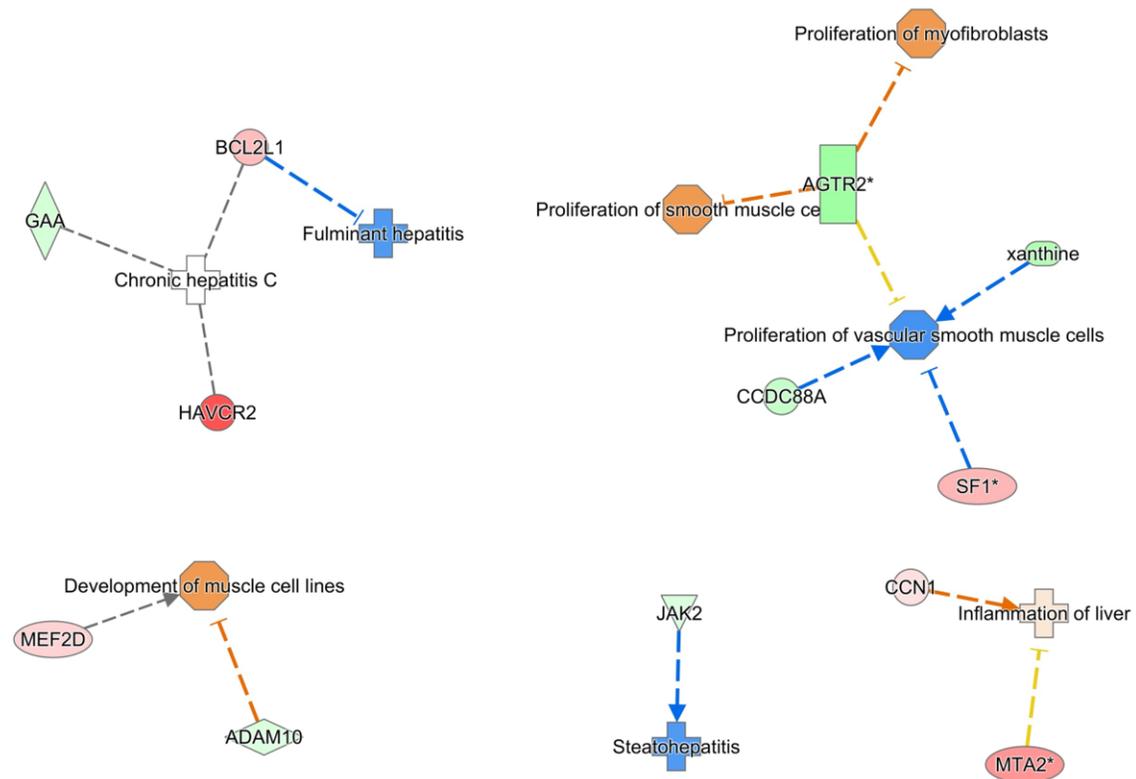


Figure 4.12. Similar to Figure 4.11, but represented as Organic View in IPA. BioProfiler analysis of GeneMetaboliteListB (ranked by the Machine Learning-based SVMAttributeEval) from hybrid striped bass. Network was generated from filtered results for muscle and liver interactions. Genes and metabolites from the input dataset are colored by fold change (fish muscle from the good-growth group/fish muscle from the poor-growth group; red: upregulated in fish muscle from the good-growth group; green: downregulated in fish muscle from the good-growth group). The filtered functions based on the signature of input genes were predicted to be inhibited (blue) or activated (orange) in fish muscle from the good-growth group relative to the fish muscle from the poor-growth group. Connections are drawn between molecules that have been found to have relationships in the literature. Arrows indicate activation, and perpendicular lines indicate inhibition. Legend is provided indicating what each color represents. Orange and blue lines indicate agreement and yellow line indicates disagreement in the published literature.

**Chapter V: The Effects of Frequent and Infrequent Feeding Regimes on the Growth of
Sunshine Hybrid Striped Bass (Male Striped Bass *Morone saxatilis* x Female White
Bass *Morone chrysops*)**

Abstract

Reduced frequency feeding was conducted to improve the feed conversion ratio (FCR) of hybrid striped bass (female *Morone chrysops* x male *M. saxatilis*) while maintaining good growth. Fish (initial weight 416 ± 7.9 g, mean \pm standard deviation) were fed floating pellets (45% protein, 16% fat) for 88 days using automated belt feeding or *ad libitum* hand feeding either three days or five days per week. Experimental groups were replicated three times in flow-through, outdoor tanks (N = 150 fish each tank). There were significant differences in mean total weight fed per fish, initial feeding rate (based on percent body weight), and final feeding rate by 2 x 2 factorial analysis of variance (ANOVA, $p \leq 0.05$). There were no statistical differences in average initial weight, final weight, daily weight gain, or FCR between the treatment groups. Although not significant, FCR was approximately 20% higher in fish fed five days per week compared to those fed three days per week. This increased FCR may have an economic impact in commercial operations and reducing frequency of feeding may lower the overall cost of feeds required to raise the fish, especially as they get larger during Phase III (final growout) production.

Keywords: feed conversion ratio, feeding frequencies, hybrid striped bass, growth rate, sunshine bass

Introduction

Hybrid striped bass refers to either the original cross between a male white bass (*Morone chrysops*) and female striped bass (*Morone saxatilis*), or the reciprocal cross between a male striped bass and a female white bass. The original cross is also known as the “palmetto bass”, which was first produced in South Carolina in the mid-1960s (Hodson and Hayes 1989b). The reciprocal cross, also known as the “sunshine bass”, is produced by most hatcheries, because white bass females are more available and are easy to spawn in captivity (Lochmann 2015). Sunshine bass may grow faster and have higher survival rates than striped bass, white bass, or palmetto bass when reared in indoor water recirculation systems or earthen ponds (Kelly and Kohler 1996; Rudacille and Kohler 2000; McEntire et al. 2015). Also, sunshine bass were found in many studies to grow more quickly and to convert feed more efficiently than striped bass, white bass and palmetto bass during some phases of the production cycle (D’Abramo and Frinsko 2008; Dasgupta and Thompson 2013; Lochmann 2015; Rudacille and Kohler 2000). Despite the performance differences, both of the hybrid crosses exhibit hybrid vigor (heterosis), manifested as improved survival and growth during the first two years, and resistance to handling stress and diseases (Kerby 1993; Kerby and Harrell 1990; Noga et al. 1994). Thus, they gained widespread acceptance either as a sportfish for stock enhancement or as an aquaculture species for food production (Hodson and Hayes 1989b; Kerby 1993; Dasgupta and Thompson 2013).

In recent years, hybrid striped bass production has dramatically grown in different parts of the U.S. (Hodson and Hayes 1989b; Rudacille and Kohler 2000). Farming of hybrid striped bass is currently the fourth largest form of finfish aquaculture in the U.S., behind

only catfish, salmonids and tilapia. In 2012, the total national production of hybrid striped bass was an estimated 8 to 10 million pounds, with a farm-gate value of approximately \$30 to 40 million USD (Lochmann 2015; Reading et al. 2018). The industry has been expanding at a rate of 10 to 15% per year for the past 20 years (Dasgupta and Thompson 2013) and presently, it is estimated to be approximately 12 million pounds with a farm-gate value of about \$50 million USD each year. The expansion potential of hybrid striped bass production, however, is negatively affected by increasing costs associated with aquafeeds. Costs of aquafeeds typically represent the largest single component of aquaculture production. Aquafeeds represent 40 to 50% of the total production costs in aquiculture (Lochmann 2015). Additionally, inconsistencies in final product flavor and texture may be attributed, in part, to feeding diets of differing ingredients. Striped bass Growers Association polls their members annually and feed cost is consistently their number one concern among producers nationwide (Gatlin et al. 2007). There are two generally considered methods to improve aquafeed costs without interfering with the growth rates of fish: 1) formulate feeds which are comprised of ingredients that cost less to produce, but provide an equivalent nutritional value and 2) optimize feeding and husbandry practices to maximize the feed conversion ratio (FCR) of the animals. FCR is the kg (or pounds, lbs) of feed required to increase the biomass in the fish by one kg (or lb, respectively) and it is a primary economic concern of fish farmers (D'Abramo and Frinsko 2008).

Several studies have focused on establishing the nutrient requirements (D'Abramo and Frinsko 2008; Rawles et al. 2012) and improving diet formulations to improve the growth rate and, hence, economics of hybrid striped bass production (Gatlin et al. 2007). These studies aim to optimize and evaluate less expensive feed formulations that provide

necessary energy and essential nutrients for growth. On the other hand, few studies have tested the response of hybrid striped bass growth performance to changes in feeding frequency or husbandry practices (Liu and Liao 1999; Webster et al. 2001; Rawles *et al.* 2012). For instance, Rawles et al. (2012) recommended the best way to reduce the feed cost without affecting the growth performance of the hybrid striped bass seems to be reducing feeding when water temperatures are not within the optimum growth range.

Henceforth, the objective of the present study was to test whether reduced frequency feeding may improve costs of production by optimizing husbandry practices and maximizing the FCR of hybrid striped bass. By increasing the growth rate through a series of trials using different feeding strategies (three day and five day per week), we attempted to improve the FCR of hybrid striped bass, while maintaining maximum growth.

Materials & Methods

Domestic hybrid striped bass (sunshine bass) were produced at North Carolina Fisheries (Aurora, NC, USA) by crossing the domestic striped bass male with the wild female white bass captured from the Tennessee River (TN, USA). The domestic fish originated from the *National Program for Genetic Improvement and Selective Breeding for the Hybrid Striped Bass Industry* at North Carolina State University Pamlico Aquaculture Field Laboratory (NCSU PAFL) in Aurora, NC, USA (Reading et al. 2018). The fish (initial weight 416 ± 7.9 g, values given as average and \pm standard deviation) were randomly sorted into 12 tanks at 120-178 fish per tank (three meters diameter; 5,800 liters) in an outdoor flow-through aquaculture system at the NCSU PAFL. To ensure unbiased initial average weights or behavioral feeding effects, fish were stocked by sorting them into tanks

15 fish at a time to ensure near equal initial average weights. Each tank received 17 to 20 liters per minute (LPM) of freshwater sourced from a well (Castle Hayne Aquifer). From stocking (March 21) until March 28, the tanks received a mixture of well water (freshwater) and South Creek water (brackish) to increase salinity for stress relief following handling of the fish (1.9 to 7.8 ppt) (Table 5.1). The fish were otherwise reared in freshwater during the experiment. Water flow rates and aeration were maintained continuously and dissolved oxygen in the water of the tanks stocked at the rates in this study did not fall below 80% saturation. Water and air-flow were not shut off at all during the entirety of the study.

Water temperature was recorded periodically throughout the study. The average water temperature was approximately 16°C in mid-March and increased to approximately 23°C by mid-June at the conclusion of the experiment. For 88 days the fish were fed using two different methods: either by hand additionally automated belt feeder. In addition, fish were fed at two different frequencies: three compared to five days per week. Thus, there was a total of four experimental groups with three replicate tanks each (Hand x 3, Hand x 5, Belt x 3, and Belt x 5). All fish were fed with a prepared floating pellet diet (1/4", 45% protein / 16% lipid) (Zeigler Brothers, Inc., Gardners, PA, USA states). The amount fed to the fish was based on observations of feeding response: if all feed was consumed rapidly, then the amount fed at the next feeding was increased. If not all feed was consumed, then the amount fed at the next feeding was decreased. The amount of feed provided to each tank was weighed and recorded daily. Also, each day, approximately 15% of the water volume was drained from the bottom of each tank by removing the standpipe. The color and turbidity of the water was observed and noted. At the end of the feeding trial (88 days)

fish were weighed and the average weight gain per day, total weight fed, and feed conversion ratio (FCR) were calculated for each of the 4 groups. The following formula was used to calculate FCR:

$$\text{FCR} = \frac{\text{Average Feed intake (g)}}{\text{Average weight gain per day (g)}}$$

Feeding rates (FR) also were evaluated as the amount of feed consumed and expressed as a percent of total body weight. The following formula was used to calculate FR at the beginning (Initial) and at the conclusion (Final) of the feeding trials:

$$\text{FR} = \frac{\text{Average feed intake (g)}}{\text{Average body weight (g)}} \times 100$$

Statistical differences in initial weight, final weight, weight gain per day, total weight fed, FCR, Initial FR, and Final FR between the treatment groups were evaluated with a 2 x 2 factorial analysis of variance (ANOVA, alpha = 0.05) using Statistical Analysis System – JMP (SAS JMP®, 14.0.0; SAS Institute Inc., Cary, NC). If the ANOVA indicated significance ($p \leq 0.05$), then treatment means and their interactions were subjected to a LSMeans Difference Tukey Honestly Significant Difference (HSD) *post hoc* test to observe where the differences occurred. In the cases in which two means (pairwise) were being compared, LSMeans Student's t-tests were performed at an alpha level of 0.05. The slopes (m) of lines representing changes in average Initial and Final FRs and the average Initial and Final weights between groups were calculated using the linear equation:

$$y = mx + b$$

Results

Unbeknownst at time of stocking, 30 fish were mistakenly stocked in replicate three of the Belt x 3 treatment instead of replicate two of the Hand x 3 treatment. Since the number of fish were unequal in the replicate tanks, all statistical comparisons were based on a per fish basis and not a per tank basis. There were no fish mortalities during the growth trials; however, it was noted that a few fish jumped out of the rearing tanks during the duration of the growth trials. The water temperature and salinity data are shown in Table 5.1. Water temperatures that were recorded in the experiment ranged from 16.0°C to 22.5°C, which is less than the optimum temperature range (25 to 27°C) recorded for the best hybrid striped bass growth rate (Hodson and Hayes 1989b), however the fish still grew well (average weight gain was about two g per day for all treatment groups).

The summary feeding data are shown in Table 5.2 and the summary statistics are shown in Table 5.3. The initial average weight of fish ranged from 410 to 422 g in the four feeding groups. There were no significant differences between feeding groups in initial weight of the sunshine bass fed either three or five days per week either by hand or by automated belt feeders (Tables 5.2, 5.3; Figure 5.1), indicating that the fish were of equivalent weight in the four treatment groups at the beginning of the growth trials. The final average weight of sunshine bass ranged from 592 to 609 g. There was no significant difference in final average weights of sunshine bass fed either three or five days per week either by hand or by automated belt feeders (Tables 5.2, 5.3; Figure 5.1). The average weight gain per day, average weight fed per fish, and FCR ranged from 1.94 to 2.22 g, 315.97 to 437.10 g, and 1.71 to 2.49, respectively. There were no significant differences between feeding groups in regards to average weight gain per day or average FCR of

sunshine bass in any of the feeding trials (Tables 5.2, 5.3; Figures 5.2A, 5.2C). There was a significant difference in average total weight fed per fish and the difference lies in the interaction between the fish fed by hand three days per week and the fish fed by automated belt feeder five days per week ($p = 0.004$) (Tables 5.2, 5.3; Figure 5.2). The fish fed five days per week by automated belt feeder consumed more feed than those that were fed three days per week by hand (g/fish).

The average Initial and Final FRs of hybrid striped bass are shown in Table 5.2. The Initial FR was significantly different between fish fed by hand and by automated belt feeder, between fish fed three or five days per week, and between the interaction of feeding method (by automated belt feeder or by hand) and feeding frequency (three or five days per week) (Table 5.3; Figure 5.3). Fish that were fed by automated belt feeder three days per week had a significantly higher average Initial FR compared to the other groups. The Belt x 3 group had a significantly higher Initial FR compared to the other groups due to range finding of the feeding rates at the beginning of the trials. There was no significant difference in the Final FR between hand fed and automated belt feeding groups (Table 5.3; Figure 5.3). There was, however, a significant difference between the feeding frequencies (Table 5.3; Figure 5.3), such that the fish fed three days per week had higher Final FR compared to fish fed five days per week. There was no significant difference between the interaction of feeding frequencies and feeding methods in regards to the Final FR (Table 5.3). There were no significant differences between the feeding groups in Initial or Final body weight (ANOVA, $\alpha = 0.05$) (Tables 5.2, 5.3; Figure 5.4).

Blackish colored water was observed from the tanks containing fish that were fed three days per week either by hand or by automated belt feeder and tan colored water

containing feces and uneaten feed was observed from the tanks containing fish that were fed five days per week after approximately 15% of the water was drained from the tanks by removing the standpipe.

Discussion & Conclusion

The results overall demonstrate that reducing the frequency of feeding in sunshine hybrid striped bass may improve production costs as the fish appear to grow similarly when fed either three or five days per week. Feed consumption increases, although not significantly, with increased frequency of feeding (Table 5.2; Figure 5.2). This increased feed consumption did not result in any benefit to growth rate as there were no differences in final body weight between groups fed at three or five days per week (Figure 5.4) or average weight gain per day (Figure 5.2). Previous studies that were conducted on other fish species (Jobling 1983) have shown that feed consumption and growth rate either measured via specific growth rate (SGR) or total body weight generally increases with feeding frequency up to a some limit (Liu and Liao 1999; Dwyer et al. 2002; El Sayed Ali et al. 2016). For instance, starvation may lead to some physical adaptation, such as hyperphagic response, but slight reduction in the frequency for feeding may not (Jobling 1982). Liu and Liao (1999), conducted feeding trials of hybrid striped bass, where they similarly show that feeding frequency directly affected the feed intake, which in return affected the total metabolic efficiency and growth rate. Their results showed that fish fed at high frequencies consumed larger quantities of feed overall compared to those fed less often, but with decreased feed conversion efficiency. This is consistent with the data presented here for domestic sunshine hybrid striped bass.

Feeding frequencies have a direct effect on the water quality (Thongprajukaew et al. 2017). According to previous studies, the quality of water is associated with the faster gastric and/or intestinal emptying rate when fish were predominantly subjected to high feeding frequencies (Andrews and Page 1975; Wang et al. 1998; Liu and Liao 1999; Dwyer et al. 2002; Tekinary et al. 2003; Li et al. 2014). As such, the food passed through the digestive track more quickly, resulting in less effective digestion and assimilation, and inevitably contributed to poor water quality (Andrews and Page 1975; Wang et al. 1998; Liu and Liao 1999; Dwyer et al. 2002; Li et al. 2014). Lim and Webster (2001), suspected that hybrid striped bass like to gorge and once their optimum feed rate drops somewhere below three or four percent body weight, then it is better to feed less often. Hybrid striped bass can and will eat up to seven percent body weight given the chance and they will empty their digestive track of the partially digested feed so as to be able to eat fresh feed. The consequences are obvious, but there is a hesitancy by growers to feed less often, even though this has been shown to be a viable approach to maintain improved water quality. This was also demonstrated here for sunshine hybrid striped bass by rapidly draining about 15% of the tank volume each weekday. The color of the discharged waste water was a good indicator of the digestion rate of the fish as blackish colored water that was observed from the tanks containing three day per week feeding groups indicated that the feed was fully digested, whereas tan colored water observed from the five day per week feeding groups indicated overfeeding with incomplete digestion and/or passage of undigested feed.

In this study, the FCR was quite variable between groups, ranging from 1.71 to 2.49 (Figure 5.2), with a substantially better FCR for fish fed by hand three days per week (Table

5.2). Even though this was not statistically significant, the average FCR of the fish that were fed only three days per week was 1.94 compared to those that were fed five days per week (2.38). Thus, the average FCR is approximately 20% higher in fish that are fed at five compared to three days per week and reducing the frequency of feeding may improve the FCR. Additionally, the average weight gain per day was not greater in the fish fed five versus three days per week. There was no significant difference in growth evaluated as a change in total body weight from the beginning of the experiment to the end in regards to feeding the fish at a frequency of three versus five days per week (Figures 5.1A, 5.1B). Therefore, it appears that the fish will grow equivocally if they are fed at a slightly lower frequency. A similar result was obtained for sea bass *Dicentrarchus labrax*, where fish fed less frequently showed better feed conversion efficiency (FCE) (Tsevis et al. 1992). According to Chua and Teng (1978), an improved FCE (or FCR), could be attributed to less physiological activity and lower endocrine function of fish. This was in agreement with Thongprajukaew et al. (2017), who investigated of the effects of different feeding frequencies on the digestive enzyme activity in Nile tilapia *Oreochromis niloticus*. Tilapia fed at longer time intervals (i.e., less frequently), retained greater activity of the intestinal total protease, amylase, and lipase activities, than those fed at shorter time intervals (i.e., more frequently). Accordingly, they concluded that the longer time intervals between the meals will lead to higher feed utilization in the long-term for enzymatic digestion. This may relate to improved FCR for fish fed at reduced frequency feeding intervals. The optimum feeding frequency may vary with species, nevertheless, finding the adequate feeding regimens for hybrid striped bass may improve digestibility, and consequently accelerate protein assimilation, which may result in better FCR. Further research will be required to

confirm these findings at larger scale; however, improved FCR of sunshine bass may reduce the overall cost of feeds required to raise the fish.

Since fishes are poikilotherms, the FCR is dependent on the water temperature as well as water quality, and food uptake and assimilation (Turano et al. 2007). Fish growth declines when temperature is not optimal, either being too cold or too warm. Therefore, growth in ponds or flow-through systems may decrease or cease during winter or summer months for many species (Keembiyehetty and Wilson 1998). Feeding of hybrid striped bass is commonly suspended when water temperature is below 15.6°C, which will affect the water quality and hence food intake efficiency (D'Abramo and Frinsko 2008). On the other hand, when water temperatures exceed 30°C, a large number of hybrid striped bass will sometimes be observed floating upside down at the pond surface. This is presumably caused by an inability to digest feed efficiently at high water temperatures (Chipps et al. 2000; Rawles et al. 2012). Previous research has shown hybrid striped bass growth performance at various temperatures and interestingly there is an increase in FCR in response to a parallel increase in temperature (D'Abramo and Frinsko 2008; Davis 2004; Keembiyehetty and Wilson 1998) indicating that the overall growth efficiency declines at warmer temperatures. At higher temperatures, hybrid striped bass overconsume feed, but do not efficiently utilize the nutrients for growth (Rawles et al. 2012). Thus, at temperatures below or beyond the optimum growth range (i.e., 25 to 27°C, Woiwode and Adelman 1983; Hodson and Hayes 1989b) a producer should either: 1) Change the feed formula to one that corresponds most appropriately to the change in temperature conditions (D'Abramo and Frinsko 2008) or 2) by employing a reduced frequency feeding strategy as suggested here to maximize the FCR of hybrid striped bass. Rawles et al. (2012),

similarly recommended the best way to reduce the feed cost without affecting the growth performance of the hybrid striped bass seems to be in reducing feeding when water temperatures are not within the optimum range. Summarily, these adaptations are of importance for husbandry techniques, for operational efficiency, and quality control of water, all of which can be improved by decreasing the frequency of feedings when the temperature is high. In applying this method, hybrid striped bass will maintain high FCR, hence maximize growth and minimize waste and improve water quality, whilst reducing labor costs, much like in the European seabass *Dicentrarchus labrax* (Peres and Oliva-Teles 1999).

In conclusion, different feeding frequencies had no statistically significant effects on growth performance of hybrid striped bass, however increasing feeding frequency was associated with an increase in individual food intake, a decrease in the FCR, and a decrease in water quality associated with presence of undigested feed in the water. In general, sunshine hybrid striped bass consumed more feed as they were provided with it more frequently (five days per week), yet the final growth rate represented as average body weight in grams showed no significant difference between that of fish fed less frequently (three days per week). The FCR was approximately 20% higher in fish that were fed more frequently (five days per week) compared to those that were fed less frequently (three days per week) indicating that this practice is more efficient, perhaps especially during times when the fish are not at optimal temperatures. The current results suggests that, reducing the frequency of feeding may improve the FCR of sunshine bass and lower the overall cost of feeds required to raise the fish, especially as they get larger during Phase III (final growout) production and when water temperatures are not at its optimum range.

References

Ali, T. E. S., Martínez-Llorens, S., Moñino, A. V., Cerdá, M. J. and Tomás-Vidal, A. (2016).

Effects of weekly feeding frequency and previous ration restriction on the compensatory growth and body composition of Nile tilapia fingerlings. *Egyptian Journal of Aquatic Research*, 42(3), 357–363.

Andrews, J. W. and Page, J. W. (1975). The effects of frequency of feeding on culture of

catfish. *Transactions of the American Fisheries Society*, 1975(2), 317–321.

Chipps, S. R., Clapp, D. F. and Wahl, D. H. (2000). Variation in routine metabolism of

juvenile muskellunge: Evidence for seasonal metabolic compensation in fishes.

Journal of Fish Biology, 56(2), 311–318.

Chua, T.-E. and Teng, S.-K. (1978). Effects of feeding frequency on the growth of young

estuary grouper, *Epinephelus tauvina* (Forsk.), cultured in floating net-cages.

Aquaculture, 14, 31–47.

D'Abramo, L.R. and M.O. Frinsko. (2008). Hybrid Striped Bass: Pond Production of Food

Fish. Southern Regional Aquaculture Center. Stoneville, Mississippi. Publication

Number 303.

Dasgupta, Siddhartha and Kenneth R. Thompson. (2013). Comparison of coast of

different Hybrid Striped Bass production systems in ponds. *Southern Regional*

Aquaculture Center, Publication Number 3000.

Davis, K. B. (2004). Temperature affects physiological stress responses to acute

confinement in sunshine bass (*Morone chrysops* × *Morone saxatilis*). *Comparative*

Biochemistry and Physiology - A Molecular and Integrative Physiology, 139(4), 433–

440.

- Dwyer, K. S., Brown, J. A., Parrish, C. and Lall, S. P.** (2002). Feeding frequency affects food consumption, feeding pattern and growth of juvenile yellowtail flounder (*Limanda ferruginea*). *Aquaculture*, 213(1-4), 279-292.
- Gatlin, Delbert M., Frederic T. Barrows, Paul Brown, Konrad Dabrowski, T. Gibson Gaylord, Ronald W. Hardy, Eliot Herman, Gongshe Hu, Åshild Krogdahl, Richard Nelson, Kenneth Overturf, Michael Rust, Wendy Sealey, Denise Skonberg, Edward J Souza, David Stone, Rich Wilson and Eve Wurtele.** (2007). *Expanding the utilization of sustainable plant products in aquafeeds: a review. Aquaculture research*, 38(6), 551-579.
- Helena Peres, A. O.-T.** (1999). Influence of temperature on protein utilization in juvenile European seabass (*Dicentrarchus labrax*). *Aquaculture*, 170, 337-348.
- Hodson, R. G. and M. Hayes.** (1989b). Hybrid Striped Bass pond production of foodfish. *Southern Regional Aquaculture Center, Publication Number 303.*
- Jobling, M.** (1982). Some observations on the effects of feeding frequency on the food intake and growth of plaice, *Pleuronectes platessa* L. *Journal of Fish Biology*, 20(4), 431-444.
- Jobling, M.** (1983). Effect of feeding frequency on food intake and growth of Arctic charr, *Salvelinus alpinus* L. *Journal of Fish Biology*, 23(2), 177-185.
- Keembiyehetty, C. N. and Wilson, R. P.** (1998). Effect of water temperature on growth and nutrient utilization of sunshine bass (*Morone chrysops* ♀ x *Morone saxatilis* ♂) fed diets containing different energy/protein ratios. *Aquaculture*, 166(1-2), 151-162.
- Kelly, Anita M. and Christopher C. Kohler.** (1996). Sunshine bass performance in ponds, cages, and indoor tanks. *The Progressive fish-culturist*, 58(1), 55-58.

- Kerby, J.H. and Harrell RM.** (1990). Hybridization, genetic manipulation, and gene pool conservation of striped bass. In: Harrell RM, Kerby JH, Minton RV eds. Culture and propagation of striped bass and its hybrids. *Bethesda, MD: American Fisheries Society*, 159–190.
- Kerby, J.H.** (1993). The hybrid striped bass and its hybrids. Culture of nonsalmonid freshwater fishes (pp. 252–255) in Robert R. Stickney editor. Boca Raton, FL: CRC Press, Inc.
- Lim, C. and Webster, C. D.** (2001). Nutrition and fish health. *New York: Food Products Press*.
- Liu, F. G. and Liao, I. C.** (1999). Effect of Feeding Regimen on the Food Consumption, Growth, and Body Composition in Hybrid Striped Bass *Morone saxatilis* x *M. chrysops*. *Fisheries science*, 65(4), 513-519.
- Li, X. F., Tian, H. Y., Zhang, D. D., Jiang, G. Z. and Liu, W. Bin.** (2014). Feeding frequency affects stress, innate immunity and disease resistance of juvenile blunt snout bream *Megalobrama amblycephala*. *Fish and Shellfish Immunology*, 38(1), 80–87.
- Lochmann, Rebecca.** (2015). Feeds and feeding of hybrid striped bass. *Southern Regional Aquaculture Center*, Publication Number 3001.
- McEntire, M., Snyder, S. and Freeman, D.** (2015). Comparison of growth between *Morone* hybrids (palmetto and sunshine) in earthen ponds. *Journal of the World Aquaculture Society*, 46(5), 557-563.
- Noga, E. J., Kerby, J. H., King, W., Aucoin, D. P. and Giesbrecht, F.** (1994). Quantitative comparison of the stress response of striped bass (*Morone saxatilis*) and hybrid

- striped bass (*Morone saxatilis* x *Morone chrysops* and *Morone saxatilis* x *Morone americana*). *American journal of veterinary research*, 55(3), 405-409.
- Peres, Helena and Oliva-Teles.** (1999). Influence of temperature on protein utilization in juvenile European seabass (*Dicentrarchus labrax*). *Aquaculture*, 170, 337–348.
- Rawles, Steven D., T. Gibson Gaylord, G. Scott Snyder and Donald W. Freeman.** (2012). The influence of protein and energy density in commercial diets on growth, body composition, and nutrient retention of sunshine bass, *Morone chrysops* ♀ x *Morone saxatilis* ♂, reared at extreme temperature. *Journal of the World Aquaculture Society*, 41(s2), 165-178.
- Reading, B.J., McGinty, A.S., Clark, R.W., Hopper, M.S., Woods III, L.C. and Baltzegar, D.A.** (2018). Genomic Enablement of Temperate Bass Aquaculture (Family Moronidae), In *Breeding and Culture of Perch and Bass* (Wang, H., and Liang, X., Eds.). Science China Press (Chinese Academy of Sciences).
- Rudacille, J. B. and Kohler, C. C.** (2000). Aquaculture performance comparison of sunshine bass, palmetto bass, and white bass. *North American Journal of Aquaculture*, 62(2), 114-124.
- Thongprajukaew, K., Kovitvadhi, S., Kovitvadhi, U. and Preprame, P.** (2017). Effects of feeding frequency on growth performance and digestive enzyme activity of sex-reversed Nile tilapia, *Oreochromis niloticus* (Linnaeus, 1758). *Agriculture and Natural Resources*, 51(4), 292–298.
- Tsevis, N., Klaoudatos, S. and Conides, A.** (1992). Food conversion budget in sea bass, *Dicentrarchus labrax*, fingerlings under two different feeding frequency patterns. *Aquaculture*, 101(3–4), 293–304.

- Turano, Marc J., Russell J. Borski and Harry V. Daniels.** (2007). Compensatory growth of pond-reared Hybrid Striped Bass, *Morone chrysops* × *Morone saxatilis*, Fingerlings. *Journal of the World Aquaculture Society*, 38(2), 250-261.
- Wang, N., Hayward, R. S. and Noltie, D. B.** (1998). Effect of feeding frequency on food consumption, growth, size variation, and feeding pattern of age-0 hybrid sunfish. *Aquaculture*, 165(3-4), 261-267.
- Webster, C.D. and Lin, C.E.** (2000). Nutrient Requirements and Feeding of Finfish for Aquaculture. CABI Publishing, New York.
- Webster, Carl D., Kenneth R. Thompson, Ann M. Morgan, Ebony J. Grisby and Siddhartha Dasgupta.** (2001). Feeding frequency affects growth, not fillet composition, of juvenile sunshine bass *Morone chrysops* × *M. saxatilis* grown in cages. *Journal of the world aquaculture society*, 32(1), 79-88.
- Woiwode, J. G. and Adelman, I. R.** (1983). Growth, food conversion efficiency and survival of the hybrid white × striped bass as a function of temperature. The aquaculture of striped bass: proceedings. Maryland Sea Grant Program, Publication UM-SG-MAP-84-01, University of Maryland, College Park, 143-150.

Table 5.1. Water temperature and salinity data during the hybrid striped bass feeding trials.

Date	Feeding Time	Temperature	Salinity
March 24	14:00	18.9	7.5
March 21	12:45	16.6	7.8
March 28	13:00	16.0	1.9
March 29	13:00	18.2	0.0
March 31	14:00	17.5	0.0
April 1	14:00	17.9	0.0
April 5	13:45	18.6	0.0
April 25	12:15	20.9	0.0
June 9	13:15	22.7	0.0
June 10	14:30	22.7	0.0

Table 5.2. Summary data of the sunshine hybrid striped bass growth trials including number of fish per replicate tank, initial weight, initial feeding rate (FR % body weight), final weight, final FR, weight fed, and feed conversion ratio (FCR) for fish either were fed by hand or automated belt feeder at either three or five days per week (Hand x 3, Belt x 3, Hand x 5, and Belt x 5).

Variable	Feeding Method and Frequency			
	Hand x 3	Belt x 3	Hand x 5	Belt x 5
Fish per tank	139 ± 17	159 ± 17	150 ± 0	150 ± 2
Initial Weight (g/fish)	414 ± 5	422 ± 10	416 ± 10	410 ± 4
Initial FR (% Body Weight)*	1.28 ± 0.19 ^b	1.96 ± 0.16 ^a	1.01 ± 0.11 ^b	1.30 ± 0.21 ^b
Final Weight (g/fish)	609 ± 78	593 ± 37	595 ± 8	592 ± 35
Final FR (% Body Weight)*	2.16 ± 0.32 ^a	1.98 ± 0.10 ^{ab}	1.46 ± 0.22 ^b	1.54 ± 0.53 ^b
Daily Weight Gain (g/fish)	2.22 ± 0.85	1.94 ± 0.33	2.01 ± 0.17	2.04 ± 0.42
Total Weight Fed (g/fish)*	316 ± 50 ^a	370 ± 47 ^{ab}	406 ± 13 ^{ab}	437 ± 25 ^b
Feed Conversion Rate (FCR)	1.71 ± 0.35	2.16 ± 0.34	2.28 ± 0.21	2.49 ± 0.60

Notes.

Values are means ± *SD* (average N = 150 fish per tank);

Asterisks indicate variables with statistically significant differences between feeding trials; mean values in a column with different superscript letters are significantly different (ANOVA, LSMeans Student's t-test, or LSMeans Difference Tukey HSD, $p \leq 0.05$).

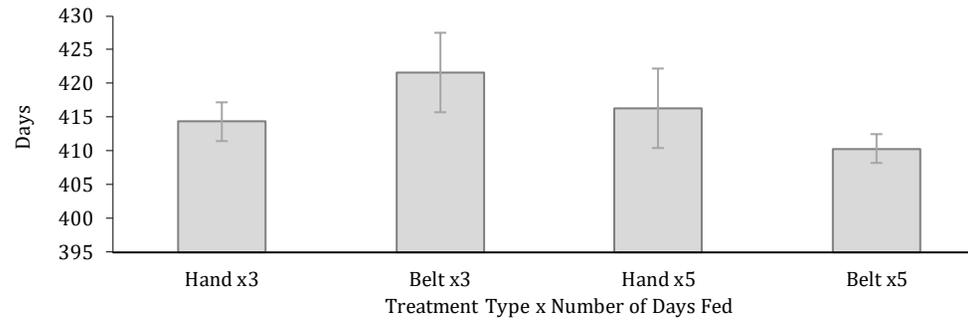
Table 5.3. Summary of the *p*-values for the sunshine hybrid striped bass growth trials including initial weight, initial feeding rate (FR % body weight), final weight, final FR, weight fed, and feed conversion ratio (FCR) for fish either were fed by hand or automated belt feeder at either three or five days per week (Hand x 3, Belt x 3, Hand x 5, and Belt x 5)

Variable	Feeding Method	Feeding Frequency	Interaction	Feeding Method X Feeding Frequency			
				Hand x 3	Belt x 3	Hand x 5	Belt x 5
Initial Weight (g/Fish)	0.8867	0.3331	0.1794	A	a	a	a
Initial FR (% Body Weight)	0.0003*	0.0004*	0.0397*	B	a	b	b
Final Weight (g/Fish)	0.7149	0.7872	0.8149	A	a	a	a
Final FR (% Body Weight)	0.6481	0.0012*	0.2893	A	a,b	b	b
Daily Weight Gain (g/Fish)	0.6853	0.8517	0.6230	A	a	a	a
Total Weight Fed (g/Fish)	0.0852	0.0040*	0.6740	B	a,b	a,b	a
Feed Conversion Rate (FCR)	0.1907	0.0890	0.6077	A	a	a	a

Notes.

Asterisks indicate *p*-values that are significantly different in columns Feeding Method and Feeding Frequency (ANOVA, LSMeans Student's *t*, $p \leq 0.05$) and Interaction (between those two variables) (LSMeans Difference Tukey HSD, $p \leq 0.05$); The letters (a,b) denote significant differences between variables such that those with the same letter superscript indicate mean values that are not significantly different.

(a) Initial weights of sunshine bass



(b) Final weights of sunshine bass

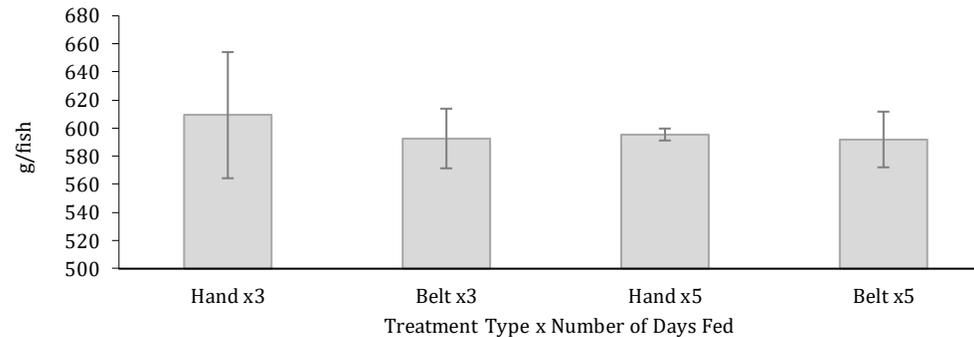
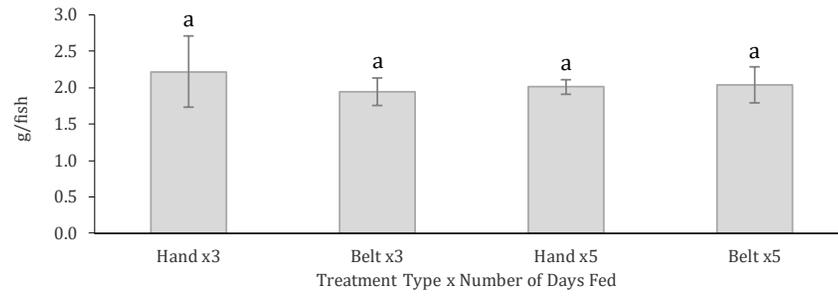


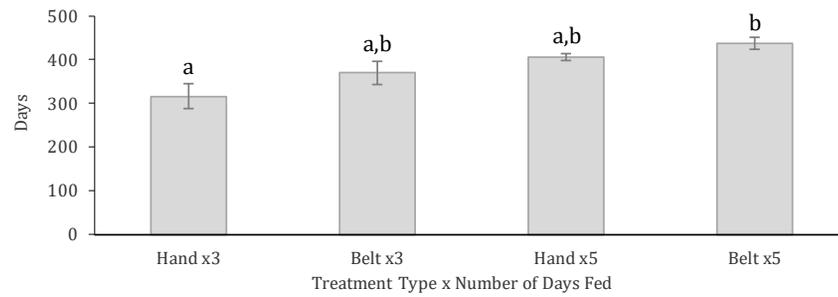
Figure 5.1. Bar charts showing the average (a) initial and (b) final weights (g/fish) of sunshine bass in the feeding trials. The fish were fed either by hand or automated belt feeder three or five days per week (Hand x 3, Hand x 5, Belt x 3, and Belt x 5). Average values are represented by each bar and the *SE* of the mean is shown by the brackets. There were three tank replicates per experimental group. There were no significant differences between groups (ANOVA, LSMeans Student's t-test or LSMeans Difference Tukey HSD, alpha = 0.05).

Figure 5.2. Bar charts showing the average (a) weight gain per day (g/fish), (b) total weight fed per fish, and (c) feed conversion ratio (FCR) of sunshine bass hybrid striped bass in the feeding trials. The fish were fed either by hand or automated belt feeder three or five days per week (Hand x 3, Hand x 5, Belt x 3, and Belt x 5). Average values are represented by each bar and the *SE* of the mean is shown by the brackets. There were three tank replicates per experimental group. Data points with the same letter superscript indicate mean values that are not significantly different at alpha = 0.05 (ANOVA, LSMeans Student's t-test or LSMeans Difference Tukey HSD). There were no significant differences between groups in average weight gain per day (g/fish) or average feed conversion ratio (FCR) (ANOVA, LSMeans Student's t-test, alpha = 0.05). There was a significant difference between Hand x 3 and Belt x 5 in average total weight fed per fish (LSMeans Difference Tukey HSD, $p \leq 0.05$).

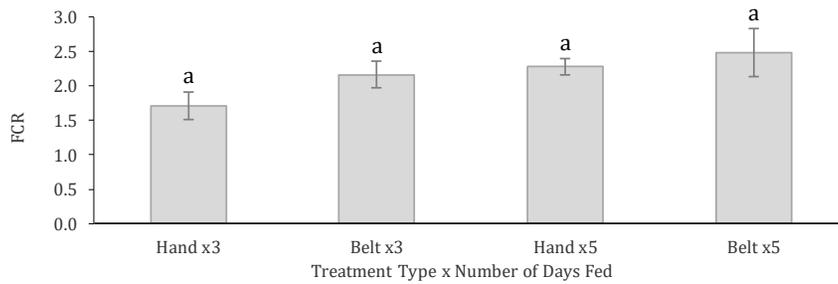
(a) Weight Gain Per Day



(b) Total Weight Fed Per Fish



(c) Feed Conversion Ratio (FCR)



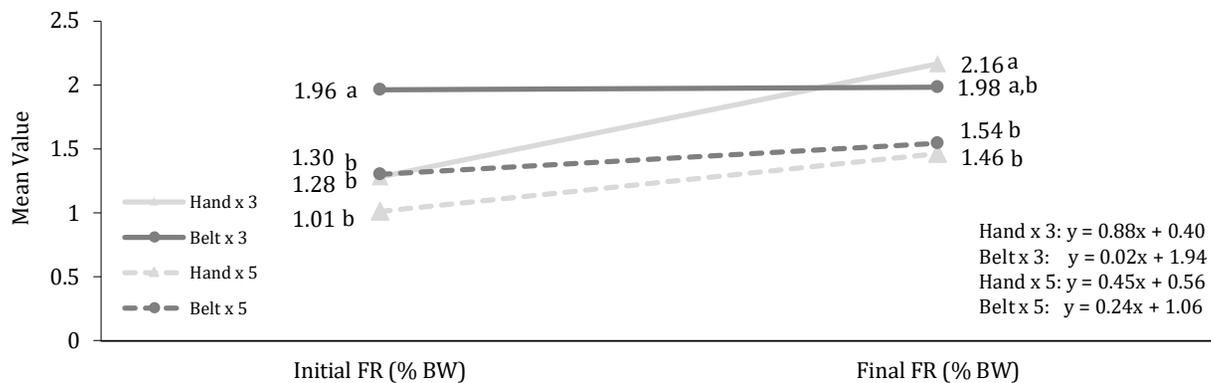


Figure 5.3. Line graph showing the average Initial feeding rates (Initial FR) and Final feeding rates (Final FR) of sunshine hybrid striped bass expressed as percent body weight (%BW) for all treatment groups. The fish were fed either by hand or automated belt feeder three or five days per week (Hand x 3, Belt x 3, Hand x 5, and Belt x 5). The values represent the mean of three replicate tanks. Data points with the same letter superscript indicate mean values that are not significantly different (ANOVA, LSMeans Student's t-test or LSMeans Difference Tukey HSD, $p \leq 0.05$). The slopes of the lines representing change in feeding rate also are provided.

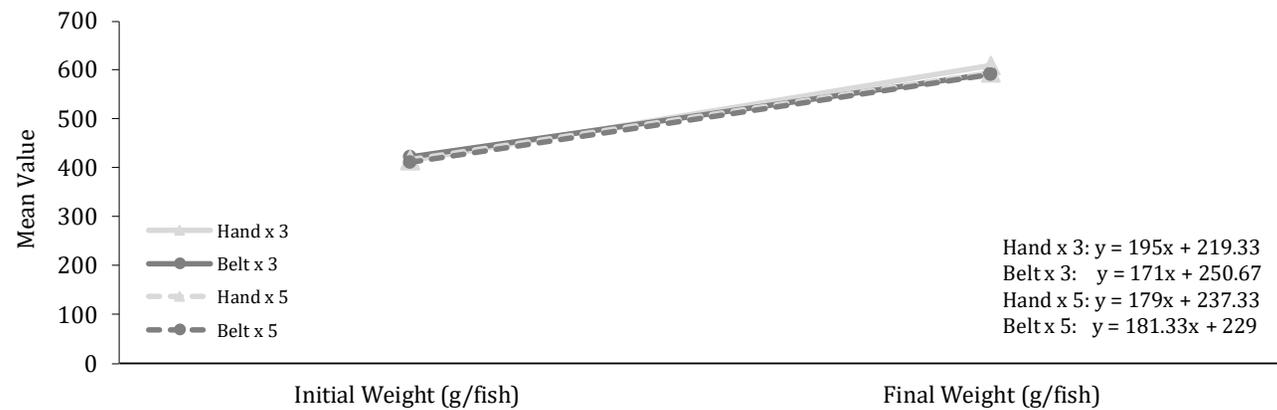


Figure 5.4. Line graph showing the average Initial and Final body weights (g/fish) of sunshine hybrid striped bass for all treatment groups. The values represent the mean of three replicate tanks. The fish were fed either by hand or automated belt feeder three or five days per week (Hand x 3, Belt x 3, Hand x 5, and Belt x 5). There were no significant differences between groups in Initial or Final body weight (ANOVA, LSMeans Student's t-test or LSMeans Difference Tukey HSD, alpha = 0.05). The slopes of the lines representing changes in body weight (i.e., growth) also are provided.

Chapter VI: Summary and Future Directions

Understanding the growth regulation of fish is critical for breeding superior animals for aquaculture production. The study described across **Chapters II** through **V** provide foundational information for doing so in hybrid striped bass (HSB), which is the fourth most important finfish aquaculture species in the United States (Reading et al., 2018). The industry has been expanding at a rate of 10-15% per year for the past 20 years with a total national production of an approximately 12 million pounds with a farm-gate value of approximately \$50 million USD (USDA, 2019). Specifically, the described study used machine learning algorithms to determine which of the identified metabolites (**Chapter II**) and gene transcripts (**Chapter III**) present in the liver (metabolites) and white muscle tissue (metabolites and gene transcripts) of HSB are the most associated with either good (top 10%) or poor (bottom 10%) growth performance (**Chapters II-IV**). The study also included an analysis of muscle tissue fibers (**Chapter II**) to identify differences in hyperplastic muscle fiber growth patterns between growth performance groups, and a feeding study to establish the effects of feeding frequency on growth performance, measured by feed conversion ratio (FCR) (**Chapter V**).

A total of 464 and 469 metabolites were identified from HSB liver and muscle samples, respectively, for which thirty of these metabolites for both the liver and muscle tissue samples were classified as the most determinant of growth performance (Figures 2.9 and 2.10). Metabolites are small molecules and are an intermediate in a pathway or end product of metabolism and they are directly involved in normal growth, development, and reproduction. The results from this study identified several metabolomic differences between HSB growth performance groups, including changes in

metabolites related to energetics (e.g., glycolysis and TCA cycle), amino acid catabolism and anabolism (i.e., growth), and biomarkers of stress and inflammation. Curiously, rather than suggesting increased nutrient availability for organismal growth, this study suggested signs of increasing energy demand in the fish from the good-growth group compared to poor-growth group, which may suggest a greater dedication of nutrients toward growth or possibly tissue maintenance. The decreased levels of liver diacylglycerols in fish from the good-growth group was noted, which could affect nutrient storage in adipose tissue. Further experiments correlating changes in nutrient intake, feed conversion, and triglyceride storage (and potentially shifts of resources away from reproduction) with metabolomic changes may shed additional light on metabolomic markers of growth potential in HSB. For example, measuring the adiposity index could be included in addition to the hepatosomatic index (HSI) and gonadosomatic index (GSI) in future studies. Additionally, HSB plasma or blood could be sampled as an alternative of the muscle and liver tissues and that method could be a better non-lethal or non-invasive approach to provide useful markers of growth. Sampling plasma and with the association of known markers, such as IGF1, growth hormone, or insulin or microbiome differences or efficiency of feed conversion, could provide additional criteria by which to assess growth changes in this and similar studies (Islam et al., 2020). Together these strategies may allow the identification of earlier metabolic indicators of growth potential.

This study also confirmed the importance of eleven genes, including their spliced variants, for HSB growth, whereby expression of these genes were elevated in the white muscle tissue of the poor-growth group (*agtr2*, *cd22*, *cebpb*, *cxcr4*, *gadd45b*, *mmp17*, *msi1*, *rgl1*, *rgs13*, *rhou*, and *rundc3a*). The cause for such elevation, whether it be genetic

inheritance, environmental stressors in the form of dietary, or husbandry factors, remain to be determined. Thus, understanding these eleven genes and their influence on growth regulatory pathways could provide information to produce superior HSB through selective breeding. Similarly, by tracing the variation (i.e., spliced variants) associated with those eleven genes identified in this study may lead to future investigations of heritability of these growth related traits. There is likely to be genetic variation between those eleven genes in the response to targeted environmental manipulation, and genomic prediction using large full-sibling families each split into groups tested with targeted environmental treatments may answer this (Figure 6.1). Therefore, selection for improved response to environmental changes could be a route to realizing genetic improvement for impact across diverse production environments such as those in ponds or tanks used in the *National Program for Genetic Improvement and Selective Breeding for the Hybrid Striped Bass Industry* at North Carolina State University Pamlico Aquaculture Field Laboratory (NCSU PAFL) in Aurora, NC, USA.

The potential overlapping pathways that were investigated in **Objective 3, Chapter IV** and the most important metabolite and gene lists (**Objective 1, Chapter II and Objective 2, Chapter III**, respectively), were both optimal when the appropriate inputs (e.g., the machine learning approach gene list compared to the inferential statistics gene list versus the whole metabolites list instead of the top thirty list), were provided to determine the correct statistical significance scores. Hence, in the future although only significantly different molecules were analyzed in Ingenuity Pathway Analysis (IPA), a complete tested list of all measured parameters could also be provided for input and additionally, any genetic data and metadata associated to the fish also could be used. The IPA Core Analysis

of each dataset provides a very detailed overview of the impact of the expression changes on the biology. However, due to the nature of cell biology, this analysis provides more results than can be feasibly tested and observed. In such a scenario, when individual analyses are studied in detail, they can support the hypothesis, but also may obscure the complete picture if not investigated sufficiently. For example, some of the results from GeneListA, the 143 identified genes based on their inferential statistics from **Objective 3, Chapter IV**, have inconsistent results with respect to growth of cells and some individual pathways and regulators do show correlation with the fish from the good-growth group phenotype, but it is contradicted when a complete picture is observed using a different approach like BioProfiler. Although there could be multiple reasons for this, some of which like sample collection stage, statistical method used, biological variation, have been mentioned in the discussion, there is no way to determine if one particular function would be more impactful than another, unless post-transcription analyses are considered. This is the strength of using metabolomics paired with transcriptomics as in this study. Further, any of the pathways and upstream regulators mentioned in the results could be good biomarkers based on predictive analysis, but this would need to be further confirmed with gene knock-out or other such appropriate experiments in HSB.

The two gene lists from **Objective 2, Chapter III** had minimal overlap (~10%), but in comparison most of the pathways, regulators and diseases, and functions were commonly represented. This indicates the importance of interpretation analysis on gene lists, since unique lists do not in any way mean that they will affect the biology differently. The growth of cells (muscle or otherwise), inflammation, and other such diseases and functions were commonly represented in both of the gene lists. This further supports the

robustness of using orthogonal biomath approaches (e.g., inferential statistics and machine learning).

The gene list generated using the machine learning approach was generally more consistent with muscle growth phenotype (i.e., hyperplasia and hypertrophy) irrespective of the pathway analysis method (Core Analysis or BioProfiler). In contrast, the *p*-value based analysis was less consistent with this particular phenotype. It was concluded that machine learning SVMAttributeEval gene rankings are a superior way of interpreting differences in gene expression for next generation sequencing data, such as RNA-seq, as compared to the conventional *p*-value approach. Such comparisons however need to be repeated in several diverse experiments with multiple variables to provide confirmation. One approach could be to run the two comparisons on multiple public datasets (already existing) so as to be efficient with both time and finances. It would also be prudent to analyze other machine learning approaches so that the optimal ranking methods can be determined. If SVMAttributeEval gene rankings are consistently observed as the optimal approach, the scientific community will greatly benefit with such a finding, since it is often seen that conventional statistics cannot reliably link gene expression to phenotype and researchers are left using post-transcriptional analysis, metabolite profiling and other approaches in hopes of completing the targeted objective.

The fish from the poor-growth group appeared in the IPA Core Analysis to be less healthy and suffer from increased apoptosis, which as a result may lead to an inflammatory response. The metabolite data also support this, where increases in eicosanoids and cortisol in fish muscle from the poor-growth group could suggest increased stress and inflammation, which has been shown to have negative effects on fish growth. Further

studies assessing markers known to contribute to growth (e.g., plasma growth hormone, microbiome differences, efficiency of feed conversion) could provide additional criteria by which to assess changes in this and similar datasets. Furthermore, conducting a characterization experiment of stress in HSB, and diversity of cortisol induction post-stress, correlated with growth potential could expand knowledge of the effects of stress on HSB aquaculture and hence improve the growth performance of the stock. Additional experimentation should be performed to confirm the increased cell death (i.e. histology and caspase staining) and to determine if there are any conditions that are evoking an inflammatory response in the fish from the poor-growth group. Coupled with the observations of increased apoptosis and inflammation, biomarkers in the fish from the poor-growth group based on activation in *sirt1* may play a role in protecting against inflammation or cell death. A number of strategies should be taken to investigate the potential role, if any, of these transcriptional regulators during muscle growth in fish. This would include examining spatiotemporal expression of these factors during the critical stages of muscle growth, such as by immunohistochemistry or *in situ* hybridization. Once a better understanding of the expression/activation pattern of these factors occur, one could treat the fish with pharmacological inhibitors or activators on these components to determine if the fish have altered muscle growth rates relative to control (i.e., creation of the phenotype).

While filtered gene lists (based on *p*-value or ranked by SVMAttributeEval) were used to run the Core Analyses in IPA, a preferred method would be to use the full dataset (otherwise, the full reference dataset against which to do the Fisher's exact test) that is represented in the experiment and filter the dataset within IPA when doing the Core

Analysis. The major difference in these, especially in the case of metabolite data, is that IPA can take into account which molecules were represented in the experiment (rather than account for a global set of molecules). Given the minimal overlap of genes in GeneListA and GeneListB (**from Objective 2, Chapter III**), having the general direction of expression of all pathway components would help to determine if additional evidence exists that supports the different growth performance of HSB. For example, GeneListA shows a strong indication of inflammatory response, but it would be useful to see how other components/targets of inflammatory pathways are expressed as well. For example, are there changes in the targets that support the hypothesis that inflammation is down-regulated in good-growth muscle?

The FCR, in the final study, was approximately 20% higher in fish that were fed more frequently (five days per week) compared to those that were fed less frequently (three days per week). This indicates that reduced frequency feeding is more efficient, perhaps especially during times when the fish are not at optimal rearing temperatures. The result suggests that, reducing the frequency of feeding may improve the FCR of sunshine bass HSB and lower the overall cost of feeds required to raise the fish, especially as they get larger during Phase III (final growout) production and when water temperatures are not at the optimum range. Hence, the best recommendation to reduce the feed cost without affecting the growth performance of the HSB seems to be in reducing feeding when water temperatures are high in the summer and low in the winter. These adaptations are of importance for husbandry techniques, for operational efficiency, and quality control of water, all of which can be improved by decreasing the frequency of feedings when the temperature is high. In applying this method, HSB will maintain a good FCR, hence

maximize growth and minimize waste and improve water quality, whilst reducing labor costs, much like in the European seabass *Dicentrarchus labrax* (Peres and Oliva-Teles 1999). Future studies should be conducted for a longer duration and also at commercial density.

In conclusion, HSB is the fourth most important finfish aquaculture species in the U.S. and the industry has been expanding over the past 20 years. There are still some key issues that need to be addressed to sustain and promote HSB production and further support this growing industry. Particularly, the underlying molecular and cellular processes and their parallel pathways related to growth rate differences amongst HSB groups, which has not yet been investigated. This study, applied, for the first time, two powerful tools (metabolomics and transcriptomics) along with a novel machine learning-based model analysis using Support Vector Machines (SVMAttributeEval), that presented an informative list of the most important metabolite and gene biomarkers and their likely integrated pathways associated with HSB growth performance. These powerful platform approaches, which ten years ago were neither available nor accessible, provided the most practical prediction of growth performance in fish and also identified key metabolites and gene transcripts that may be indicators of growth performance. The integrated pathways of these components provided a much greater insight into the merits and mechanisms of muscle development, which will help improve HSB aquaculture production and assure the successful continuation of the industry. Finally, the feeding study conducted to evaluate the impacts of reduced frequency feeding regimes on improved feed conversion efficiency in HSB, showed a better FCR performance with fish fed less frequently (three days per week) compared to those fed more frequent (five days per week). Which suggest a reduced

feeding frequency may improve the FCR of sunshine bass and lower the overall cost of feeds required to raise the fish, especially as they get larger during Phase III (final growout) production and when water temperatures are not at its optimum range. In applying this method, HSB will maintain a good FCR, hence maximize growth and minimize waste and improve water quality, whilst reducing labor costs.

References

- Georges, M., Charlier, C., and Hayes, B.** (2019). Harnessing genomic information for livestock improvement. *Nature Reviews Genetics*, 20(3), 135-156.
- Houston, R.D., Bean, T.P., Macqueen, D.J., Gundappa, M.K., Jin, Y.H., Jenkins, T.L., Selly, S.L.C., Martin, S.A., Stevens, J.R., Santos, E.M. and Davie, A.** (2020). Harnessing genomics to fast-track genetic improvement in aquaculture. *Nature Reviews Genetics*, 1-21.
- Islam, M. J., Kunzmann, A., Bögner, M., Meyer, A., Thiele, R., & Slater, M. J.** (2020). Metabolic and molecular stress responses of European seabass, *Dicentrarchus labrax* at low and high temperature extremes. *Ecological Indicators*, 112, 106118.
- Peres, Helena and Oliva-Teles.** (1999). Influence of temperature on protein utilization in juvenile European seabass (*Dicentrarchus labrax*). *Aquaculture*, 170, 337–348.
- Reading, B.J., McGinty, A.S., Clark, R.W., Hopper, M.S., Woods III, L.C., and Baltzegar, D.A.** 2018. Genomic enablement of temperate bass aquaculture (Family Moronidae). In: H. Wang and X. Liang (Eds.) *Breeding and Culture of Perch and Bass*. Science China Press (Chinese Academy of Sciences), Beijing, China.
- USDA (United States Department of Agriculture).** 2019. 2017 Census of Agriculture, 2018 Census of Aquaculture, 3(2). Publication AC-17-SS-2.

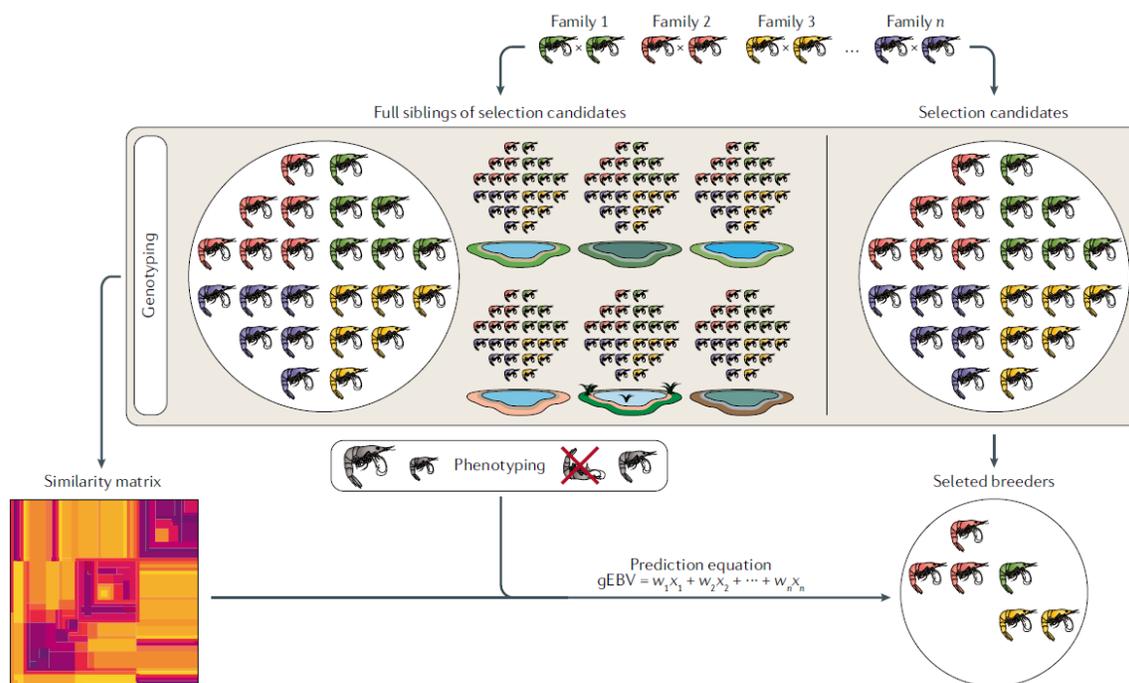


Figure 6.1. Genomic selection within an aquaculture breeding program. Full siblings from a number of families are split into selection candidates and animals for phenotypic evaluation. These full siblings of the selection candidates can be grown in different environmental conditions and phenotyped for different traits, for example, to measure growth performance traits in ponds and tanks. The selection candidates and their phenotyped full siblings are all genotyped, and a genomic relationship matrix reflecting the genetic similarity between each pair of animals is built. This relationship matrix and the collected phenotypes enable the estimation of breeding values for the selection candidates through the use of genomic selection models such as GBLUP (genomic best linear unbiased prediction) or Bayesian models (Georges et al., 2019). gEBV, genomic estimated breeding value. (from Houston et al., 2020).