ABSTRACT

Title of Dissertation: THE GENETIC BASIS OF PIGMENT PATTERN DIFFERENTIATION IN LAKE MALAWI AFRICAN CICHLIDS

Claire T. O'Quin, Doctor of Philosophy, 2014

Directed By: Dr. Thomas D. Kocher, Professor, Department of Biology

The cichlids of East Africa are a well-known group of fishes that display a wide range of phenotypic diversity, including differences in tooth shape, facial and body morphology, visual palettes, and body coloration. This diversity in phenotype was generated within the last 10 million years via an adaptive radiation. The cichlids of Lake Malawi are particularly known for their wide variation in male color patterns, which is thought to have been driven by sexual selection. Given their recent evolution, different species of cichlids can be intercrossed in the lab to identify the underlying genetic basis of phenotypic traits via a forward genetics approach. Using this methodology, I created a hybrid cross between two cichlid species, *Metriaclima zebra* and *M. mbenjii* that differ in several pigmentation traits in an attempt to identify the underlying genetic basis of these traits. After quantifying these pigmentation traits, I was able to use the Castle-Wright equation to estimate that a small number of genes underlie these traits. I was then able to identify quantitative trait loci (QTL) for

three traits, dorsal fin xanthophores, caudal fin xanthophores, and pelvic fin melanophores. The QTL for dorsal and caudal fin xanthophores were found to overlap on the same linkage group. I was able to identify and preliminarily analyze the candidate gene, AAK1, for this shared QTL region. In addition to this pigmentation work, I was also able to identify a genomic region where a potential XY sex determiner is located and analyze a candidate gene, GSDF.

THE GENETIC BASIS OF PIGMENT PATTERN DIFFERENTIATION IN LAKE MALAWI AFRICAN CICHLIDS

By

Claire T. O'Quin

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2014

Advisory Committee:
Professor Thomas D. Kocher, Chair
Associate Professor Karen L. Carleton
Associate Professor Alexandra Bely
Associate Professor Carlos A. Machado
Professor Leslie Pick

Preface

“Science and everyday life cannot and should not be separated.”

Rosalind Franklin

Dedication

To my family and friends who never stopped believing in me: thank you for your love and support through the years.

Kelly and Alana O'Quin

Michael, Dawn, and Leia Tamplain

Jerry, Linda, Casey, and Katie O'Quin

All of my labmates, past and present

Thomas Kocher

Acknowledgements

First, I would like to thank my parents for fostering my curiosity for the world around me. They never flinched when I had jars of bugs or brought home a baby bird in a shoebox. Next, I thank my husband Kelly O'Quin, who has been by my side through thick and thin. I would have never finished this work if not for his love and constant encouragement. I also thank my daughter, Alana O'Quin, who "helped" mommy write for the first six months of her young life. I am forever indebted to my advisor, Tom Kocher, who gave me a second chance at getting my PhD and patiently taught me all I know about genetics. I thank each of my committee members-Karen Carleton, Alexa Bely, Leslie Pick, and Carlos Machado-for their time and helpful advice throughout this process. Finally, I thank the following individuals, institutions, and funding sources that supported the research contained in each chapter:

Chapter 2. This chapter was previously published in the *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* and is reprinted here under John Wiley and Sons license no. 3304261411902. I once again thank my co-authors for their hard work and help on this chapter. Together we also thank Jennifer Ser for help with animal husbandry, Karen Carleton for assistance with analysis of bar spacing and comments on the manuscript, and Kelly O'Quin for assistance with statistical analysis and comments on the manuscript.

Chapter 3. This chapter was previously published in the journal *BMC Genomics* and I reprint it here under the Creative Commons Attribution 2.0 Generic License. I would like to thank my co-workers for their hard work and help in this chapter.

Together we also thank Kelly O'Quin for help with the QTL analysis and Sarah Robbins, Ginni LaRosa, and Amanda Auerbach for help with microsatellite genotyping. TDK was supported by National Science Foundation grant DEB-1143920. CTO was supported by an award from the Guy Jordan Endowment Fund of the American Cichlid Association.

Chapter 4. CTO was supported by an Ann G. Wylie Dissertation Fellowship from the University of Maryland during the writing of this chapter.

Chapter 5. I once again thank my co-authors, two of which were undergraduate volunteers. CTO was supported by an Ann G. Wylie Dissertation Fellowship from the University of Maryland during the writing of this chapter.

Table of Contents

List of Figures

List of Tables

Chapter 1: A Review of Pigmentation Genetics in Fishes

Introduction

Evolution has produced a diverse array of animal pigment patterns. Colors and patterns like the stripes on a zebra or patches on a giraffe often serve to confuse predators. Pigmentation can communicate to a predatory bird that a butterfly is distasteful. Mimetic species take advantage of this by sharing these color patterns to gain protection from these predators (Beldade and Brakefield, 2002). In cichlids, strawberry poison frogs, and many other organisms, pigment patterns alone are thought to provide visual cues for females to choose their conspecific in mating assays (Kidd et al., 2006; Summers et al., 1999; Candolin, 2003). Dwarf chameleons have been shown to adjust their color response for camouflage based on the predator pursuing them (Stuart-Fox et al., 2008). These diverse roles of pigmentation have led to it being studied from a variety of perspectives in a number of model and non-model systems.

Fishes are becoming an increasingly useful model to study pigmentation not only because they possess a variety of pigment cells types, but also because they possess transparent embryos, have short generation times, are able to be reared in the lab, and display an array of pigment patterns in closely related taxa (Parichy et al., 2006a). The majority of pigmentation work done in fishes has been done in the model organism, *Danio rerio* (Kelsh et al., 2009). The ricefish, *Oryzias latipes*, is also a useful model system for pigmentation studies since it possesses an additional cell type, the leucophore, that is only found on the zebrafish tail, but is widespread on the

ricefish body (Kelsh, 2004).

This chapter will review our present knowledge on pigmentation genetics in fishes. I will first briefly provide background information about the major pigment cell types found in fishes. Next, I will talk about the cellular origin of pigment cells, their differentiation, and subsequent migration. I will then talk about the formation of the pigment containing organelles. Next, I discuss work that examines how pigment cell interactions contribute to forming the adult pigment pattern of fishes. Finally, I will discuss the role hormones can play in changing pigment patterns.

Pigment Cell Types

Endotherms, which include birds and mammals, possess only one cell type, the melanocyte. These cells produce the black to brown pigment melanin. In mammals, the coloration differences that we observe results from the production of two main melanin types: phaeomelanin (red to brown pigment) or eumelanin (black pigment), or the absence of pigment (Mills and Patterson, 2009). It is the distribution and spacing of pigment deposition from the melanophore onto the hair that produces the varied fur patterning that we observe (Kelsh, 2004). Birds also produce the phaeomelanin and eumelanin in their melanophores. However, the bright colors we observe are derived in one of two ways. Bright reds, yellows, and oranges are due to the presence of carotenoid pigments obtained from dietary sources. Colors such as blue or green are actually the result of structural color, where short-wavelength light is reflected off nanostructures in the feather itself (Mills and Patterson, 2009).

Ectotherms, such as fishes, amphibians, and reptiles, possess several different pigment cell types. These include melanophores, which contain the black pigment

melanin; xanthophores, which contain yellow, carotenoid or pteridine-derived pigment; iridiophores, which contain silvery, purine derived pigment; and three lesser-known cell types, white leucophores, red erythrophores and blue cyanophores (Mills and Patterson, 2009; Parichy et al., 2006a). Fish chromatophores maintain their pigment intracellulary (Mills and Patterson, 2009). While endotherms generate pigmention variation by varying the amount of melanin produced, ectotherms can do this for the other pigment types as well, and can even modify the arrangement of their chromatophores to achieve variation in their pigmentation patterns (Mills and Patterson, 2009). Electron microscopy studies have shown that the pigment cells are arranged in a unique fashion in the dermis. In zebrafish, the interstripe regions are composed of xanthophores underlain by type S iridophores (iridophores with small, reflecting platelets). In the stripes, the type S iridophores are underlain by melanophores and type L iridophores (iridophores with large reflecting platelets) (Kelsh et al., 2009).

Melanophores

Melanophores are chromatophores that contain the black pigment melanin (Fig. 1-1A&B). Melanin is synthesized from tyrosine within the melanosome, an organelle in the melanophore (Braash et al., 2007). Mutations that cause disruptions in the melanin synthesis pathway have been linked to lighter coloration across a variety of taxa (Hoekstra, 2006).

Xanthophores and Erythophores

In fishes, yellow and red patterns are generated by two pigment cell types, xanthophores and erythrophores, respectively. Xanthophores possess pteridine

pigments synthesized from GTP or carotenoid pigments (Braash et al., 2007) (Fig. 1-1A). There has been some debate as to whether these are two distinct cells types. Resolution of this debate has not been readily achieved for several reasons. First, both of these red and yellow colors are produced by pteridine and carotenoid pigments, and both of these cell types can contain both pigment types (Mills and Patterson, 2009). Also, xanthophore color has been shown to vary in closely related species of *Danio* fishes. Lastly, the models used for xanthophore study lack erythrophores (Mills and Patterson, 2009).

Iridophores and Leucophores

Iridophores and leucophores are the reflective pigment cells (Fig. 1-1B). Their appearance is due to the presence of guanine and hypoxanthine crystals or platelets within the cells. Leucophores possess short, vertically arranged platelets and appear white, while iridophores possess long, horizontally arranged platelets and appear iridescent. Blue coloration in fish is mostly due to iridophores underlain by melanophores (Mills and Patterson, 2009).

Pigment cell origins

Neural crest cells are a transient, migratory cell population that arise from the neural tube and migrate extensively throughout the body. In fishes, the neural tube arises as a ventral thickening of the dorsomedial ectoderm (Parichy et al., 2006a). These cells differentiate to form many structures, including neurons, glia, ganglia, secretory cells, smooth muscle cells, bone, cartilage, and pigment cells (Sauka-Spengler and Bronner-Fraser, 2008). There are several processes involved in neural crest cell development, some which overlap in time. First, specified precursors must

be generated from the multipotent neural crest cells. Then, these precursors must be patterned in the embryo. This patterning involves movement in a migratory pathway and then localization to a specific site. Once the cells are localized, they often depend on a variety of trophic factors to survive. Finally, cells must express specific differentiation products to form specific cell types (Fig. 1-2) (Kelsh et al., 1996).

There is evidence that pigment cell precursor specification in fishes occurs very early on at the neural tube. This is based on observations of the specific migration paths that the pigment cells in fish follow. Iridoblasts travel along the medial pathway, xanthoblasts travel along the lateral pathway, and melanoblasts travel along the ventral and lateral pathways (Kelsh et al., 2009). The transcription factor Sox10 plays a key role in pigment cell specification since zebrafish possessing a mutation in this gene display acute chromatophore defects in all pigment cell types (Parichy et al., 2006a). Below I will discuss genes that are involved in the specification and migration of each class of pigment cells.

Melanophore Specification

Most of what is known about pigment cell development is based on studies of melanophores, because these cells are present in the dominant model organisms, such as mice (Bennett and Lamoreux, 2003; Kelsh et al., 2009). Several of the genes involved in melanophore development are conserved across taxa, including *mitf,* which encodes a basic helix-loop-helix/leucine zipper protein key to the early specification of melanophores (Fig. 1-2). Zebrafish *nacre* mutants, who possess a premature stop codon in *mitf*, never develop body melanophores (Fig. 1-3A). However, the retinal epithelium is still pigmented in nacre mutants, indicating that the

melanin synthesis pathway is still intact (Lister et al., 1999).

Other genes are necessary for the survival of melanophores at different points of development. Zebrafish *sparse* mutants carry a mutation in *kit*, which encodes a type III receptor tyrosine kinase. Zebrafish *sparse* mutants exhibit half the normal number of stripe melanophores compared to wild-type fish, which is a very similar phenotype to *kit* mutants in mammals (Fig. 1-3B). Parichy et al. (1999) found that *sparse* mutants have an accumulation of melanophores near their site of origin, posterior to the otocysts and at the lateral margins of the branchial arches. This is in contrast to wild-type embryos in which melanophores disperse dorsally and anteriorly. Additionally, mutant larvae had TUNEL positive cells (labeling apoptosis) beginning 4 days post fertilization and continuing through at least 11 days post fertilization. These results indicate that *kit* is needed for survival of the embryonic/early larval melanophores (Parichy et al., 1999).

Zebrafish *picasso* mutants have a normal early larval pigment pattern but the adult pigment pattern is disrupted due to a smaller population of metamorphic melanophores (Fig. 1-3C). *Picasso* mutants are the result of mutations in *erbb3b*, an epidermal growth factor receptor-like tyrosine kinase. Budi et al. 2008 showed that wild-type embryos treated with pharmacological inhibitors of ErbB signaling develop a normal early larval pigment pattern. However, after metamorphosis, they could not be distinguished from the *picasso* mutants. Wild-type fish treated later in development at the early larval stage could not be distinguished from the controls. This suggests that ErbB signals are needed in embryos for the adult pigmentation pattern to form properly (Budi et al., 2008). Subsequent experiments showed that

adult pigmentation formation was most affected between 14 to 22 hours post fertilization. This time period corresponds to the time period when neural crest cells are migrating. This study points to the need for ErbB signaling in early development so that metamorphic melanophores can be generated later in development (Budi et al., 2008).

Rose mutants in zebrafish contain mutations in endothelin receptor b, 21-amino acid proteins that are ligands for seven pass G protein-couple endothelin receptors (Parichy et al., 2000b). These mutants display a normal larval pigment pattern early in development; however, the adult pattern is perturbed in that they possess stripes dorsally but spots ventrally (Fig. 1-3D). These fish develop fewer melanophores than their wild-type counter parts during late metamorphosis. This in the end causes adult *rose* mutants to have only half the number of melanocytes in their stripes. This indicates that the expression of *ednrb1* is essential for the development of late metamorphic melanocytes. *Rose* mutants also have a defect in the number of iridophores that they possess (Parichy et al., 2000b). Expression studies using *in situ* hybridization resulted in two models to explain the role of *ednrb1* during melanophore development. The first model suggests that *ednrb1* is only responsible for the development of a certain set of pigment cell populations and their precursors that form the late metamorphic melanophores, and is not used to generate pigment cells for the larval or early metamorphic patterns. The ability to generate these genetically distinct populations of precursors can be accomplished by the redundancy of gene activity due to the presence of *ednrb1* paralogues. The second model suggests that *ednrb1* is needed for the development of a single population of pigment cell

precursors that are then recruited at different times during development to differentiate into different cell types. This population can contribute to early larval, early metamorphic, and late metamorphic melanophores. In mutants, only enough precursors are generated to contribute to the early stages of development, but not the later stages. More work is needed to determine the role of *ednrb1* during pigment pattern development (Parichy et al., 2000b).

Iridophore Specification

Two recent studies using zebrafish have provided further insight into the genes that play a role in the development of iridophores. Zebrafish *shady* mutants have reduced numbers of iridophores, with an allelic series showing gradation in the number of iridophores present, as well as embryonic survival. Homozygotes with strong alleles have very few iridophores and die as larvae, while mutants for weaker alleles show reduced numbers of differentiated iridophores. Adult viable mutants were identified and display no problems with the embryonic phenotype when homozygous for their mutant alleles. Since other neural crest derived structures appeared normal in the mutants, it was concluded that *shady* acts specifically on the iridophore neural crest lineage (Lopes et al., 2008).

Positional cloning determined that *shady* mutations corresponded to the *ltk* (leukocyte tyrosine kinase) gene, an insulin receptor-like receptor tyrosine kinase (Fig. 1-4A). Function of *ltk* was confirmed by both morpholino injections and *in situ* hybridization in wild-type embryos. Injection of wild-type embryos with an *ltk* morpholino resulted in *shady* mutant phenocopies. Using *in situs*, it was determined that *ltk* expressing cells corresponded to the iridophore locations. Expression studies

of *ltk* during the stages when neural crest cells are specified and begin to differentiate in wild-type versus mutant zebrafish embryos revealed that mutant embryos showed reduced numbers of *ltk*-positive cell, indicating a failure of iridoblast specification in the *shady* mutants (Lopes et al., 2008).

Work by Curran et al. (2010) provided further insight into the differentiation process of iridophores using zebrafish as their model. Work from their lab suggests that some neural crest cells develop directly into either melanophores or iridophores, while other neural crest cells pass through a bi-potent stage in which they can develop into either one or the other cell type. These bi-potent cells express *mitfa*. The expression of Foxd3 than pushes the cell towards one pigment cell fate or the other. They suggest that Foxd3 represses *mitfa* and promotes expression of *pnp4a*, thus pushing the cell into an iridophore state. In cells where Foxd3 is not expressed and *mitfa* expression is maintained, melanophores develop. This whole process occurs very early in development, between 18-24 hours post fertilization. While these two studies have proven valuable in trying to understand more about the development of this cell type, more work is needed to determine if any other genes play a role in the development of this cell type.

Xanthophore Specification

Very little is known about xanthophore specification. The zebrafish *panther* mutant was mapped to the interval containing the gene *fms*, a type III receptor tyrosine kinase (Parichy et al., 2000a). Zebrafish with this mutation have a normal pattern of larval melanophores, but the adult melanophores fail to organize themselves into stripes (Fig. 1-4B). Surprisingly, this mutation does not affect the

formation of melanophores, but instead causes the loss of embryonic and adult xanthophores. This indicates that interactions between melanophores and xanthophores are probably necessary for the adult stripes to form properly (Parichy et al., 2003a). Experimental results indicate that *fms* acts non-autonomously on melanophores to arrange them into stripes and autonomously on pigment cell lineages to promote xanthophore development (Parichy et al., 2003a). During early pigment pattern metamorphosis, *fms* is needed to recruit precursor stem cells to a xanthophore fate. During the middle stages of pigment pattern metamorphosis, *fms* expressing cells of the xanthophore lineage help to organize melanophores into stripes. From the last stages of metamorphosis and onward, *fms* is needed maintain the xanthophore lineage and thus the melanophore stripes (Parichy et al., 2003a). It should be noted here that *fms* is also known as *csf1r*.

Two other genes known to be important in xanthophores development are *pax3* and *pax7*. *Pax3* was found to function in early specification of xanthophores. In zebrafish where *pax3* was knocked down using morpholinos, four known xanthoblast markers, were absent, indicating an absence of xanthophores precursors in those fish. Additionally, it was found that *pax3* is down regulated as migration commences. *Pax3* is also thought to have a role in later development, since fish with their *pax3* expression knocked down by morpholinos also have marked reduction in the trunk of yellow pigment normally produced in xanthophores, indicating it to be required for xanthophores terminal differentiation to produce pterinosome organaelles and yellow pigment. Also, *pax3* was found to be needed for the maintenance of *sox10* expression during the later stages of neural crest development, since the neural crest display

defects in those fish in which *pax3* was knocked down (Minchin and Hughes, 2008). *Pax7* was found to have a role later in development for xanthophores. *Pax7* marks xanthoblasts as they migrate on the lateral pathway and accumulates as xanthophore differentiation advances. Additionally, in fish where *pax7* expression was down regulated, they had reduced yellow pigmentation, however xanthophores specification proceeded, providing further evidence for a later role in development for this gene (Minchin and Hughes, 2008).

Biogenesis of Pigment Organelles

Within the different fish pigment cells, pigment is contained in what are known as lysosome-related organelles (Fig. 1-5). For example, melanophores contain melanosomes, xanthophores contain xanthosomes, and iridophores contain iridosomes (Navarro et al., 2008). Most of what is known about lysosome-related organelle development comes from studies of melanosomes. Melanosomes arise from the trafficking of different soluable and membrane proteins into endosomes (Navarro et al., 2008). Melansosomes then go through four stages of development. Stage I melanosomes contain internal membranous vesicles and an irregular fibrous structure (Marks and Seabra, 2001). Stage II melanosomes have a striated appearance due to the presence of intralumenal fibers forming parallel arrays. At stage III, the melansome darkens and thickens with the deposition of melanin, and by stage IV, the intralumenal structure is masked by the accumulated melanin (Fig. 1-5) (Marks and Seabra, 2001).

Several mutants in zebrafish and one mutant in medaka have pigmentation defects that cause their melanophores to appear diluted. These mutations are in genes

for proteins in multimeric complexes that play important roles in the formation or functions of the melanosome. The first set of mutants discussed has a reduced number of melanosomes. The zebrafish *fading vision* mutant was mapped to the protein Pmel17 (Fig. 1-6A). This protein is the main component of the intralumenal fibers seen in stage II melanosomes. Melanosomes are also reduced in mutants for Class C vesicular protein sorting subunits. These core proteins associate with effector proteins to form a complex known as HOPS. These proteins are thought to be important in mediating vesicular traffic to melanosomes (Navarro et al., 2008).

Another set of dilution mutants are those that affect melanosome function and vesicular traffic via mutations in genes endcoded vacuolar ATPases, lysosomal intetral membrane proteins, and coat protein complexes. Seven zebrafish mutants were identified in vacuolar ATPases. Vacuolar ATPases are necessary for creating acidic environments, which are required for melanosome function. Zebrafish mutants for lysosomal integral membrane protein 2 appear diluted up to 3 days post fertilization, at which time they return to normal pigmentation. While its function is not fully understood, it plays a role in targeting enzymes to lysomes. The recovery of phenotype is thought to occur due to the presence of three orthologs in the zebrafish genome that rescue the phenotype. Finally, mutants with mutations in the coat protein complex I subunits have melanocytes that proliferate and migrate properly, but are not pigmented. These genes help form nonclathrin coats for traffic vesicles (Navarro et al., 2008).

Two other mutants also display defects in melanophore formation, *touchtone* and *lockjaw* mutants. Zebrafish *touchtone* mutants possess mutations in transient

receptor potential melastatin 7 (*trpm7*) and display defects in embryonic melanophores (Fig. 1-6B). Mutant embryos are indistinguishable from their wild type siblings until about 30 hpf, when melanophores appear smaller in size and number. By 48 hpf, melanophores and melanoblasts are absent, which suggests death of this cell population. However, at approximately 72 hpf, melanophores begin to appear again (McNeill et al., 2007). It was found that during the early development of the embryo, cells are dying by necrosis, as opposed to apoptosis. Upon further investigation, it appears that melanophores in the mutants are dying as a result of cytotoxic intermediates of the melanin synthesis pathway. Trpm7 is thought to play in a role in the breakdown of these intermediates between 30-72 hpf. At 72 hpf, it is hypothesized that a homologous protein takes over the role of trpm7, and melanin synthesis and melanophore development are restored (McNeill et al., 2007).

Zebrafish *lockjaw* mutants have defects in the gene transcription factor AP-2. These mutants not only have defects in their pharyngeal skeleton, but also in their pigment cells. By 25 hpf, these mutants display a smaller number of melanophores (Fig. 1-6C). Iridophores were also reduced in these mutants. However, xanthophores developed normally. This suggests that it is important in the early specification neural crest cell populations into sublineages. *Lockjaw* mutants also have a defect in the spatial arrangement of their pigment cells. For example, in the tail, the dorsal stripe is normal, but pigment cells are absent from where the lateral and ventral stripes form. This suggests that transcription factor AP-2 is also important in the proper migration of pigment cells, with evidence that this is accomplished via regulation of *kit* (Knight et al., 2004).

Pigment Cell Interactions and Patterning

Work in zebrafish has shown not only the importance of specific genes in pigment patterning, but also the interactions of different pigment cells themselves in the creation of the striped pattern observed in these fish. Zebrafish possessing *leopard* alleles exhibit a range of patterns including undulating stripes, broken stripes, and spots (Fig. 1-7A). Through positional cloning, Watanabe et al. 2006 determined that the *leopard* allele corresponds to the *connexin41.8* gene. Connexins are components of gap junction channels. These channels are composed of two connexons, which in turn are composed of six connexin subunits. Gap junction channels allow for the direct intracellular transfer of ions, small metabolites, and secondary messengers (Watanabe et al., 2006). Experimental results indicate that *connexin41.8* might be crucial in pattern formation due to the role in plays in cell-cell interactions. Not only could it be crucial for melanophore-melanophore and xanthophore-xanthophore interactions, but it could also play a role in melanophore-xanthophore interactions, which are thought to play an important role in stripe formation (Watanabe et al., 2006).

Heterozygous *jaguar*/*obelix* mutants display fewer, broader stripes compared to wild-type fish, while homozygous *jaguar/obelix* mutants have broader yellow xanthophore bands and thinner melanophore stripes (Fig. 1-7B). Through linkage analysis, a candidate gene, *Kir7.1*, was identified for this mutant phenotype (Iwashita et al., 2006). *Kir7.1* belongs to the inwardly rectifying potassium channel family, and mutations in this gene abolish potassium conductance (Iwashita et al., 2006). Mutants subjected to rapid background adaptation, which is under neural control, had

melanophores that did not disperse in response to a change from light to dark backgrounds. This implies that the mechanism of melanophore dispersion and aggregation is somehow impaired in the *jaguar*/*obelix* mutants (Iwashita et al., 2006). It is still not well understood how this translates to changes in the width of stripes. Further work is needed to characterize the complex interactions among pigment cell types (Iwashita et al., 2006).

Zebrafish *choker* mutants have disorganization of the mytomes in the anterior somites, as well as pigmentation defects, with a band of melanophores that bridge the dorsal and ventral stripes (Fig. 1-7D). These defects are the result of an altered expression pattern of stromal cell-derived factor 1a (Sdf1a). Sdf1a acts as a chemoattractant for migrating melanophores. The absence of this ligand causes issues with migration of xanthophores and melanophores along the lateral pathway. In fact, *choker* mutants had melanophores migrating into the lateral pathway at times when wild-type individuals had melanophores restricted to one of the pigment stripes. This causes an accumulation of melanophores in the collar region. Additionally, xanthophores cleared from this region at approximately the same time melanophores accumulated. While both events happen at the same time, they are independent of each other. Iridophores in this mutant appear normal (Svetic et al., 2007).

In order to better understand how melanophore and xanthophore interactions might contribute to pattern formation in zebrafish, Nakamasu et al. (2009) performed four laser ablation experiments. The first experiment sought to examine long-range interactions and their influence on the development of new pigment cells. When melanophores were ablated in a black stripe, melanophores formed to regenerate the

pattern. However, when xanthophores were ablated in addition to the melanophores, the melanophores that regenerated in the black stripe were decreased. When just xanthophores were ablated in a yellow stripe, xanthophores developed to regenerate the pattern. When melanophores in the neighboring black stripe were ablated at the same time as the xanthophores, melanophores developed to take the place of the xanthophores in the yellow stripe. The authors concluded that melanophores in neighboring stripes have repressive actions on the development of melanophores in distant places.

In the second experiment, the authors sought to determine how cells from neighboring stripes influence pigment cell survival. They first ablated melanophores from two black stripes that surrounded a yellow stripe. Xanthophore survival was normal. When xanthophores from two yellow stripes that surrounded a black stripe were ablated, the melanophores in the black stripe showed a reduced survival rate. This led to the conclusion that xanthophores in neighboring stripes are needed for melanophore survival and that xanthophores have influence over a long distance (Nakamasu et al., 2009).

The third experiment set out to determine the short-range effect of adjacent cells. When lone melanophores were observed in the regions that would become yellow stripes, they either experienced cell death or migrated to a black stripe region. When xanthophores were ablated from the neighboring regions of the lone melanophores, the melanophores remained in the yellow stripe region and little cell death was observed. This indicates that xanthophores in the control experiment caused elimination of the melanophores (Nakamasu et al., 2009).

The fourth and last experiment examined the short-range effects of melanophores on xanthophores. To enhance survival of melanophores in regions where they normally disappear, melanophores were ablated in neighboring black-striped regions. This allowed some melanophore colonies to survive. Xanthophores near the melanophore colonies disappeared, which suggest that melanophores can act at a short range on xanthophores by competing with them (Nakamasu et al., 2009).

As a result of these experiments, the authors were able to develop a model to explain how the melanophores and xanthophores interact with each other. Activation of melanophores inhibits xanthophores locally and in turn further activates melanophores in the region. Inhibiting xanthophores can cause a decrease in melanophores distally. These results prove consistent with computer modeling in that there is long-range activation but local inhibition. While this study helped to provide a better understanding of how pigment cell actually interact with one anther, there is still work that needs to be done to understand the molecular basis of this interaction and how these can create creating Turing patterns (Nakamasu et al., 2009).

Zebrafish mutants for *basonuclin-2* (*bnc2*) have also been used to investigate stripe formation. These mutants have fewer melanophores, xanthophores, and iridophores and as a result have no body stripes or interstripes (Fig. 1-7C). RT-PCR showed reduced expression of the receptor ligands *kitlga* and *csf1a* and *csf1b* in *bnc2* mutants compared to wild type fish. *Kitlga* is known to be important for the development of melanophores, while *csf1a* and *csf1b* are known to be important in the development of xanthophores. Restoring expression of these genes helped rescue these two pigment cell types; however, the stripes were not recovered, indicating

some other factor must play a role in stripe organization (Patterson and Parichy, 2013).

Further investigation showed that iridophores are important in promoting the development of xanthophores and melanophores. Iridophores appear to provide a localized source of *csf1a* and *csf1b* that help in the specification of xanthophores within the interstripe. In experiments in *csf1r* mutants, where xanthophores are absent, melanophores mis-patterened, indicating that iridophores influence melanophores independently of xanthophores. Taken together, these results indicate a role for iridophores in the normal development of stripes and interstripes (Patterson and Parichy, 2013).

Frohnhofer et al. 2013 also investigated the role of iridophores in stripe formation by observing the interactions of pigment cells in single and double zebrafish mutants for each of the three pigment cell types. They developed a model for the interactions between pigment cells that lead to stripe formation. S-iridophores support high melanophore numbers in their neighborhood and cause their aggregation into stripe regions. This is the result of short-range repulsion between iridophores and melanophores. Xanthophores and S-iridophores display a mutual attraction. Xanthophores might cause a positive signal in order to maintain melanophores at a distance. Overall, it appears that xanthophores and iridophores support aggregation of melanophores by reducing their tendancy to avoid each other. The authors also suggest that horizontal myoseptum serves as a prepattern for the first interstripe (Frohnhofer et al., 2013).

Neural and Hormonal Pigmentation Changes

Neural signaling and hormones can also play a role in the development and appearance of pigment cells. There are two main mechanisms thought to control color change in teleosts. During physiological color change, pigment moves rapidly within the cell, either through contraction or expansion. This is a rapid response and the results can last anywhere from a few seconds to a few hours based on if the response was mediated via neural or hormonal mechanisms. This response could be either contraction or expansion of the cell. For example, in *Astatotilipia (Haplochromis) burtoni*, barred males expand their melanophores while barless males contract their melanophores in response to the same concentration of the neurotransmitter norepinephrine (Muske and Fernald, 1987).

Morphological color change occurs over a longer period of time, usually weeks or months. Pigmentation is either accumulated or reduced by changing the amount of pigment within the cells or the actual number of pigment cells. This change is usually mediated by a hormonal response. In teleost fish that were allowed to adapt to a black background, an increase of α-MSH was observed. This hormone has been shown to increase melanophore density, as well as cause dispersal of melanophores. When adapting to lighter backgrounds, studies in several fish showed that melanophores degenerated and melanin was discharged at the skin surface. It is thought that norepinephrine depletion causes apoptosis of melanophores (Sugimoto, 2002).

The medaka mutant *color interfere* has a mutation in the gene encoding somatolactin, a hormone in the growth hormone/prolactin family. This mutation

causes severe truncation of the resulting protein. These mutants have a decreased number of xanthophores and an increased number of leucophores. The other two cell types, iridophores and melanophores, have no change in number. Observations of somatolactin transcription during morphological background adaptation showed a decrease of somatolactin transcription in wild-type fish kept on a white background and no up-regulation when they were placed against a black background. It was concluded that somatolactin suppresses the growth and differentiation of leucophores (Fukamachi et al., 2004).

Dissertation Goals

Most of what we understand about pigmentation in fishes has been learned from mutational studies. Very little work has been done to understand if the same genes found to be important in these studies are the same genes utilized in natural populations to create the diversity of pigment patterns observed. To investigate this question, I utilized the African cichlid system. Within the last 10 million years approximately 2,000 species of cichlid fishes have evolved in the Great African Rift Valley Lakes, making them the most rapidly speciating group of vertebrates on earth. One of these lakes, Lake Malawi contains more than 500 species of cichlids. The cichlids from Lake Malawi are well known for their diversity of color patterns, which is a target for sexual selection in these species (Kocher, 2004). The genus *Metriaclima* is one of the most diverse groups of fish in Lake Malawi, based on the variety of their pigmentation patterns. Due to their recent evolutionary divergence, we can exploit the natural variation found within the lake to create hybrids to perform QTL analysis.

The goals of my doctoral research are to determine the genes underlying a set of pigmentation traits in male African cichlids. I first identified traits that possess quantifiable differences in male pigmentation between *Metriaclima zebra* and *Metriaclima mbenjii*. I then utilized a hybrid cross between these two species to first, determine chromosomal regions associated with these traits and second, to determine candidate genes associated with male pigmentation. Additionally, I investigated sex determination in my hybrid cross. These analyses will provide important insight into the genes underlying male pigmentation, a trait thought to be under sexual selection and helping to drive speciation of African cichlids in Lake Malawi.

Figures

Figure 1-1. Examples of the three main pigment cell types found in African cichlids. A) Xanthophores form the orange background and the black dots are melanophores. B) Iridophores form the light blue backbground and the black dots are melanophores. The scale bars represent 0.50 mm.

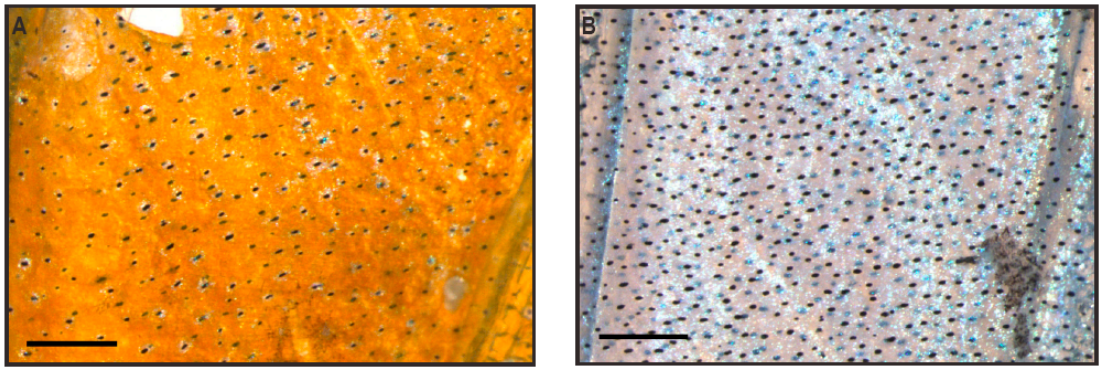


Figure 1-2. Model of pigment cell specification pathways. All pigment cells originate from the neural crest. Neural crest cells that express *sox10* and *foxd3* are pigment cell precursors that can become one of three pigment cell types in the embryo. Cells that express *pax3* and *csf1r* become xanthophores. Those that express *ltk* and *pnp4a* become iridophores. Cells that express *mitfa* and *kita* become melanophores. Melanophores and iridophores can also form from a bipotent precursor that is *mitfa*+ and *foxd3*-. In cells where *foxd3* becomes expressed, *mitfa* is inhibited and the cell is fated to become an iridophore. Otherwise, the cell will express *mitfa* and become a melanophore. There are also neural crest cells that will express *erbb3b* and serve as a population of metamorphic pigment cell precursors. These cells will eventually express *ednrb1* and become bi-potent precursors that can become metamorphic melanophores or iridophores. Model based on information found in Budi et al., 2011 and Curren et al., 2010.

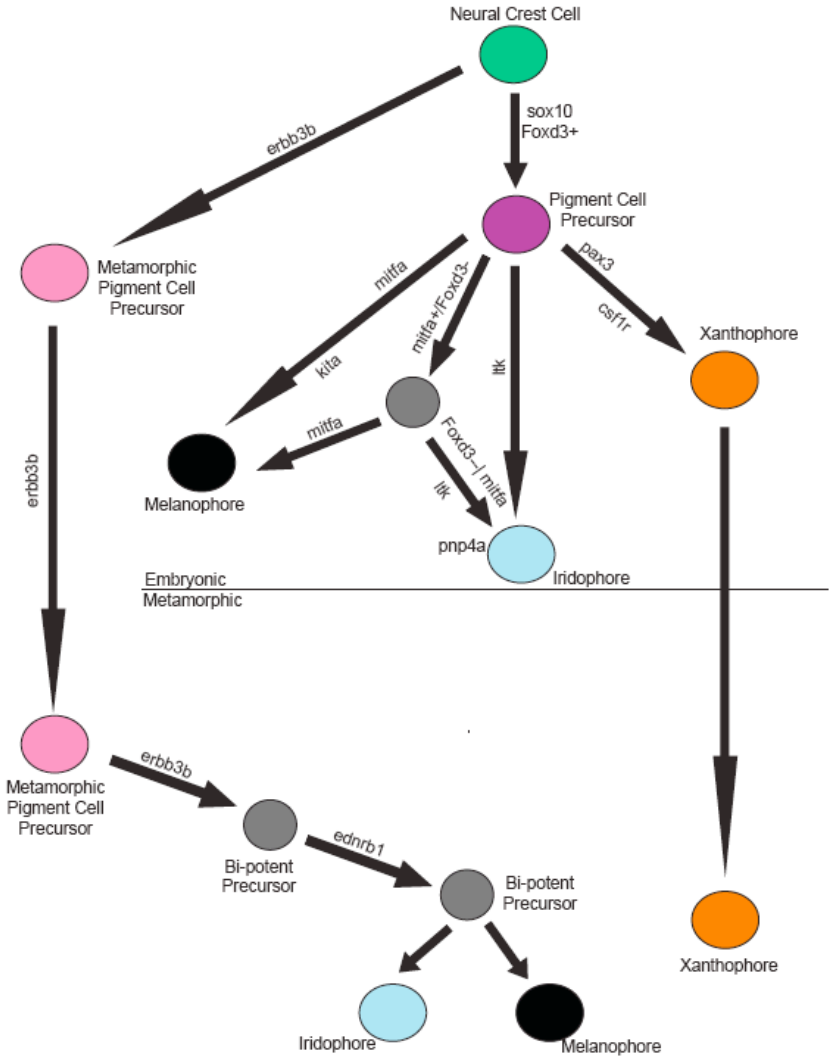


Figure 1-3. Melanophore specification zebrafish mutants. A) *nacre* mutant possessing a mutation in *mitf*. B) *sparse* mutant possessing a mutation in *kit*. C) *picasso* mutant possessing a mutation in *erbb3b*. D) *rose* mutant possessing a mutation in *ednrb1*. All images from the Parichy lab website (http://faculty.washington.edu/dparichy/Zebrafish_mutants.html).

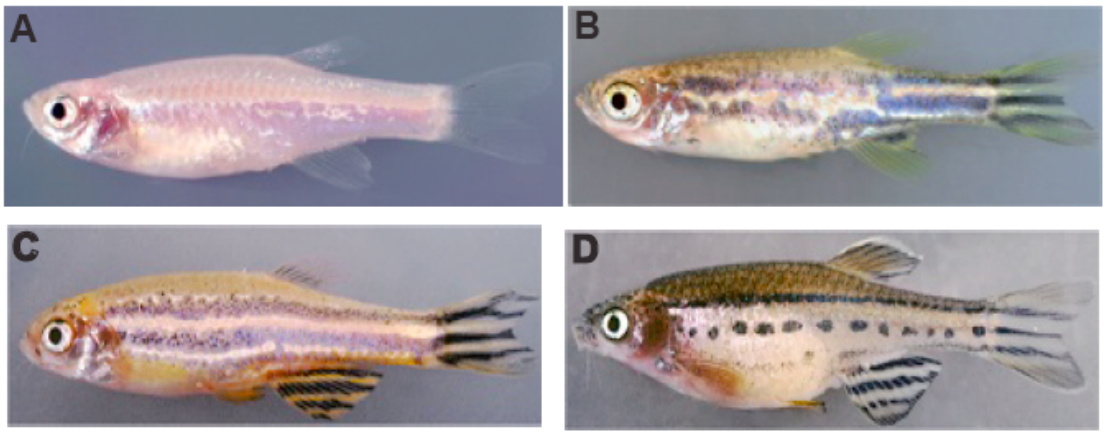


Figure 1-4. Iridophore and xanthophore specification zebrafish mutants. A) *shady* mutant possessing a mutation in *ltk*. B) *panther* mutant possessing a mutation in *csf1r*. All images from the Parichy lab website (http://faculty.washington.edu/dparichy/Zebrafish_mutants.html).

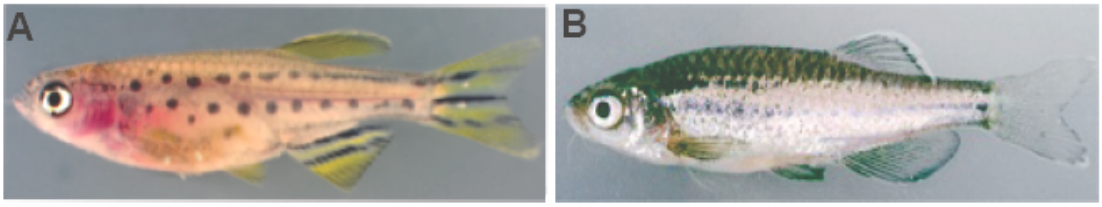


Figure 1-5. Pigment organelle biogenesis pathway. Genes where mutations have been found in fish are listed here and labeled on the figure: COPI, Pmel17, HOPS, V-ATPase, Limp2.

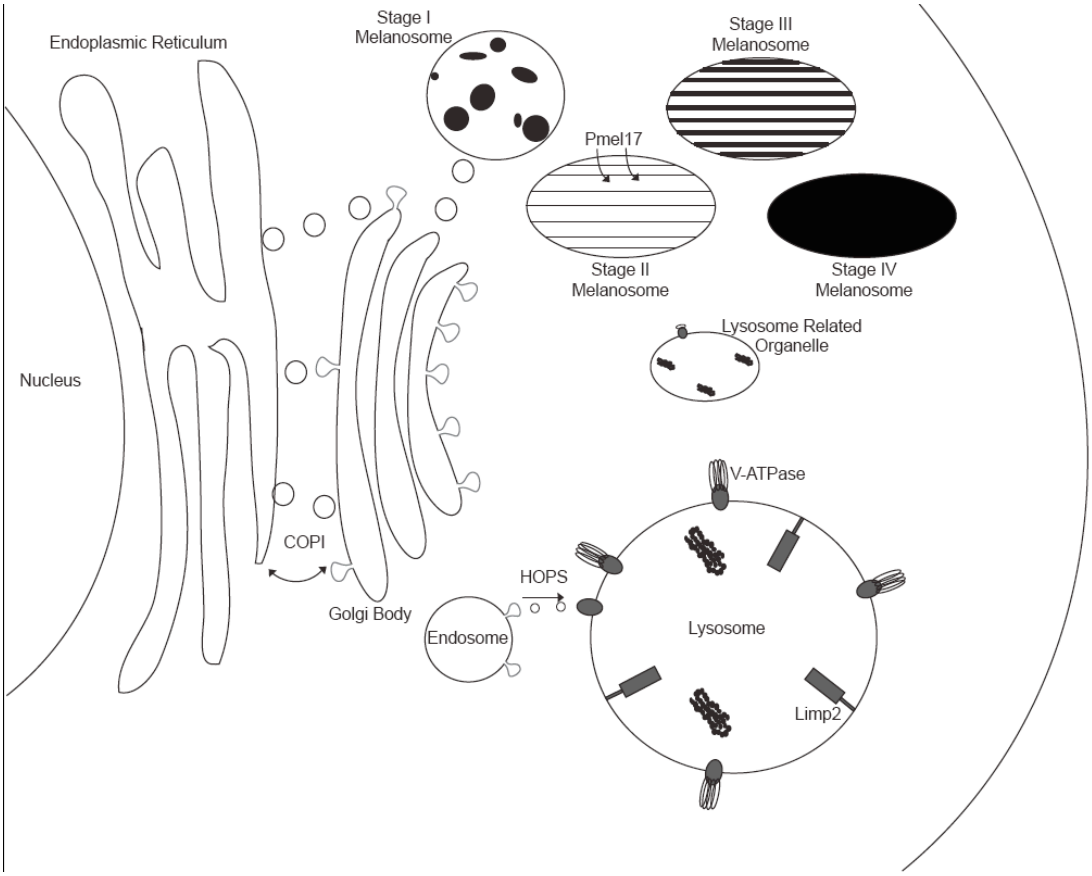


Figure 1-6. Pigment organelle biogenesis zebrafish mutants. A) Top panel is a wild type zebrafish embryo, while the bottom panel is the *fading vision* mutant that possesses a mutation in *pmel17* (Schonthaler et al., 2005). B) Left panel shows wild type zebrafish embryos. Right panels shows *touchtone* mutants, which possess a mutation in *trpm7* (Arduini and Henion, 2004). C) Left panel shows wild type zebrafish embryos. Right panels shows *lockjaw* mutants, which possess a mutation in *tfap2* (Knight et al., 2004).

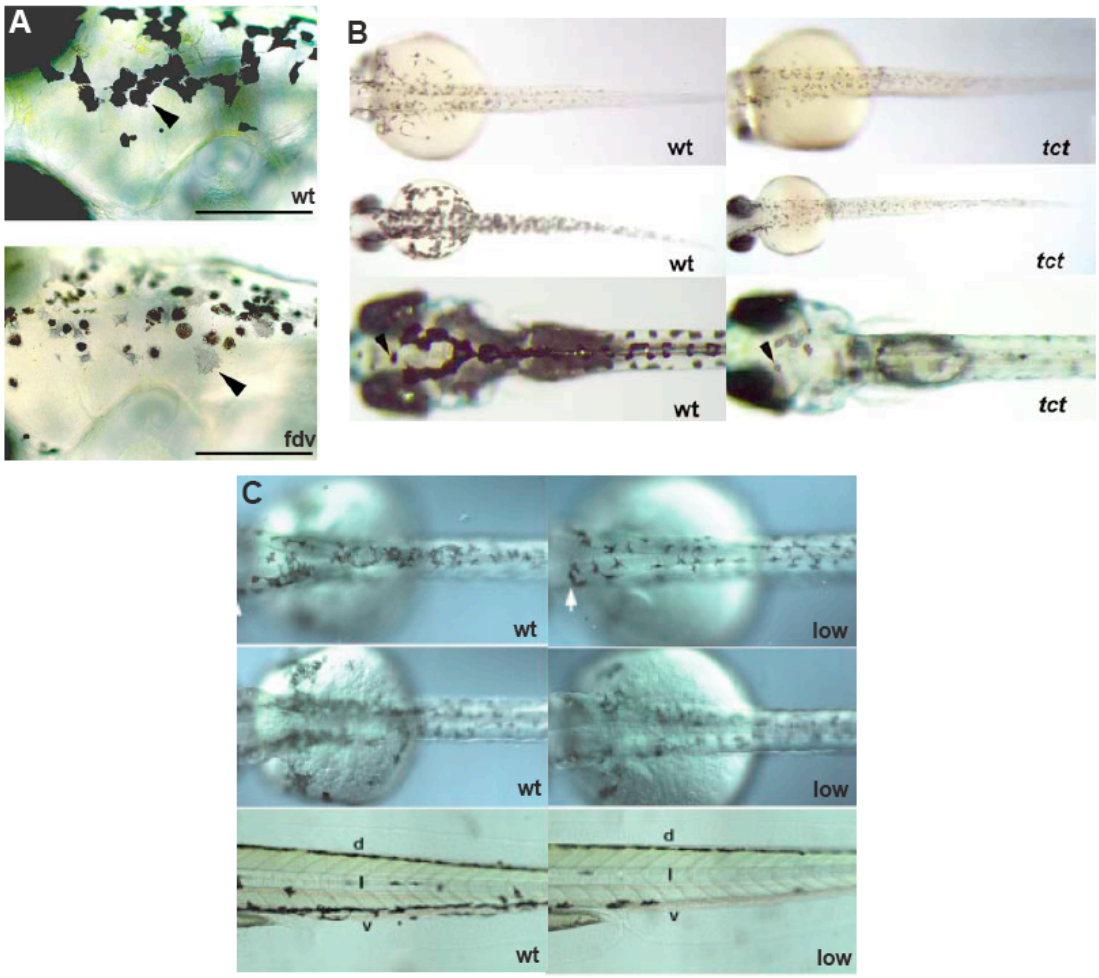
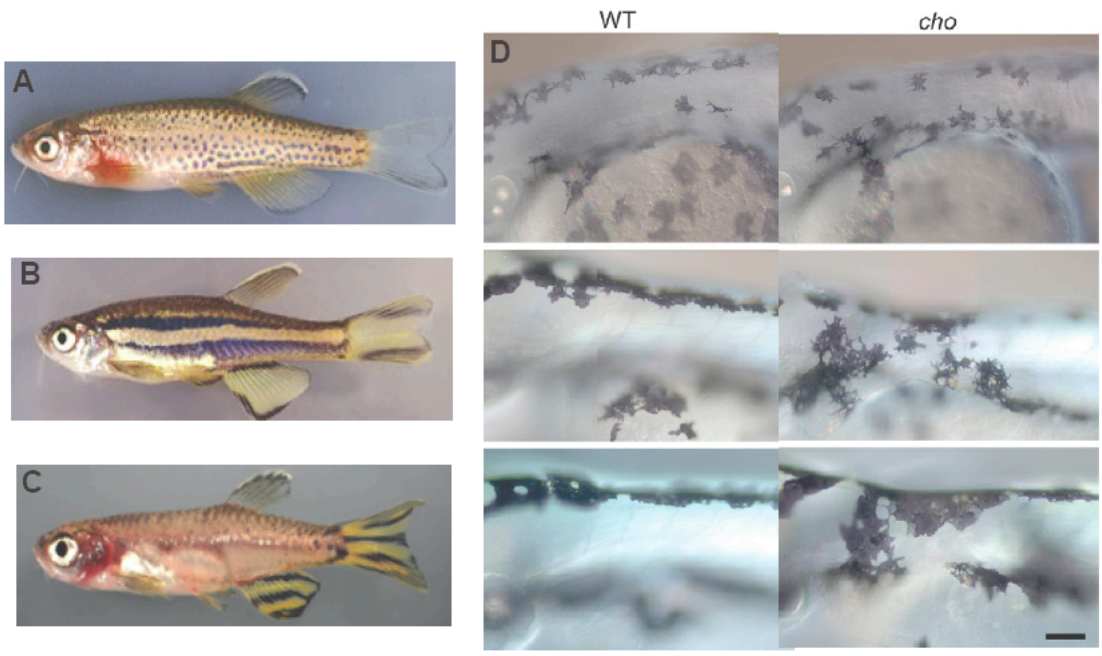


Figure 1-7. Pigment patterning zebrafish mutants. A) *leopard* mutant possessing a mutation in *connexin41.8* (Watanabe et al., 2006). B) *jaguar* mutant possessing a mutation in *kir7.1* (Iwashita et al., 2006). C) *bonaparte* mutant possessing a mutation in *basonuclin-2* (Patterson and Parichy, 2013). D) *choker* mutant possessing a mutation in *sdf1a*. Wild type embryos are on the left, while mutant embryos are on the right (Svetic et al., 2007).



Chapter 2: A Small Number of Genes Underlie Male Pigmentation Traits in Lake Malawi Cichlid Fishes

Claire T. O'Quin, Alexi C. Drilea, Reade B. Roberts, and Thomas D. Kocher

See Appendix 1 for all supplementary tables and figures referenced in this chapter.

This chapter is published as:

O'Quin CT, Drilea AC, Roberts RB, Kocher TD (2012) A small number of genes underlie male pigmentation traits in Lake Malawi Cichlid Fishes. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 318(3): 199-208.

Abstract

Pigmentation patterns are one of the most recognizable forms of phenotypic diversity and an important component of organismal fitness. While much progress has been made in understanding the genes controlling pigmentation in model systems, many questions remain about the genetic basis of pigment traits observed in nature. Lake Malawi cichlid fishes are known for their diversity of male pigmentation patterns, which have been shaped by sexual selection. To begin the process of identifying the genes underlying this diversity, we quantified the number of pigment cells on the body and fins of two species of the genus *Metriaclima* and their hybrids. We then used the Castle–Wright equation to estimate that differences in individual pigmentation traits between these species are controlled by one to four genes each. Different pigmentation traits are highly correlated in the F_2, suggesting shared developmental pathways and genetic pleiotropy. Melanophore and xanthophore traits fall on opposite ends of the first principal component axis of the F_2 phenotypes, suggesting a tradeoff during the development of these two pigment cell types.

Introduction

The cichlid fishes found in the three Great Lakes of East Africa represent one of the most stunning radiations of vertebrates on earth. Some 2,000 new species have evolved in lakes Malawi, Tanganyika, and Victoria over the last ten million years (Genner et al. 2007). The species flock in Lake Malawi is less than two million years old, but contains more than 500 species of cichlids known for their variety of behavior, trophic morphology, and color pattern (Turner et al. 2001; Kocher, 2004). It is thought that female preference for color pattern has been a driving force in the speciation of these fishes (Dominey, 1984; Takimoto, *et al.* 2000; Turner and Burrows, 1995). Behavioral assays have shown that visual cues alone are sufficient for proper mate choice among closely-related species (Seehausen and van Alphen, 1998; Kidd et al., 2006), and population genetics supports the role of pigmentation patterns in maintaining reproductive isolation in sympatry (van Oppen et al., 1998). Despite its importance to the radiation of cichlid species in Lake Malawi, surprisingly little work has been directed at understanding the genetic and cellular basis of cichlid pigment patterns.

Endotherms have just one pigment cell type (melanocytes), while ectotherms have at least three major cell types: melanophores, xanthophores, and iridophores (Mills and Patterson, 2009). Melanophores contain the black pigment melanin, xanthophores contain yellow pteridine- or carotenoid-derived pigments, and iridophores contain reflective purine-derived pigment (Mills and Patterson, 2009). Mutational analysis of zebrafish has identified several of the genes involved in the development of these pigment cells (Kelsh et al., 1996; Odenthal et al., 1996; Lister et

al., 1999; Parichy et al., 1999, 2000a,b, 2003b; Lopes et al., 2008). Genetic analysis has also been used to study pigment patterns in natural *Danio* populations (Quigley et al., 2005; Parichy, 2006b; Mills et al., 2007).

The diversity of pigment patterns in Lake Malawi cichlids provides an opportunity to investigate the genetic basis of animal pigment patterns generally. Closely related species of cichlid recapitulate the variety of patterning seen across vertebrates, including bars, stripes, and spots. Cichlids are easily raised and crossed in the lab, allowing genetic analysis of a wide range of colors and patterns arising from the expansion and patterning of pigment cell lineages. Here we use a hybrid cross to estimate the number of genetic factors underlying differences in male pigmentation between two closely related Lake Malawi cichlid species, *Metriaclima zebra* and *M. mbenjii.* These species are morphologically similar, but differ in coloration, including differences in body barring that may be a sufficient signal to maintain reproductive isolation of congeners in sympatry (Kidd et al., 2006). Previous genetic studies of cichlid pigmentation have analyzed images of the whole body (Barson et al., 2007; Magalhaes and Seehausen, 2010). We have taken a different approach by quantifying pigment cell numbers in individual fins and body regions, which gives us greater statistical power and makes it easier to develop hypotheses about the function of specific genes in pigment cell development.

The ultimate goals of this study are two-fold. First, we are interested in understanding the genetic basis of pigment patterns throughout the cichlid flock of Lake Malawi. Many of the pigment patterns observed in our parental species are repeated across several genera of Lake Malawi cichlids. How many genes cause these

pigment patterns? Are the same genes used to create the same patterns in different species and genera? Second, we want to estimate how many F_2 individuals will be needed to identify quantitative trait loci for each of the pigmentation traits. This information will be used to design future studies to map and identify the genes underlying variation in pigment patterns among these species.

Materials and Methods

Animals

We scored pigmentation traits for 20 male individuals of each parental species, 20 F_1 hybrid males, and 100 F_2 males. The progeny were the result of crossing a single *M. zebra* female to a single *M. mbenjii* male. One F_1 family resulting from this cross was raised together and intercrossed to produce the F_2 generation. A single F_1 male, mated with several sibling females, sired all the F_2 males used for this analysis.

The dominance hierarchy present in the tank can influence the expression of cichlid pigmentation. To ensure that we scored males in full breeding color, only the most dominant male was pulled from a tank each day. Males were placed individually in clear tanks and allowed to visually interact with other males during the several days of data collection. After data collection, the fish were euthanized with an overdose of MS-222, and the sex of the fish was confirmed via dissection.

Pigment Phenotypes

Malawi cichlids have three pigment cell types: black melanophores, silver iridophores, and a third type containing yellow to orange alcohol soluble pigment, which we have called xanthophores. The blue coloration in our fish is formed by

structural interference in iridophores. On the face and scales, iridophores are tightly associated with melanophores, with iridophores overlaying the melanophores. On the fin web, there are fields of iridophores, with melanophores interspersed throughout. All animal procedures were approved by the University of Maryland IACUC (Protocol no. R0771). Photographs of whole fish were taken with a Nikon D50 digital camera (Nikon,Melville, NY) in an enclosed box with reflected fluorescent light. Images of scales, fins, and cheeks were captured on a Leica MZ 16FA stereomicroscope (Leica Microsystems, Buffalo Grove, IL). To facilitate counting of individual chromatophores, the fish were treated with epinephrine (500 mg/L), which aggregates melanosomes. For analysis of melanophores on the scales, dorsal and pelvic fins, and xanthophores on all fins, the epinephrine was dissolved in phosphate buffered saline (pH 7.0), and the scales and fins were removed and immersed in the solution for 1 hr. For analysis of melanophores on the cheek and caudal fin, the epinephrine was dissolved in tank water, and the fish immersed in this solution approximately 5–10 min until the melanophores were contracted. Before imaging, fish were anesthetized with 40 mg/L MS-222 (Argent Chemical Labs, Redmond, WA).

Melanophore phenotypes were scored from the scales; the dorsal, caudal, and pelvic fins; and the cheek of each fish. One scale from the first full bar of the body and one scale from the space following that bar from both the left and right side of the body were removed from just above the lateral line. This could be accomplished also in unbarred fish due to the fact that they show faint bars when stressed. After treatment with epinephrine to contract the melanophores, scales were immediately

fixed in 10% buffered formalin. Each scale was photographed, and melanophores within a 0.25-mm2 area were counted manually in the program Image J (Rasband, 2009). To quantify melanophores in the dorsal fin, we counted 12 fin rays anterior of the last soft fin ray and removed the two fin webs immediately anterior of this dorsal spine. After epinephrine treatment, the fin web was photographed, and melanophores in a 0.25-mm2 area were counted. To count melanophores in the caudal fin, the fish were bathed in epinephrine solution. Once the melanophores were contracted, we took a photograph between the two branching points of the fin rays on the center caudal fin web and again counted melanophores in a 0.25-mm2 area. Melanophores in a region of the cheek approximately 1.5-mm below the eye were contracted with epinephrine and imaged under the microscope. A picture was taken and melanophores in a 0.25-mm2 area were counted. To score the melanophores on the pelvic fin, the fish were euthanized with MS-222, and the pelvic fins immediately removed and placed into epinephrine solution. A picture was taken of the fin web at the leading edge of the pelvic fin, and melanophores counted in a 0.25-mm2 area.

Xanthophore phenotypes were scored from the dorsal and caudal fins of each fish. Fish were anesthetized and the lower half of the caudal fin was amputated and placed in epinephrine solution. The third set of fin rays dorsal from the cut edge was photographed and a 0.32-mm2 area of the image was analyzed. A 1.0-mm2 region of the same photograph that was obtained to analyze the number of dorsal melanophores was also analyzed for xanthophores. Variation in morphology, coloration, and cell overlap confounded counting of individual xanthophores, so image analysis was performed using Photoshop (Adobe Systems, San Jose, CA). A CMYK (C = cyan; M

= magenta; Y = yellow; K = key) histogram was created and the mean yellow value was used as a proxy for the overall number of xanthophores present. The output for this program gives low scores for those images that possess more yellow. To make the results easier to interpret, with higher values corresponding to more yellow in the image, the yellow value was subtracted from 255 (the highest obtainable score). It should be noted that the yellow value is also affected by the number of melanophores and iridophores. The amount of yellow present on the gular (area of skin that extends forward from the gill covers and over the throat) was based on a qualitative scoring system, with one representing no yellow present, two representing the presence of some yellow, and three indicating the presence of intense yellow.

Bar spacing was scored from whole-fish images in Image J. A line was drawn along the length of the fish above the lateral line and below the insertion of the dorsal fin. The grayscale intensity of pixels along this line was captured with the plot profile command. The intensity data were imported into Excel and fit to a sine wave model using the Solver plug-in to determine the period of the sine wave. Bar spacing data were standardized by dividing the period by standard length.

Statistical Analysis

Data were first checked for adherence to normality. Data that did not display a normal distribution were either log10 transformed (bar, space and cheek melanophores, and bar spacing data) or square-root transformed (pelvic melanophores). Means and variances of each trait were calculated for each of the parentals, the F_1 and F_2. Two-tailed *t*-tests were performed to determine if parental means were significantly different at $P < 0.05$. The number of loci controlling

pigmentation traits was estimated using the Castle–Wright equation (Lynch and Walsh, 1998):

$$\hat{n}_e = \frac{[\bar{z}(P_1) - \bar{z}(P_2)]^2 - Var[\bar{z}(P_1)] - Var[\bar{z}(P_2)]}{8(Var(F_2) - Var(F_1))}$$

Correlation and principal component (PC) analyses were performed using the R Statistical Software (R Development Core Team, 2008). Significance of the correlations was determined by a Student's *t*-test with $P \leq 0.05$.

Dominance Calculations

Dominance calculations were performed using the mean phenotypic values for traits of the parental and F_1 generations. We based calculations on the assumption that the parental species were differentially fixed for alleles. Additive genotypic values (a), dominance genotypic values (d), and degree of dominance (d/a) were determined based on the following equations (Falconer and Mackay, '96; Cheverud, 2000).

$$a = \frac{(G_{22} - G_{11})}{2},\ d = G_{12} - \frac{(G_{11} + G_{22})}{2}$$

Calculations were always performed with G22 being assigned as the parent that possessed alleles that would increase the phenotypic trait. For example, for melanophore traits, *M. zebra* was assigned G22, since they possessed more melanophores, but for xanthophore traits, *M. mbenjii* was assigned G22 since they possessed more xanthophores.

Results

Overall Phenotype

Behaviorally dominant *M. mbenjii* males have a plain blue body, orange dorsal and caudal fins, and a silver-white leading edge on the pelvic fin (Fig. 2-1A).

Metriaclima zebra males have a blue body with black vertical bars, blue dorsal and caudal fins, dark cheeks, a black interorbital bar, and a black leading edge on their pelvic fin (Fig. 2-1B). The overall phenotype of the F_1 males was generally intermediate between the parental species. F_1 males had a blue body with faint black bars. The dorsal and caudal fins graded from blue toward the front of the fish to yellow toward the back of the fish. The pelvic fin had either a weak black streak or no black streak on the leading edge fin web (Fig. S2-1A). The black cheeks characteristic of *M. zebra* were not observed in any of the F_1 hybrids. Notably, several F_1 hybrids possessed a yellow gular that was not observed in either of the parents. The F_2 males displayed a range of phenotypes generally intermediate between the parental species (Fig. S2-1B).

Melanophores

Metriaclima zebra possessed significantly more melanophores on both bar and space scales compared to *M. mbenjii* (Table 2-1). The number of melanophores was significantly higher on bars than spaces in *M. zebra*, but not *M. mbenjii* (Table S2-1). The relatively uniform melanophore number across bars and spaces contributes to the "unbarred" body coloration observed in *M. mbenjii*. The observed difference in body coloration of the two species is the result of a lower overall number of melanophores in *M. mbenjii*, and the presence of additional melanophores in the bars of *M. zebra*. F_1 hybrids had a mean number of melanophores in their bars and spaces closer to that of *M. mbenjii* (Table 2-1; Fig. S2-2A and B). There was no significant difference in scale melanophore number between bars and spaces in F_1 hybrids (Table S2-1). F_2 hybrids displayed a range of phenotypes that encompassed the range of

melanophore counts observed on the scales in the parents (Fig. 2-2A and B). The Castle–Wright equation provides an estimate of 2.04 genes (SE = 3.90) controlling the number of scale melanophores for bar one and 0.26 genes (SE = 0.19) for the number of scale melanophores for space one.

Metriaclima zebra possessed a significantly greater mean number of melanophores than *M. mbenjii* males for all fins examined (Figs. 2-1C–H, 2-1C–E; Table 2-1). For the dorsal and pelvic fin, F_1 hybrid males had a mean melanophore number that fell closer to that of *M. mbenjii*, while the mean melanophore number for the caudal fin in F_1 hybrid males was approximately intermediate (Fig. S2-2C–E; Table 2-1). Melanophore numbers of F_2 hybrid males encompassed the range of the two parental species for all fins (Fig. 2-2C–E) and had a higher variance than the F_1. The Castle–Wright estimates for the different fins are as follows: 0.13 genetic factors (SE = 0.05) controlling the number of dorsal melanophores; −0.73 genetic factors (SE = 6.09) controlling the number of melanophores present on the caudal fin; and 2.56 genetic factors (SE = 6.05) controlling the number of melanophores present on the leading edge of the pelvic fin. The low (or negative) estimates of genetic factors suggest that one or more assumptions underlying the Castle–Wright equation are violated in this cross.

Metriaclima zebra males had a mean cheek melanophore number significantly greater than *M. mbenjii* males (Table 2-1). F_1 hybrids possessed a mean melanophore number intermediate to the parents, and the F_2 had a greater variance than the F_1 (Fig. 2-2F; Fig. S2-2F). The Castle–Wright equation gave an estimate 0.18 (SE = 0.14) genetic factors, an underestimate likely arising from the high variance in *M.*

zebra relative to the mean difference between the parental species.

Correlation analysis reveals that all melanophore traits examined are highly correlated with each other in the F_2 (Table 2-2). The highest correlation exists between scale melanophores on and off bars, with an r^2 value of 0.75. All other melanophore traits have significant correlations with each other, with r^2 values ranging from 0.59 to 0.44. This suggests genetic pleiotropy and common pathways involved the development of these traits. Dominance calculations based on the mean phenotypes for the parental species and F_1 hybrids suggests that the *M. mbenjii* alleles are dominant for all melanophore traits (Table S2-3).

Xanthophores

Image analysis revealed that *M. mbenjii* had a significantly greater yellow score for their dorsal and caudal fins when compared to *M. zebra* (Figs. 2-1C–F, 2-2G and H; Table 2-1). F_1 hybrid males had a mean yellow score that fell closer to that of *M. zebra* for both of these fin traits (Fig. S2-2G and H). Although the F_2 variance was greater than that of the F_1 in both fins, the high-end yellow values observed in the *M. mbenjii* males were not fully recovered in any of the F_2 (Fig. 2-2G and H). The Castle-Wright estimate suggests that the xanthophore pigmentation of the dorsal fin is controlled by a minimum of 0.99 genes (SE = 0.06), while the xanthophores pigmentation of the caudal fin is controlled by at least 3.23 genetic factors (SE = 24.24). Correlation analysis in the F_2 reveals these two traits are significantly correlated with an r^2 value of 0.61, once again suggesting a common developmental pathway controlling these traits (Table 2-2). Dominance calculations based on the mean phenotypes for the parental species and F_1 hybrids suggests that the *M. zebra*

alleles are dominant for all xanthophore traits (Table S2-3).

Xanthophores were also observed on two other locations on the body, the pelvic fin and the gular. A xanthophore streak present on the third pelvic fin web was observed in several individuals of the cross. Because it was not differentially fixed in the parents, we cannot estimate the number of genes controlling this trait, however, we note that the intensity of the streak varies in a way that suggests simple Mendelian inheritance. For the gular, the majority of specimens of both parental species scored a one (no yellow), with only six specimens receiving the score of two (some yellow). Of the 20 F_1 individuals, six individuals received a score of one, nine individuals received a score of two, and five individuals received a score of three (intense yellow). Among the 100 F_2 scored for this study, eight individuals received a score of two, while the remaining individuals received a score of one. These data indicate transgressive segregation in the F_1 and suggest that several genes contribute to this trait.

Bar Spacing

The bars of *M. zebra* are more closely spaced than in *M. mbenjii* males; however, the means of these two species were not significantly different (Table 2-1). The spacing in F_1 hybrid males was more similar to the *M. zebra* parent. F_2 hybrid males possessed a range of periods that encompassed the parental values, and the F_2 had a greater variance than the F_1, suggesting the segregation of genes underlying this trait. However, because the means of the parents were not significantly different, the Castle–Wright equation gives an estimate of −3.74 genetic factors for this trait.

Principal Components Analysis

Principal components (PC) analysis identified five components that each explained 5% or more of the variance. The first PC explained 40.1% of the variance. Dorsal fin xanthophore score and caudal fin xanthophores score had the highest positive loadings on the first PC, while the melanophore traits showed generally high negative loadings (Fig. 2-3; Table S2-2). PC2 explained 13.8% of the variance. Bar spacing had the highest positive loading scores, while caudal fin melanophores, caudal fin xanthophores, dorsal fin xanthophores, and cheek melanophores all had relatively high negative loading scores on axis 2 (Table S2-2). PC3 explained 11.9% of the variance. Dorsal, caudal, and pelvic fin xanthophores had the three highest negative loading scores on this axis. PC4 explained 9.3% of the variance, with dorsal fin xanthophores having the highest positive loading score and pelvic fin xanthophores and bar spacing having the highest negative loading score (Table S2-2). PC5 explained 7.2% of the variance. Pelvic xanthophores had the highest positive loading score (Table S2-2). The Castle–Wright estimates were 1.56 genetic factors (SE = 1.84) for PC1, 1.06 (SE = 13.50) for PC2, and 0.92 (SE = 19.92) for PC3 (Table S2-2). Castle–Wright estimates for PC4 and PC5 were near zero.

Discussion

Castle-Wright Estimates

Our results suggest that the male pigmentation traits we studied are controlled by a small number of genes. We estimate that the number of melanophores on the body and the fins are controlled by one to three genetic factors, while xanthophores on the fins are controlled by one to four genetic factors. Xanthophore and

melanophore traits fell at opposite ends of PC1. This may be indicative of a developmental tradeoff between melanophores and xanthophores. Castle–Wright estimates for the PC1 lead us to hypothesize that a minimum of two genes could be controlling a tradeoff between melanophore and xanthophore cell number. Considering the PC analysis as a whole, as few as four genes control over half the variance.

The low or negative Castle–Wright estimates for melanophore number between body bars, on the cheek, and on the caudal and dorsal fins are probably due to violations of the assumptions of the Castle–Wright estimator. The Castle–Wright equation assumes fixed differences between the parents, additive gene action, unlinked loci, and equality of allelic effects (Lynch and Walsh, 1998). The assumption most likely violated is that of additive gene action, especially since that dominance calculations indicate partial dominance for each trait examined (Table S2-3). However, unequal allelic effects and genetic linkage have also been shown to contribute to large biases of the estimator (Zeng, 1992). Due to probable violations of these assumptions, our estimates of the number of genetic factors controlling the traits studied should be considered lower bounds on the number of genes controlling each trait.

Our findings are consistent with previous genetic analyses that suggested that the pigmentation of male cichlids is controlled by a small number of genes. Barson et al. (2007) concluded that four to seven genetic factors control differences in body color between blue and gold morphs of the Lake Malawi cichlid *Pseudotropheus (Metriaclima) zebra*, with dominance and epistasis playing a role. Magalhaes and

Seehausen (2010) estimated that differences in red coloration of the Lake Victoria cichlids *Pundamilia pundamilia* and *P. nyererei* are caused by two to four genetic factors, again with dominance and epistasis playing a role. They also suggested that differences in yellow coloration are controlled by a single locus with complete dominance.

While the two previous studies estimate the number of genetic factors responsible for overall color differences from images of the whole fish, we believe it is important to also consider these traits in terms of their cellular basis. By quantifying specific cell types involved in pigmentation, we have broken down composite phenotypes, such as body color, into their discrete components. By measuring quantitative differences in cell number, we hope to make more informed hypotheses about possible candidate genes and pleiotropy among traits. This study also allowed us to estimate the number of F_2 offspring needed to detect quantitative trait loci for these phenotypes. We estimate that fewer than 200 F_2 will be needed to map the quantitative trait loci responsible for the differences in melanophore and xanthophore number between these species.

Pigmentation Modules

Careful study of the diversity of color patterns of Lake Malawi cichlids suggests a number of common modules that are used in different combinations among species. Allender et al. (2003) described four core color patterns that account for the coloration of 90% of rock-dwelling cichlids in Lake Malawi: all blue, all yellow, yellow chest, and yellow/orange dorsal fin. In addition, body bars are found on many species of rock-dwelling cichlids. These core traits are observed in different

combinations throughout the lake. For example, *M. pyrsonotus* possesses barring on the body similar to *M. zebra*, but has orange fins like those observed in *M. mbenjii*.

Our correlation analyses provide insight into the genetic structure of these modules. All of the melanophore traits are significantly correlated with each other. This is consistent with a common developmental pathway upregulating the number of melanophores throughout the body and fins. The strong positive correlation of xanthophore score in the dorsal and caudal fins is consistent with a common developmental pathway and suggests a mechanistic basis for the frequent correlation of these traits among species (Konings, 2007). Dominance calculations also suggest pleiotropy among traits, because *M. mbenjii* alleles are dominant for all of the melanophore traits, and *M. zebra* alleles are dominant for all of the xanthophore traits (Table S2-3).

An orange-blotch pigment morph observed in several genera of Lake Malawi cichlid females is due to a single *pax7* allele that is shared among several species across multiple genera (Roberts et al., 2009). Because of the recent origin of these species, similar pigment pattern elements may be due to retention of ancestral allelic polymorphisms, or sharing of alleles among species via hybridization. Indeed, rather than de novo mutation of novel color genes, reuse of a handful of pigment alleles may better explain the spectacular number of color transitions in the flock (Allender et al., 2003). By identifying the genes responsible for pigmentation differences between *M. zebra* and *M. mbenjii*, we may identify the genes controlling pigmentation modules across the radiation of cichlid species in Lake Malawi.

Cell Physiology

Changes in pigment cell physiology contribute to the appearance of the bars on the body and the dark cheeks on the face. Two main mechanisms are thought to control rapid color changes in teleosts. The first is a rapid response known as physiological color change, which involves the rapid movement of the pigment within the cells. We took advantage of this mechanism when we used epinephrine to cause melanophore contraction prior to counting them. The second form of color change is known as morphological color change. This form of color change involves the gradual reduction or increase of either pigment within the cell or the number of pigment cells themselves (Muske and Fernald, 1987).

While our results show a genetic basis for melanophore number on the scales and possibly the cheek, we cannot ignore the role that pigment cell physiology plays in pigment pattern displays in males. Indeed, subdominant or stressed *M. mbenjii* males display faint bars on the body. Also, our results indicate that there is no significant difference in the number of melanophores on the scales of the bars and spaces in the F_1 males; however, these males display faint bars on the body. This indicates that cellular response to hormones could play an important role in the pattern formation in these fishes. Melanophores of barred and barless males of *Astatotilipia* (*Haplochromis*) *burtoni* respond differently to the same concentration of norepinephrine, with barred males having expanded melanophores and barless males having contracted melanophores (Muske and Fernald, 1987). We hypothesize that a similar response could be occurring in our two study species, with *M. zebra* responding to a hormonal signal with melanophore expansion, intensifying dark bars,

and *M. mbenjii* responding with melanophore contraction, increasing overall light coloration.

Conclusion

This study advances our understanding of the number of genes that create the diversity of pigmentation phenotypes observed in East African cichlids. All studies thus far suggest that a small number of genes are responsible for the differences among species. More work is needed to identify the genetic loci responsible for these differences through genetic mapping.

Figures

Figure 2-1. Male parental phenotypes. (A) *Metriaclima mbenjii* adult male; (B) *M. zebra* adult male; (C) *M. mbenjii* dorsal fin; (D) *M. zebra* dorsal fin; (E) *M. mbenjii* tail fin; (F) *M. zebra* tail fin; (G) *M. mbenjii* pelvic fin; (H) *M. zebra* pelvic fin. Scale bars for A and B represent 1.0 cm. Scale bars for C–H represent 0.50 mm.

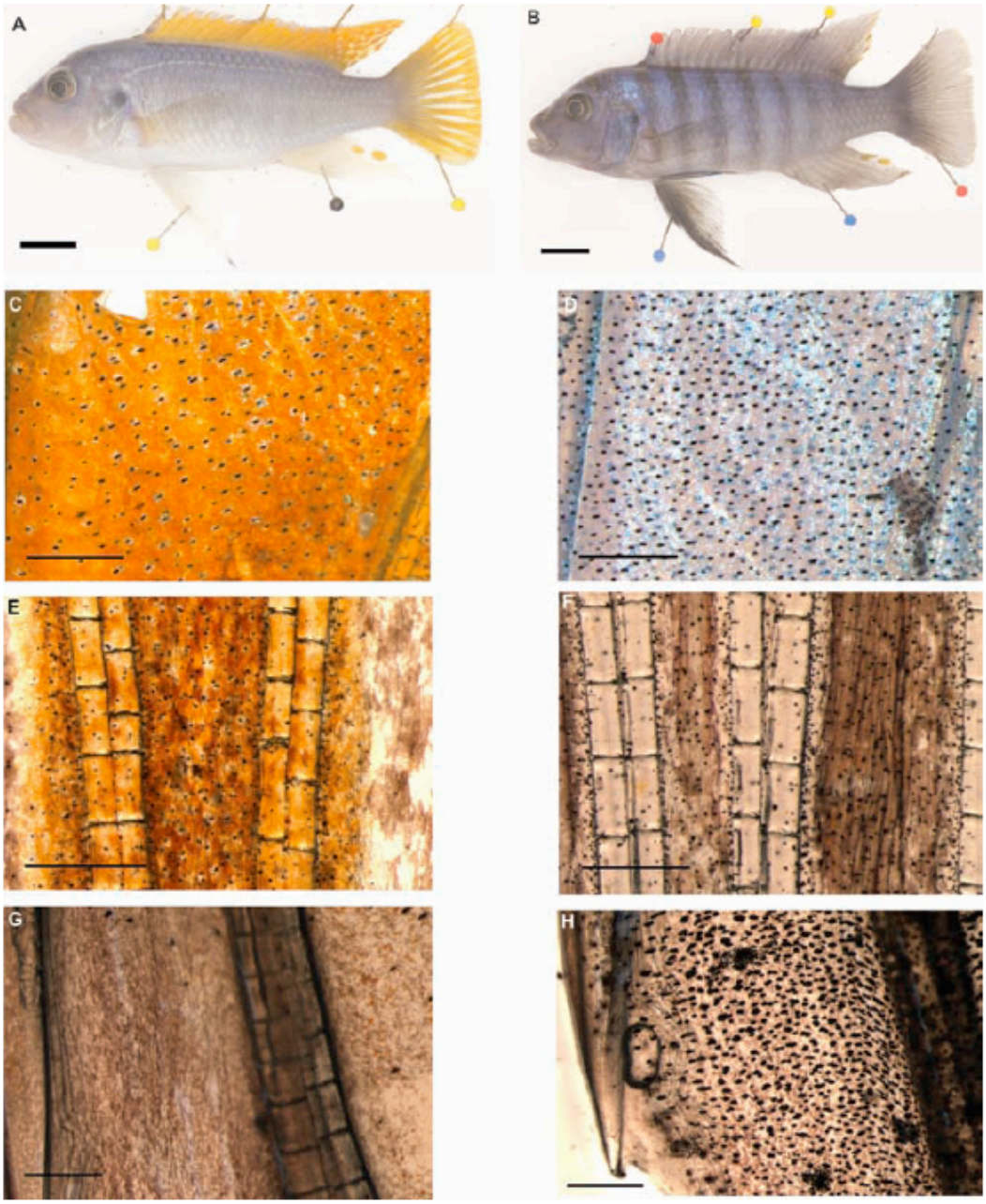


Figure 2-2. Boxplots of phenotypic data from the parental, F_1, and F_2 generations. The line within the box represents the median of the data, the areas of the box above and below the line represents the upper and lower quartiles, and the whiskers show the range of the data. (A) Bar scale melanophores; (B) space scale melanophores; (C) dorsal fin melanophores; (D) caudal fin melanophores; (E) pelvic fin melanophores; (F) cheek melanophores; (G) dorsal fin xanthophores; (H) caudal fin xanthophores.

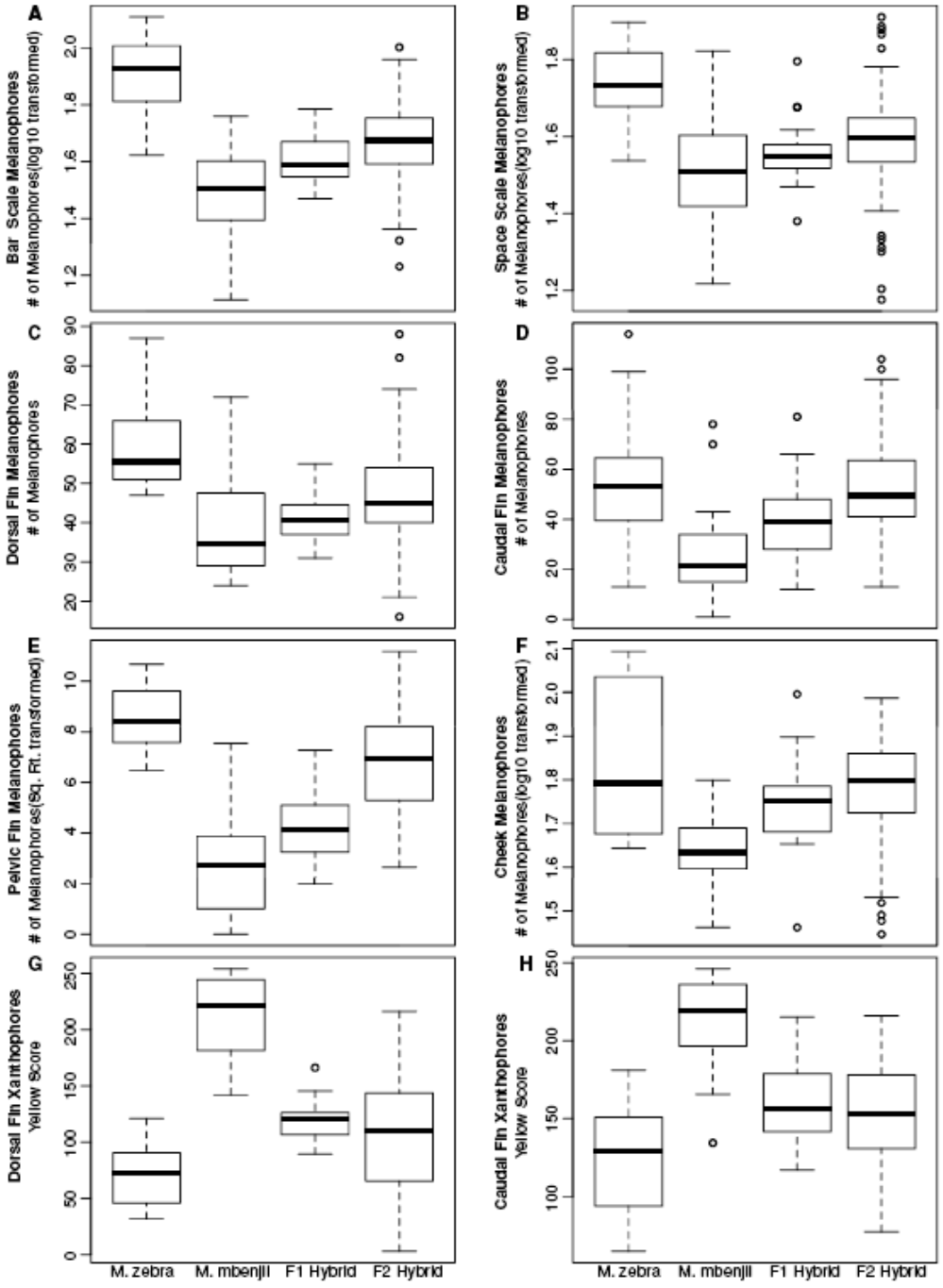
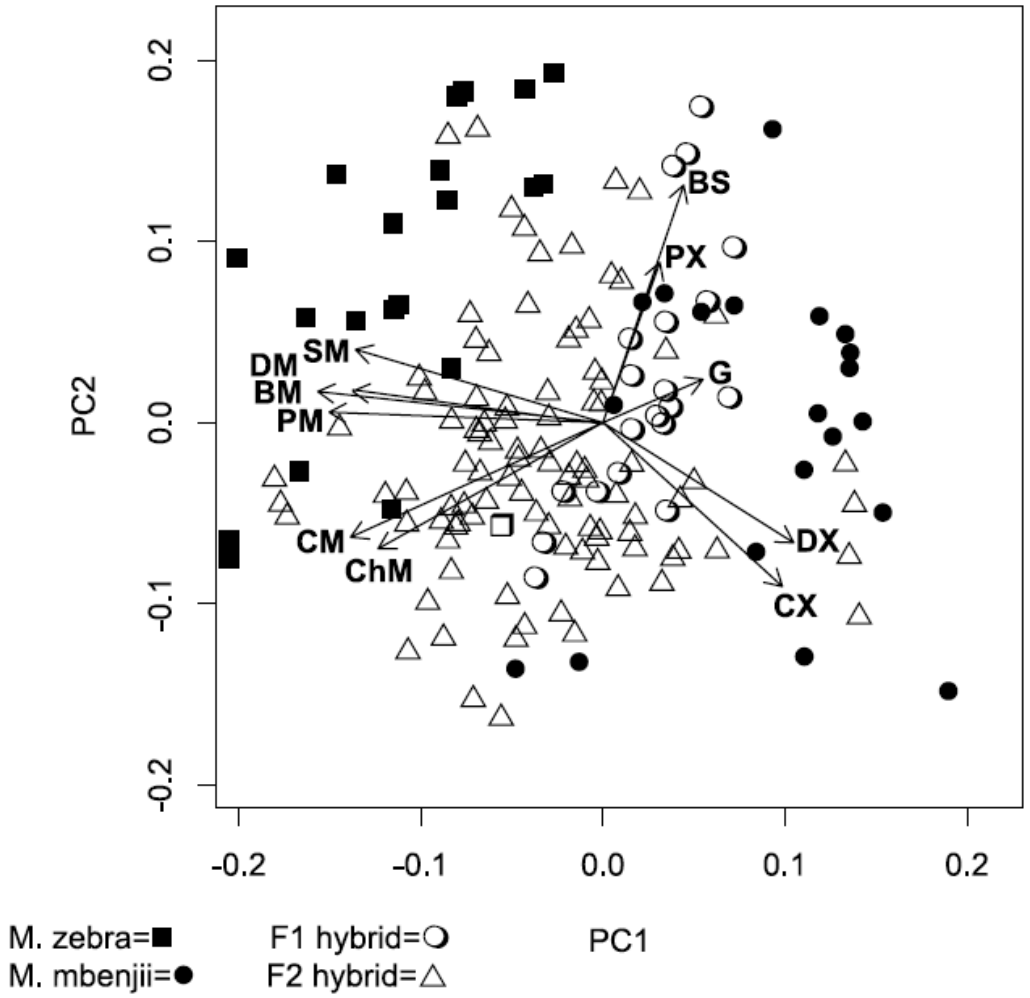


Figure 2-3. *K*-means plot of Principal Component 1 (PC1) and PC2. BM, bar melanophores; BS, bar spacing; ChM, cheek melanophores; CX, caudal xanthophores; DM, dorsal melanophores; DX, dorsal xanthophores; G, gular; PM, pelvic melanophores; PX, pelvic xanthophores; SM, space melanophores.



Tables

Table 2-1. Phenotypic means and variances (in parentheses)

Phenotype	M. zebra	M. mbenjii	p value	F_1 hybrid	F_2 hybrid	Castle-Wright
Bar Scale Melanophore	85.03 (665.30)	34.43 (163.43)	$7.90e^{-10}$	41.43 (73.93)	48.90 (196.30)	1.71
Space Scale Melanophore	56.50 (135.87)	35.40 (182.33)	$6.01e^{-06}$	36.98 (66.20)	40.56 (134.25)	0.26
Dorsal Fin Melanophore	60.80 (172.69)	38.90 (177.78)	$6.43e^{-06}$	41.25 (47.99)	46.82 (167.76)	0.13
Caudal Fin Melanophore	54.35 (621.08)	26.00 (401.89)	$3.31e^{-04}$	38.95 (275.10)	52.10 (312.54)	-0.73
Pelvic Fin Melanophore	75.80 (451.75)	10.80 (186.06)	$5.80e^{-13}$	21.15 (202.87)	51.34 (722.61)	2.56
Cheek Melanophore	74.25 (819.36)	44.60 (94.46)	$4.12e^{-05}$	56.55 (203.52)	62.47 (260.55)	0.18
Dorsal Fin Yellow Score	70.85 (760.16)	213.73 (1233.30)	$2.20e^{-16}$	118.94 (338.94)	106.87 (2661.23)	0.99
Caudal Fin Yellow Score	122.14 (1139.40)	213.07 (848.40)	$5.03e^{-11}$	161.98 (803.33)	152.70 (1046.20)	3.23
Bar Spacing-Period	25.98 (32.03)	24.14 (50.22)	0.35	22.01 (9.01)	19.85 (9.17)	-3.74

Table 2-2. Correlation analysis of F_2 phenotypic traits.

	Bar 1 Melano	Space 1 Melano	Dorsal Melano	Caudal Melano	Pelvic Melano	Cheek Melano	Dorsal Xantho	Caudal Xantho	Pelvic Xantho	Gular
Space 1 Melano	**0.75***									
Dorsal Melano	**0.59***	**0.51***								
Caudal Melano	**0.59***	**0.49***	**0.53***							
Pelvic Melano	**0.49***	**0.42***	**0.51***	**0.44***						
Cheek Melano	**0.58***	**0.40***	**0.37***	**0.52***	**0.52***					
Dorsal Xantho	-0.06	-0.09	-0.05	-0.07	**-0.20**	-0.06				
Caudal Xantho	**-0.25**	**-0.24**	**-0.31**	-0.06	**-0.21**	0.02	**0.61***			
Pelvic Xantho	-0.17	0.03	**-0.22**	**-0.23**	-0.02	**-0.22**	0.06	**0.20**		
Gular	-0.17	-0.06	**-0.21**	-0.02	**-0.22**	-0.14	0.04	0.05	0.10	
Bar Spacing	0.03	**0.22**	0.12	-0.06	0.10	-0.15	-0.07	-0.19	**0.29**	0.16

Correlations significant at $p<0.05$ are bolded. Correlations significant after Bonferonni correction are starred.

Chapter 3: Mapping of pigmentation QTL on an anchored genome assembly of the cichlid fish, *Metriaclima zebra*

Claire T. O'Quin, Alexi C. Drilea, Matthew A. Conte, and Thomas D. Kocher

See Appendix 2 for all supplementary tables and figures referenced in this chapter.

This chapter is published as:

O'Quin CT, Drilea AC, Conte MA, Kocher TD (2013) Mapping of pigmentation QTL on anchored genome assembly of the cichlid fish, *Metriaclima zebra*. *BMC Genomics* 14: 287.

Abstract

Pigmentation patterns are one of the most recognizable phenotypes across the animal kingdom. They play an important role in camouflage, communication, mate recognition and mate choice. Most progress on understanding the genetics of pigmentation has been achieved via mutational analysis, with relatively little work done to understand variation in natural populations. Pigment patterns vary dramatically among species of cichlid fish from Lake Malawi, and are thought to be important in speciation. In this study, we crossed two species, *Metriaclima zebra* and *M. mbenjii*, that differ in several aspects of their body and fin color. We genotyped 798 SNPs in 160 F_2 male individuals to construct a linkage map consisting of 834 markers in 22 linkage groups that spanned over 1,933 cM that was used to identify quantitative trait loci (QTL) associated with the pigmentation traits of interest. QTL analysis detected one QTL each for dorsal fin xanthophores, caudal fin xanthophores, and pelvic fin melanophores. Dorsal fin and caudal fin xanthophores share a QTL on LG12, while pelvic fin melanophores have a QTL on LG11. We also used the linkage map to anchor portions of the M. zebra genome assembly. We used the mapped markers to anchor 66.5% of the M. zebra genome assembly. Within each QTL interval we identified several candidate genes that might play a role in pigment cell development. This is one of a few studies to identify QTL for natural variation in fish pigmentation. The QTL intervals we identified did not contain any pigmentation genes previously identified by mutagenesis studies in other species. We expect that further work on these intervals will identify new genes involved in pigment cell development in natural populations.

Background

Most vertebrate species display a complex and species specific pigment pattern that enhances organismal fitness by contributing to crypsis, signaling, or mate recognition. Variation in pigmentation arises through differences in development (Parichy and Johnson, 2001; Parichy, 2007), nutrition (Paripatananont et al., 1999), and physiology (Nery and Castrucci, 1997). Many species use neural mechanisms to rapidly alter their pigment pattern in response to social cues (Muske and Fernald, 1987).

Early work to understand the genetic basis of variation in pigmentation focused on the analysis of mutant mice. Recently, fish have become an attractive model system due to their short generation times, large numbers of offspring, and the significant genetic resources availablefor some species. Zebrafish and medaka have been valuable models for identifying genes integral to pigment pattern formation via mutational analysis (Fukamachi et al., 2004; Lister et al., 1999; Parichy et al., 2000a; Lopes et al., 2008). Additional work has connected these genes to variation in pigment phenotypes among species of cyprinid fishes (Parichy and Johnson, 2001). Despite these advances, there is still much to be learned about the genetic basis of pigment patterning in natural populations.

Fishes display some of the most spectacular pigmentation observed in nature. Not only can a variety of colors be found, but also patterning including bars, stripes, spots, concentric rings, and blotches (Kelsh, 2004). These pigment patterns are formed by a diversity of pigment cells derived from the neural crest (Sauka-Spangler and Bronner-Fraser, 2008). While mammals and birds possess only the melanin-

containing melanocytes, fish have been found to have three basic pigment cell types: black melanophores, containing melanin, yellow to orange xanthophores, containing carotenoid or pteridine derived pigments, and highly reflective iridophores, which contain guanine platelets (Mills and Patterson, 2009). Additional cell types have been identified in some fish, including blue cyanophores, red erythrophores, and white leucophores (Mills and Patterson, 2009). Positioning of these cells in relation to each other produces the varied colors and patterns seen in nature.

The cichlid fishes of East Africa provide an excellent example of fish pigment pattern diversity and evolution. Endemic radiations of cichlids have arisen in each of the three major Rift Valley lakes (Malawi, Victoria, and Tanganyika). Of the three radiations, Lake Malawi is of particular interest because it is thought that most of the 500+ species of cichlids in the lake have arisen over the last two million years (Kocher, 2004). Diversification of pigment patterns dominates the most recent stage in the radiation, which has been driven by sexual selection (Danley and Kocher, 2001). Despite the importance of pigmentation to cichlid speciation, surprisingly little has been done to identify the genes associated with the diverse color patterns in these fishes (Barson et al., 2007; Magalhaes and Seehausen, 2010; Takahashi et al. 2013).

Because many of the species in Lake Malawi can be hybridized, it is possible to use a forward genetics approach to map genes underlying phenotypic diversity (Albertson et al., 1999). We previously analyzed an F_2 hybrid cross that suggested only a small number of genes underlie pigmentation differences between two Lake Malawi African cichlids, *Metriaclima zebra* and *M. mbenjii* (O'Quin et al., 2012). In the present study, we have identified several hundred single nucleotide

polymorphisms (SNPs) by sequencing restriction site associated DNA markers (RADSeq). We genotyped these markers to construct a linkage map for the hybrid cross and identify quantitative trait loci (QTL) underlying the pigmentation traits. Finally, we used the marker sequences to anchor the M. zebra genome sequence assembly to the linkage map, in order to identify candidate genes within each QTL interval.

Methods

Phenotypes

An F_2 hybrid cross was generated by crossing a single male *M. mbenjii* to a single female *M. zebra*. This resulted in a single F_1 family that had a single male intercrossed to sibling females to produce the F_2 offspring. While both male and female F_2 offspring were produced, only sexually mature, dominant male F_2 were analyzed. The two grandparent species differ in several aspects of male pigmentation. Male *M. mbenjii* have a light blue body with orange dorsal and caudal fins. Their pelvic fins are clear with an iridophore streak on the leading edge (Figure 3-1A). Male *M. zebra* have a light blue body with black bars and blue dorsal and caudal fins. Their pelvic fins have a melanophore streak on the leading edge. They also have dark cheeks and a black bar between the eyes (Figure 3-1B). While the xanthophores on the caudal fin of *M. mbenjii* are found in stellate form and overlap to form a uniform field, *M. zebra* often possess small punctate xanthophores on their caudal fins. The following quantitative phenotypes were collected from 160 F_2 male fish: # of melanophores on scales from the dark body bars, # of melanophores on scales from the spaces between the bars, # of dorsal fin melanophores, # of caudal fin

melanophores, # of pelvic fin melanophores, # of cheek melanophores, dorsal fin xanthophore area, caudal fin xanthophores area, pelvic fin xanthophores yellow score, gular xanthophores yellow score, and principal component one of a multivariate analysis of all ten traits. A detailed description of the measurement of these traits was published previously (O'Quin et al., 2012). The reader is referred to Additional file 1 found at http://www.biomedcentral.com/1471-2164/14/287/additional if they would like to access the phenotype data for all individuals used in the QTL study. All animal procedures were approved by the University of Maryland IACUC (Protocol no. R-10-73).

Genotypes

SNPs were identified and genotyped via restriction site associated DNA sequencing (RAD-seq) (Miller et al., 2007). Reduced representation DNA libraries were created using the protocol of Baird et al., 2011. Five libraries each containing multiplexed barcoded DNA for 32 individuals (160 F_2 progeny total), were sequenced in separate lanes of an Illumina HiSeq 1000. The F_0 grandparents were sequenced in an additional lane with 9 other individually barcoded samples. Reads were quality filtered by requiring Sanger quality scores of at least Q20 across 90% of the read. Reads were processed for individual barcodes and then assembled de novo into loci using the software pipeline Stacks v. 0.996 (Catchen et al., 2011). A minimum of 10 identical reads was required to create a "stack" in the parents. A minimum of 3 identical reads was required to create a stack in progeny individuals. One mismatch between loci was allowed when building the stacks "catalog". Resulting loci that were differentially fixed between the grandparents, in Hardy-Weinberg equilibrium,

and successfully genotyped in 100 or more F_2 individuals were chosen for creating the linkage map.

Additional microsatellite markers for the putative sex locus and selected candidate genes were added by selecting previously described microsatellite primer sequences from GenBank or by designing new primers to microsatellites identified in the *M. zebra* genome browser (www.bouillabase.org). Information on these markers can be found in Table S3-1.

Linkage map construction and QTL scan

A linkage map was created using JOINMAP 3.0 (Van Ooijen and Voorrips, 2001). The locus file consisted of genotypes at 798 SNPs and 37 microsatellites for 160 F_2 male progeny derived from a single F_1 family. The grouping module of JOINMAP assigned 834 of these markers to 22 linkage groups using a logarithm of odds (LOD) score of 5.0. We built a genetic map with the mapping module of JOINMAP, using the Kosambi mapping function, a recombination threshold of 0.450, and a jump threshold of 5.0. Linkage group numbers were assigned based on homology to the existing linkage groups of tilapia (Lee et al., 2005; Guyon et al., 2012).

QTL were detected using R/qtl (Broman et al., 2003; R Development Core Team, 2008). First, one thousand permutations were run using a Haley-Knott regression. The 5% significance level corresponded to an average LOD threshold of 4.1 for all phenotypes. We used a stepwise QTL detection algorithm that allowed for the detection of up to ten QTL for each phenotype, with the possibility of interactions. QTL intervals were further examined for significance by determining the Bayesian

credible interval. Genes in these credible intervals were compiled using the gene annotations found on the *M. zebra* genome browser (www.bouillabase.org). Candidate genes were identified via literature searches using the gene names.

Anchoring the M. zebra genome assembly

The locations of mapped markers on the M. zebra genome assembly version 0 were determined via BLAST. Assembly scaffolds were placed into anchored linkage groups if there were at least two markers from the same linkage group that blasted to that scaffold. The order and orientation of these scaffolds within each linkage group was then determined based on the BLAST location of markers relative to one another.

Results and Discussion

Rad-tag sequencing

A total of 743,486,491 reads were produced from the 5 lanes of Illumina HiSeq for the 160 F_2 progeny. 662,570,635 (89.1%) passed our Q20 filter. A total of 624,492,015 reads (94.3% of the filtered reads) were successfully processed for barcodes by Stacks. This corresponds to approximately 7.7 million reads per F_2 progeny. After filtering and barcode processing, there were 7,535,127 reads for the F_0 female and 18,850,602 reads for the F_0 male used for the cross.

Map construction and anchoring

We scored 834 genetic markers in 160 male F_2 progeny from the *M. zebra* x *M. mbenjii* cross. The average coverage of each genotype SNP was 49.9x (range of 6.9x-201.7x) in the F_2 progeny, 254.9x in the male F_0, and 105x in the female F_0. The average genotype completeness was 77% and the frequencies of each genotype class were 27.7% AA, 45.5% AB, and 26.9% BB (A designated for grandfather alleles and

B for grandmother alleles). 60 individuals had between 0–100 genotypes missing, 42 individuals had 101–200 genotypes missing, 25 individuals had between 201–300 genotypes missing, 22 individuals had between 301–400 genotypes missing, 7 individuals had between 401–550 genotypes missing, and 4 individuals had greater than 500 genotypes missing. The average number of missing genotypes per individual was 182. High numbers of missing genotypes can be attributed to low coverage in some of the F_2 individuals, with coverage ranging from 357,000 reads to 10,500,000 reads. We obtained a linkage map that contained 22 linkage groups and spanned over 1,933 cM. This agrees with previous work indicating that there are 22 chromosomes in cichlids (Guyon et al., 2012). Marker density was approximately one marker per 2.5 cM. Additional file 1 at http://www.biomedcentral.com/1471-2164/14/287/additional contains the information used to create the linkage map.

This linkage map was then used to order scaffolds of the *M. zebra* genome assembly. To be included in the anchored map, we required that scaffolds be anchored by at least two markers in the linkage map. We found 114 scaffolds (6.5 per linkage group) that met this criterion. The average size of these scaffolds was 3,918,467 bp for a total of 564,259,264 anchored bp. This represents 66.5% of the 848,776,495 bpM. zebra genome assembly. An additional 110 scaffolds had a single hit to a particular linkage group and were not included in the anchoring. If these single scaffolds were included in the anchoring, 92.3% of the assembly would become anchored. Additional file 3 at http://www.biomedcentral.com/1471-2164/14/287/additional provides the placement of scaffolds on the linkage groups.

QTL scan and detection

A genome wide scan resulted in the identification of three QTL. Phenotypes with significant LOD scores included dorsal fin xanthophores, caudal fin xanthophores, and pelvic fin melanophores (Figure 3-2). The failure to detect QTL for the remaining pigmentation traits is most likely due to the small difference in the parental means for the other pigmentation traits. QTL were detected for the three traits for which parental means were 2.8 or greater standard deviations apart. The mapping population of 160 male F_2 did not have enough power to detect QTL for traits for which the difference in parental means was smaller than 2.8 standard deviations.

Dorsal fin xanthophores

One QTL with a LOD score of 11.15 was detected for dorsal fin xanthophores on LG12 and explained 27.4% of the variance (Figure 3-3A). The Bayesian credible interval for this QTL spans from 62–73 cM along LG12. This matches the estimate from our previous work that a minimum of one QTL would be detected for this trait (O'Quin et al., 2012). The effect plot for the marker with the highest LOD score shows incomplete dominance. Individuals with more orange on their dorsal fin possess at least one allele from the *M. mbenjii* grandfather (Figure 4A). The results of the effect plot are consistent with the inheritance of orange fins from the *M. mbenjii* grandfather (O'Quin et al., 2012).

Using the annotated genome assembly for *M. zebra*, we were able to identify candidate genes within the credible interval. Little work on the genetics of xanthophores development and carotenoid formation has been done in fishes. *Csf1r*,

the most obvious candidate gene for xanthophores traits (Parichy et al., 2000), is not present in the interval. Several other genes that might be involved in pigment cell development are present, including genes involved in vesicle formation, carotenoid synthesis, and cell aggregation (Table S3-2). TRPM 6/1, a member of a gene family previously linked to pigment cell development in zebrafish, is located in the interval. However, this gene has only been demonstrated to be important in melanophore development. Fish possessing a mutation in this gene experience melanophore death, while the other pigment cell lineages (xanthophores and iridophores) appear to develop normally (Arduini and Henion, 2004; Iuga and Lerner, 2007). We do not consider TPRM6/1 a primary candidate gene for dorsal fin color since the difference between the grandparent fins appears to be a trade-off between the presence of xanthophores versus iridophores.

Caudal fin xanthophores

One QTL with a LOD score of 5.21 was identified for caudal fin xanthophores on LG 12 in a region that overlaps with, but is broader than, the QTL region identified for dorsal fin xanthophores. The Bayesian credible interval spans from 56–81 cM along LG 12 (Fig. 3-3B). Identification of a shared QTL is not surprising since our previous work indicated a strong correlation between these traits (O'Quin et al., 2012). The QTL plot for caudal fin xanthophores is broader, possibly because this trait shows less variance in the caudal fin than it does in the dorsal fin. While the xanthophores in the caudal fin of *M. mbenjji* are found in stellate form and overlap to form a uniform field, *M. zebra* often possess small punctate xanthophores on their caudal fins.

Although the QTL region for dorsal fin and caudal fin xanthophores is shared, the percent variance explained for each trait is different, with 27.4% of the variance explained for the dorsal fin xanthophores, but only 14% of the variance explained for the caudal fin xanthophores. This difference could be due to the number of genes predicted to control each of these traits. While the Castle-Wright estimator predicted that one gene controls dorsal xanthophores, a minimum of three genes was predicted to control caudal fin xanthophores (O'Quin et al., 2012). We were probably only able to detect one gene for caudal fin xanthophores due to the power limitations discussed previously.

Similar to the effect plot for dorsal xanthophores, individuals with more orange on their caudal fin possess at least one of the *M. mbenjii* grandfather's alleles (Figure 3-4B). It should be noted that while the trait appears to show overdominance, the mean for the individuals homozygous for the *M. mbenjii* alleles and the mean for heterozygotes are not significantly different. Candidate genes considered for this region are the same as those considered for the dorsal fin xanthophores.

Pelvic fin melanophores

One QTL with a LOD score of 5.01 was identified for pelvic fin melanophores, with a Bayesian credible interval spanning from 68–82 cM on LG 11 (Figure 3-3C). This QTL explains 13.4% of the variance for this trait. We previously estimated that this trait was controlled by a minimum of three genes. The relatively low power of our mapping population may explain why only one QTL was found for this trait. The jagged QTL curve for this trait may be due to a combination of high marker density and occasional missing genotypes.

Effect plots show that individuals with more melanophores on their pelvic fin are homozygous for the *M. zebra*'s grandmaternal alleles (Figure 3-4C). This is consistent with the fact that *M. zebra* males possess more melanophores on their pelvic fin compared to *M. mbenjii* males (O'Quin et al., 2012). No obvious candidate genes were identified, however various genes in the interval play a role in human skin disease, interact with other genes involved in melanophore development, or are involved in vesicle formation, packaging, and trafficking (Table S3-2).

Conclusions

Our study confirms that RADSeq is an effective method for rapidly identifying SNPs and genotyping hybrid crosses. We were able to obtain a high density of markers throughout the genome, with one marker approximately every 2.5 cM. Using these makers, we were able to create a linkage map, and subsequently anchor 66.5% of the current *M. zebra* genome assembly. We were also able to identify QTL regions for three of the eleven pigmentation traits studied. One region on LG12 contained a shared QTL for dorsal and caudal fin xanthophores. A second region on LG11 contains a QTL for pelvic fin melanophores.

The number of QTL identified and the percent variance explained appears to be consistent with our previous work. For dorsal fin xanthophores, we predicted we would identify one gene controlling this trait. This trait was the one for which we identified the highest LOD score, had the highest percent variance explained, and had the narrowest QTL peak. The other pigment traits were predicted to be controlled by multiple genes. Not surprisingly, these traits had lower LOD scores, explained a smaller portion of the variance, and had broader QTL peaks. Despite our high marker

density, the size of the mapping population limited our ability to narrow our QTL regions to less than 11 cM.

Finally, analysis of the predicted genes within our intervals showed several genes that could play a role in the development of pigmentation. None of them correspond to well-known zebrafish genes previously known to play a role in pigment cell development. Thus, these QTL represent an opportunity to learn something new about the genes underlying variation in pigmentation among fishes. We are particularly excited to have identified a major QTL contributing to xanthophores development, about which so little is known.

Figures

Figure 3-1. F_0 parents of the hybrid cross. A) *M. mbenjii* male. Note the orange dorsal and caudal fins, plain blue body, and white pelvic fin. B) *M. zebra* male. Note the blue dorsal and caudal fins, barred body, and black pelvic fin.

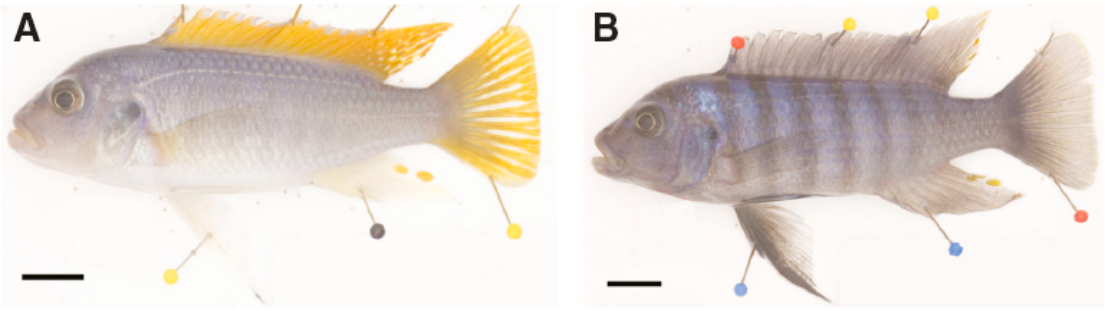


Figure 3-2. Genome wide distribution of LOD scores for each phenotype examined. Most traits fell below the 5% LOD significance threshold of 4.1, with the exception of pelvic fin melanophores, caudal fin xanthophores, and dorsal fin xanthophores.

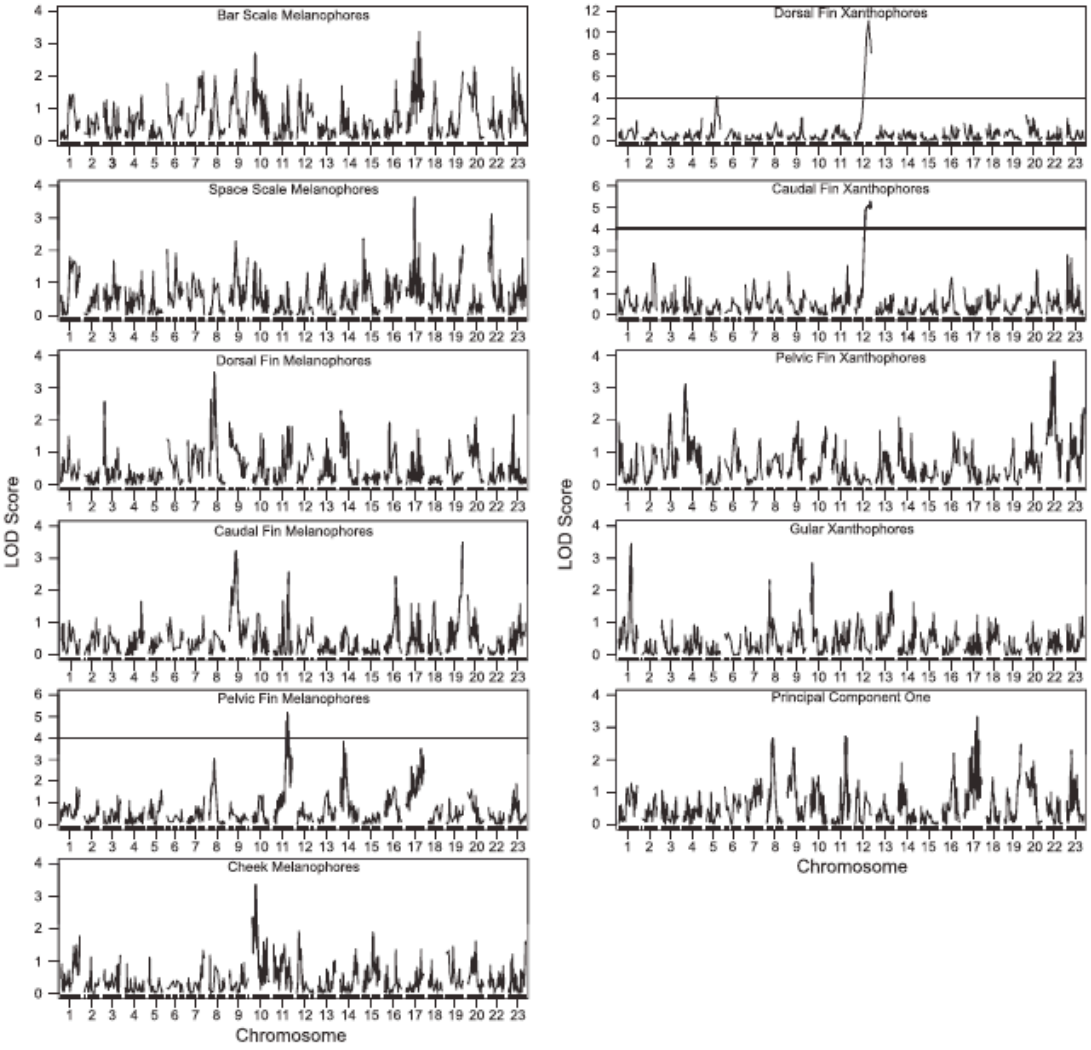


Figure 3-3. QTL plots for each trait that exceeded the significant LOD threshold. Shaded area indicates the Bayesian credible interval. The colored bars on the x-axis represent different genomic scaffolds. A) Dorsal fin xanthophores on LG12. B) Caudal fin xanthophores on LG12. C) Pelvic fin melanophores on LG11.

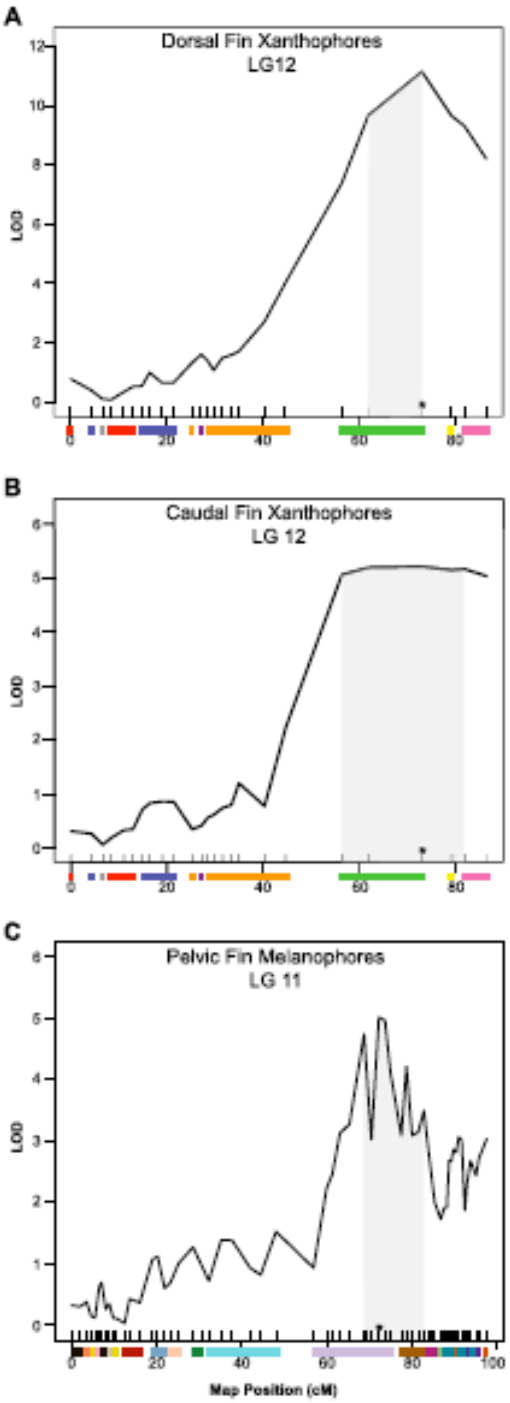
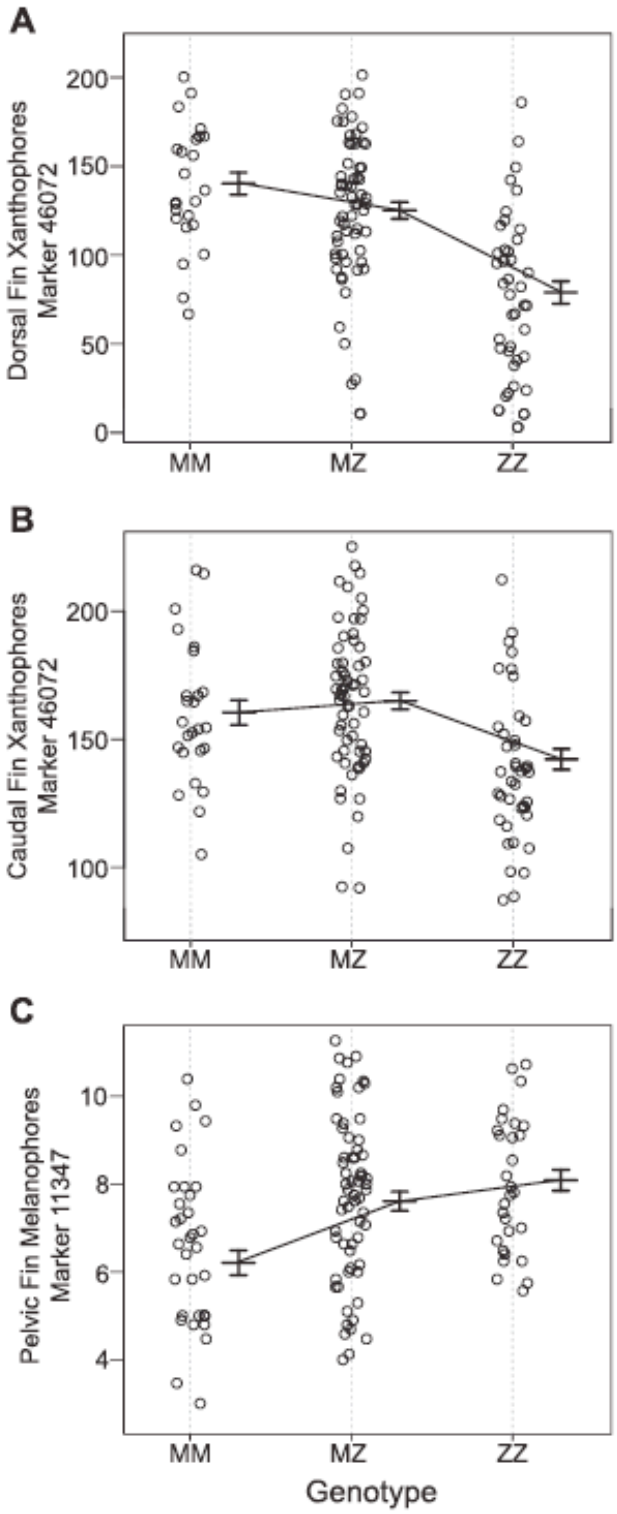


Figure 3-4. Effect plots for each trait at the marker with the highest LOD score. For each plot, "M" represents *M. mbenjii* alleles and "Z" represents *M. zebra* alleles. A) Dorsal fin xanthophores B) Caudal fin xanthophores C) Pelvic fin melanophores.



Chapter 4: Preliminary Analysis of a Pigmentation Candidate Gene

Claire O'Quin

See Appendix 3 for supplementary table referenced in this chapter.

Abstract

Elucidating the connection between genotype and phenotype has been accomplished using a variety of techniques, including mutational analyses, genome wide association studies, and quantitative trait locus (QTL) studies. While originally used in agriculture, QTL studies have proven to be a useful technique in identifying genomic regions associated with phenotypes in model and non-model species. Previously, we identified a QTL region for dorsal and caudal fin xanthophores in a hybrid cross of two cichlid species, *M. zebra* and *M. mbenjii*. Here, we determine the time point at which we hypothesize a candidate gene would be acting, and identify and begin to analyze the sequence and structure of the candidate gene, AAK1, for this trait. However, we were only able to preliminarily analyze the candidate gene, AAK1, due to the presence of gaps found in the promoter and first intron of this gene. Further work is still needed to determine if this gene plays a role in causing differences in fin coloration.

Introduction

Understanding the connection between genotype and phenotype has long been a topic of interest since the discovery of DNA. However, finding these connections is not always easily accomplished, and a variety of techniques have been used in an effort to identify genes underlying different traits of interest. Studies using mutants have been extremely valuable in identifying gene function; however, some of these mutants display dramatic phenotypes not found in the wild, and in some cases are lethal early during development. Genome wide association studies in humans have been helpful to identify susceptibility loci for disease; however, these loci often end up explaining only a small percentage of the variation observed for those traits, pointing to the complexity of human disease (Manolio et al., 2009).

Another method being utilized is quantitative trait locus studies. Originally used in agricultural genetics, this technique is now being used to identify genomic regions associated with phenotypes in both model and non-model species (Lynch and Walsh, 1998). While similar to genome wide association studies, QTL studies utilize large families where genomic linkage is greater, thus making it easier to map traits to particular genomic regions. QTL studies have been used to map a number of traits including pigmentation (Hoekstra, 2006).

Often one of the first things we notice about an organism is its color or the unique pigment patterning found on the body. These colors and patterns can play an important role in the survival of the organism. Endotherms only possess one pigmentation cell, the black to brown melanocyte (Mills and Patterson, 2009). Color variation in these organisms is achieved by variation in the color or the amount of

melanin deposited (Kelsh, 2004). Ectotherms possess three major pigment cell types, the black melanophore, the orange/yellow xanthophores, and the silver iridescent iridophore (Mills and Patterson, 2009; Parichy et al., 2006a). Of all the ectotherms, fish display one of the largest ranges of color and pattern combinations. Understanding the genetics underlying these patterns has been a topic of recent study.

Most of what we know about pigmentation in fishes comes from mutational studies in zebrafish and medaka (Kelsh et al., 2009; Kelsh et al., 2004). We have learned a great deal about the genes involved with specification of each of the three main pigment cell types and have begun to link these genes to natural variation in pigmentation patterns. However, a lot still needs to be learned about how genetic variation produces phenotypes observed in natural populations.

Using QTL methods, we previously identified a shared QTL region for dorsal fin xanthophores and caudal fin xanthophores in a hybrid cross between different species of African cichlids, *M. zebra* and *M. mbenjii* (O'Quin et al., 2013). These two species differ in several aspect of their male color pattern. Here we describe experiments aimed at determining the underlying genetic basis for this phenotype. To do this, we first identified the developmental time point when the gene controlling this trait might be active. We then confirmed the QTL peak with microsatellites and additional F_2 individuals. Next, we identified a candidate gene, AAK1, for this QTL and began a preliminary analysis of this gene via sequencing. Finally, we began to investigate the structure of AAK1.

Materials and Methods

Developmental Time Point Imaging

Broods of *M. zebra* and *M. mbenjii* were collected at 3dpf and raised together. For each brood, individuals were selected and isolated and raised individually in a tank. These individuals were photographed daily for the first month of life, and then weekly up to three months of age. If an individual died during a photographic session, a new individual from the same brood was selected to replace it. It should be noted here that we were not able to determine the sex of the individuals; so multiple individuals from multiple broods were followed to gain an overall idea of the pigment patterning process. Before hatching, fish were placed in a 3% methylcellulose solution to allow for better positioning for photography. After hatching, fish were first anesthetized with MS-222 (40 mg/L) before being placed in the methylcellulose. Fish were imaged using a Leica MZ 16FA stereomicroscope. Fish were then placed back in a dish of fish system water in order to recover before being returned to their tank.

Microsatellites and Single Marker Analysis

Seven microsatellites that were differentially fixed in the grandparents of the hybrid cross were identified and genotyped in the grandparents and 219 male F_2 offspring and 160 F_2 female offspring. These included previously developed microsatellites and new microsatellites identified using the *M. zebra* genome assembly (Table S4-1). PCR amplification was performed on DNA extracted from caudal fin clips. The size of the amplified microsatellite produced was determined on an ABI 3730 DNA sequencer using GeneMapper.

We compared mean phenotypes among the different genotypic classes at each marker using ANOVA and assessed significance. Sex was coded as a covariate. Due to the high level of significance across the different markers, the F value was used to plot the genomic location of the QTL.

Population Samples

20 individuals of each sex were sampled from the field for *M. zebra* at Mazinzi Reef and 11 male individuals and 17 female individuals for *M. mbenjii* at Mbenjii Island. Individuals were dissected and sex was confirmed via inspection of the gonads. Fin clips were preserved in DMSO-salt buffer (Seutin et al., 1991). DNA was extracted from fin clips using standard phenol-chloroform/EtOH precipitation. DNA was quantified by fluorescence with Picogreen (Life Technologies, Grand Island NY).

PCR and Sequencing

Due to the presence of genome assembly gaps within the vicinity of the AAK1 gene, sequencing was undertaken to fill in these gaps. Using the *M. zebra* genome assembly (www.bouillabase.org), primers were designed to amplify the gap regions of interest in the grandparents of the cross. For larger gaps, additional primers were designed as sequence was obtained across the gap. PCR products were gel purified for sequencing. Sequences were aligned and polymorphisms identified using Sequencher. Primers sequences can be found in Table S4-1. Gaps were named based on the gap numbers assigned to them in the genome assembly.

PCR Mismatch discrimination

The SNP identified between the grandparents for Gap 259 was genotyped via PCR mismatch discrimination using techniques previously described by Dieffenbach and Dveksler (1995). Primers (Table S1) were first designed to amplify an approximately 500 base pair region where the SNP was located. PCR reactions were allowed to run for 20 cycles at 55°C. This PCR product was then used as the DNA template for a second round of PCR where the two separate primer pairs were run for each individual. These primer pairs were used to discriminate between the SNPs. The grandfather's allele produced a fragment approximately 300 bp in size and the grandmother's allele produced a fragment approximately 150 bp in size. The second round PCR was run for 20 cycles at 60°C and the fragment sizes were visualized on a 1% agarose gel.

Cloning

We followed the cloning protocol for the pGem-T easy vector systems kit (Promega, Madison WI).

RNA extraction and cDNA creation

RNA was extracted from the dorsal and caudal fins of pooled individuals from the same family using an RNeasy kit (Qiagen, Valencia CA). cDNA was created by reverse transcription using the SuperScript III Reverse Transcriptase kit (Invitrogen, Grand Island NY).

Results

Developmental Time Points

In order to perform future functional analyses, we decided to determine the developmental time point where differences in dorsal fin color are clearly established in *M. zebra* individuals versus *M. mbenjii* individuals. At one month of age, the dorsal fin of both species looks the same, with all three major pigment cell types present (Fig. 4-1A). By two months of age, the fins of *M. mbenjii* individuals begins to fill in with xanthophores, while the number of iridophores present on the *M. zebra* fins begins to increase (Fig. 4-1B). By three months of age, the fins are clearly differentiated; with *M. mbenjii* dorsal fins possessing a field of orange xanthophores with black melanophores dispersed throughout, and *M. zebra* possessing dorsal fins with a field of iridophores with black melanophores dispersed throughout (Fig. 4-1C).

Single Marker Analysis

In order to increase power in confirming the previously identified QTL curve for dorsal xanthophores (O'Quin et al., 2013), more male individuals were analyzed, in addition to a set of female F_2 individuals. It should be noted that female indivuals also varied in the amount of orange on their fins. Six markers on Scaffold 9 and one marker on Scaffold 128 were used (Table 4-1). Sex was included as a covariate and was found to be significantly associated with dorsal xanthophores at each marker. Marker MZ8983 at 4.1Mb on Scaffold 9 had the highest level of association (Fig. 4-2). Investigation of this region of the *M. zebra* genome assembly revealed that this marker was directly upstream of AP2-associated protein kinase 1-like (AAK1). Due to this high association and the possible role this gene could play in pigmentation

organelle biogenesis, we focused our analyses on AAK1 as a primary candidate gene. We decided to focus our gene analysis first on the upstream region and in the first intron since these are locations where regulatory sequences controlling gene expression are typically found.

AAK1 Gaps

When we went to sequence the areas of interest in AAK1, we discovered three gaps in the genome assembly (Fig. 4-3). The first gap, Gap 259, is just upstream of AAK1 and was predicted to be 338 base pairs long. The other two gaps, Gap 260 and Gap 261, are within the first intron of AAK1 and are predicted to be 1,142 base pairs and 3,902 base pairs long. We decided to sequence these gaps and look for polymorphisms between the grandparents of the cross.

Gap 259

Sequencing of Gap 259 was undertaken first since it was predicted to be the smallest of the three gaps. Sequencing results indicated that the gap was 735 base pairs long. Comparison of the grandmother and grandfather sequences revealed a SNP at base position 351 in the gap. The grandmother possessed a "C" at this location, while the grandfather possessed a "T". I then genotyped this SNP, using PCR mismatch, in natural population samples of male and female *M. zebra* and *M. mbenjii* to see if the SNP segregated between these populations. Using this method, the grandmother allele produced a band around 150 bp, while the grandfather allele produced a band around 300 bp (Fig. 4-4A) An example of the mismatch gel in the offspring can be seen in Figure 4-4B. Table 4-2 shows the distributions for these

alleles in both sexes in these two species. While the SNP appeared to segregate randomly in the *M. zebra*, all but two of the *M. mbenjii* samples had a T/T genotype.

Gap 260

Sequencing of Gap 260 by conventional methods did not work due to the presence of highly repetitive sequence. Therefore, we resorted to cloning in order to sequence across the gap. Starting from the 5' sequence ending right before the gap, 19 bases of sequences were obtained before hitting a string of "A's" (with length of this string varying across the clones). After this string of "A's", we obtained another 1,129 base pairs up onto the sequence found after the gap. We identified two potential SNPs, however, these would need to be confirmed via resquencing.

Gap 261

Sequencing of Gap 261 required a few different techniques. Our original method used existing sequence before and after the gap in order to design PCR primers to amplify and sequence the DNA. This produced a single band for each grandparent, with the grandmother having a larger band than the grandfather. Once we obtained sequence further into the gap, we continued to design new primer pairs to continue a stepwise sequencing process across the gap. However, we eventually ran into issues where we would get multiple bands from our PCR product.

Concerned that we might be amplifying multiple regions in the genome, we decided to use a new approach. The original PCR product where we amplified across the entire gap was gel purified and used as the template DNA for all subsequent reactions. Once again, we eventually ran into issues where we started to obtain multiple bands. We finally decided to clone a DNA segment from the gap in an effort

to get better sequence results. We cloned from a piece of DNA that was amplified from an internal primer in the gap and a primer located outside of the gap that produced a single band.

By using this combination of methods, we obtained close to 2,300 bases of sequence working from the 5' end of the gap towards the 3' end. However, this sequence did not come without problems. First, we obtained 1,886 of sequence that consistently aligned between grandparents. Within this sequence, we identified a few sites where one grandparent was heterozygous and the other was homozygous or the grandparents had a fixed difference. After this point, we started to obtain disparate sequence for the grandmother between the cloned sequence and the sequence we obtained from my stepwise sequencing method. When sequencing from the 3' end towards the gap, we were only able to obtain sequence from the fragments cloned in the grandmother. On this side of the gap, we obtained 670 bases.

In an effort to determine why it was difficult to sequence this gap, I performed an alignment of the *M. zebra* sequence to the genome of three other cichlid species that did not possess this gap in the first intron: *Oreochromis niloticus*, *Astatotilapia burtoni*, and *Neolamprologus brichardi*. Before this gap, the *M. zebra* sequence aligns well to *A. burtoni* and *N. brichardi*. The *O. niloticus* sequence was divergent compared to these three sequences. Within the gap, *M. zebra* only matched the other two genomes for 23 bases before it completely diverged. Also, based on the alignment of sequence from *M. zebra* before Gap261 and after Gap 261 to the other species without the gap, the size of Gap 261 was predicted to only be 261 bases. This is much smaller than the amount of sequence we obtained.

AAK1 cDNA and Gene Structure

In order to start analyzing the sequence of AAK1, we decided to first work with cDNA due to the large number of exons present in the gene. When we first started working with the gene, there was only one gene prediction available, one for *O. niloticus*. We first attempted to create cDNA from the fins of adult *M. zebra* and *M. mbenjii* male individuals. To determine the presence of the cDNA we designed a set of primers that would amplify from Exon 1 to Exon 6, Exon 7 to Exon 12, and Exon 13 to Exon 19. The fragment sizes predicted to be amplified in both genomic and cDNA can be found in Table 3. While my β-actin controls showed that we had successfully created cDNA (Fig. 4-5A), it appears that AAK1 was either not present at the adult time point or my primers were not amplifying my DNA segments of interest (Fig. 4-5B).

To work around these two issues, we first created new primer sets to amplify smaller segments of the gene. We designed sets to amplify from Exon 2 to Exon 6, Exon 6 to Exon 11, Exon 12 to Exon 15, and Exon 16 to Exon 19. Once again, these fragment sizes for both genomic and cDNA can be found in Table 3. We then created some new sets of cDNA from the dorsal fins of one month, two month, and three month old fish, since this is when I determined gene expression might be important. We finally got good amplification at the correct sizes, with the exception of the fragment for Exon 16 to Exon 19, in one and two month old fish (Fig. 4-6A&B). Amplification was weaker in 3 month old fish (Fig. 4-6B). We also appeared to get amplification in adults for Exon 6 to Exon 11 and Exon 12 to Exon 15 (Fig. 4-6C).

Once we finally confirmed that we had created cDNA, we decided to investigate the structure of AAK1. At this point in our study, there were three different predictions for the structure of AAK1 based on NCBI Refseq (Fig. 4-7; Table 4-4). The prediction for *O. niloticus* has one predicted structure consisting of nineteen exons. *M. zebra* has two gene structure predictions, with one having nineteen exons and the second having twenty. The two *M. zebra* models shared the approximate same size first exon, while the *O. niloticus* was predicted to have a shorter Exon 1. Exons 2 through 17 were predicted to be the same size in all of the models. Exon 18 in the *O. niloticus* model was predicted to be the same as Exon 19 in the second *M. zebra* model, and Exon 19 was predicted to be unique in *O. niloticus*. Exon 18 in the first *M. zebra* model was predicted to be the same as Exon 18 in the second *M. zebra* model, with Exon 19 being unique in the first *M. zebra* model. Finally, Exon 20 was predicted to be unique to the second *M. zebra* model.

We decided to try and determine the structure of AAK1 using our previously made cDNA. Multiple PCR primers were designed to help distinguish the multiple models from each other and controls were run to show the size differences between the predicted cDNA PCR fragment sizes and the genomic PCR fragment sizes. Unfortunately, results were inconclusive; we either got multiple bands for PCR products, or got bands that were the same size for both the cDNA and genomic DNA, even when controls showed size differences between the two DNA types (Fig. 4-8).

Discussion

The single marker analysis with additional individuals microsatellites confirmed the previously identified QTL and helped to direct us to a more specific

area of the genome in which to look for candidate genes. When first looking for candidate genes, we focused our search on previously identified genes known to play a role in pigmentation in fishes. In particular, we looked for *csf1ra* and *ltk*. *Csf1ra* was previously identified to play a major role in xanthophore specification (Parichy et al., 2003a), while *ltk* was shown to play a major role in iridophore specification (Lopes et al., 2008). We hypothesized these two genes could play a role due to the fact that the main difference between the fin color in our two species of interest was the presence of xanthophores on the fins in *M. mbenjii* and iridorphores on the fins of *M. zebra*. However, neither of these genes or any other previously identified known pigmentation genes were found in the QTL interval.

We decided to focus on AAK1 as a candidate because the single marker analysis showed the most significant association just upstream of this gene and because of the evidence that AAK1 plays an important role in pigment organelle biogenesis. AAK1 stands for adaptor-associated kinase 1. This protein was found to bind directly to and modify via phosphorylation the μ subunit of adaptor protein 2 (AP2) (Conner and Schmid, 2002). Adaptor protein complexes, which include AP2, play an important role in the endosomal sorting system. In mice, AP2 has been shown to be important in sorting to early endosomes (Hearing, 2005). The endosomal sorting system is used in melanosome formation, and in zebrafish, it has been shown that mutations in different components of this system can cause the pigmentation of the fish to appear diluted (Navarro et al., 2008). Given its role in melanosome formation, it could be feasible that AAK1 could also play a role in the formation of other pigment organelles, such as xanthophores or iridophores.

Given that regions upstream of a gene are important in the binding of regulatory elements (Pierce, 2010), we decided to examine the promoter region of AAK1. When examining the genome assembly, we found the presence of a small gap in the proximal promoter. We were able to sequence across the gap in the proximal promoter, Gap 259, and identified one SNP between the grandparents of the cross. While this SNP was not fixed in natural populations of Mazinzi *M. zebra*, all but two individuals of *M. mbenjii* were fixed for the same genotype as the *M. mbenjii* grandfather, T/T. For future work, we would follow up on this SNP to see if it segregates in the hybrid cross based on pigmentation phenotype. We would also sequence this SNP in more *M. mbenjii* individuals. We could also sequence the existing upstream sequence and look for additional SNPs. It should be noted that there is another gap, Gap 258, which is found further upstream from Gap 259 and is approximately 5,000 bp long. We could also attempt to sequence this gap to look for fixed differences between the grandparents of the cross.

We also found two large gaps in the first intron, Gaps 260 and 261, which could also contain binding sites for transcription factors, thus causing changes in gene expression (Fedorova and Fedorov, 2003). However, sequencing across these gaps proved challenging, mostly due to the presence of repetitive sequence. This repetitive sequence is most likely the culprit of the gap being present in the assembly. While we were able to sequence completely across Gap 260, we were not able to sequence across the larger Gap 261. Gap 261 is especially interesting because there appears to be a size difference in the gap between the two grandparents. This could possibly play a role in the regulation of expression of AAK1, perhaps with certain regulatory

elements present in one but not the other. Aligning other fish genomes in this region did not reveal the presence of any conserved regions within the first intron. Additional work should be undertaken to see if this size difference appears in additional members of *M. zebra* and *M. mbenjii* to determine if this is indeed species specific.

We also found that we obtained much more sequence for Gap 261 than what was predicted from examining the sequence from other cichlid species. One possible explanation for this is the presence of a transposable element. Blasting the sequence we obtained from *M. zebra* to the NCBI database produced multiple hits to many cichlid genomes. While it didn't hit a transposable element, the multiple hits to several other cichlid genomes might indicate it is a repetitive element. Also, we could not discern tandem repeats that would indicate the presence of a transposon just upon visual inspection of the sequence, but it should not be ruled out as a possibility.

Early on, we had several issues with amplifying cDNA for AAK1. One reason for this could be that there are simply low levels of AAK1 cDNA in the adult fins of *M. mbenjii* and *M. zebra*. Also, while our initial primer sets should have been able to amplify our DNA regions of interest without issue, we designed new primer sets to ensure that this would not be an issue. We also decided to switch to working with cDNA from younger fish at the time points we hypothesized our candidate gene would be expressed. By making these changes, we were finally able to get cDNA amplification of AAK1 in one, two, and three month old fish, as well as in adults.

The last thing our study tried to determine was the structure of AAK1 due to three different structures being predicted. Our work from trying to amplify cDNA

tells us that exons 2-15 are in fact at the predicted sizes for all three models. However, when we tried to amplify exons 16-19, we did not obtain PCR product, which could indicate that these predicted exons don't exist. We created additional primers that were specific to the different predictions of the gene structure at the end of AAK1, since this was variable amongst the models. This also produced inconclusive results, as we ran into issues with amplified cDNA being the same size as the genomic DNA. Future work on this aspect of the project will involve designing primers to determine the structure of the gene for the first exon, as well as exons after exon 15.

Overall, this project has begun to analyze the feasibility of AAK1 as a candidate gene for dorsal and caudal fin coloration differences between *M. zebra* and *M. mbenjii*. While we ran into several issues during the process, we did begin to make progress in the analysis of this gene. In addition to suggested future work mentioned previously, there are several other types of analyses we can do. The first one is that of qPCR. By using this method, we could determine whether or not there are differences in expression in fish at different ages. Based on the developmental time study, it suggests that differences in gene expression might be observed between two and three months of age, when the differences in pigment cells populations between the two species becomes evident. Another technique we could use is *in situ* hybridization. This would allow us to visualize when and where expression of AAK1 might be occurring. Currently, we predict expression would occur within the fins themselves around 2-3 months of age. Overall, AAK1 needs further analysis until it is either shown to be or not to be a viable candidate.

Figures

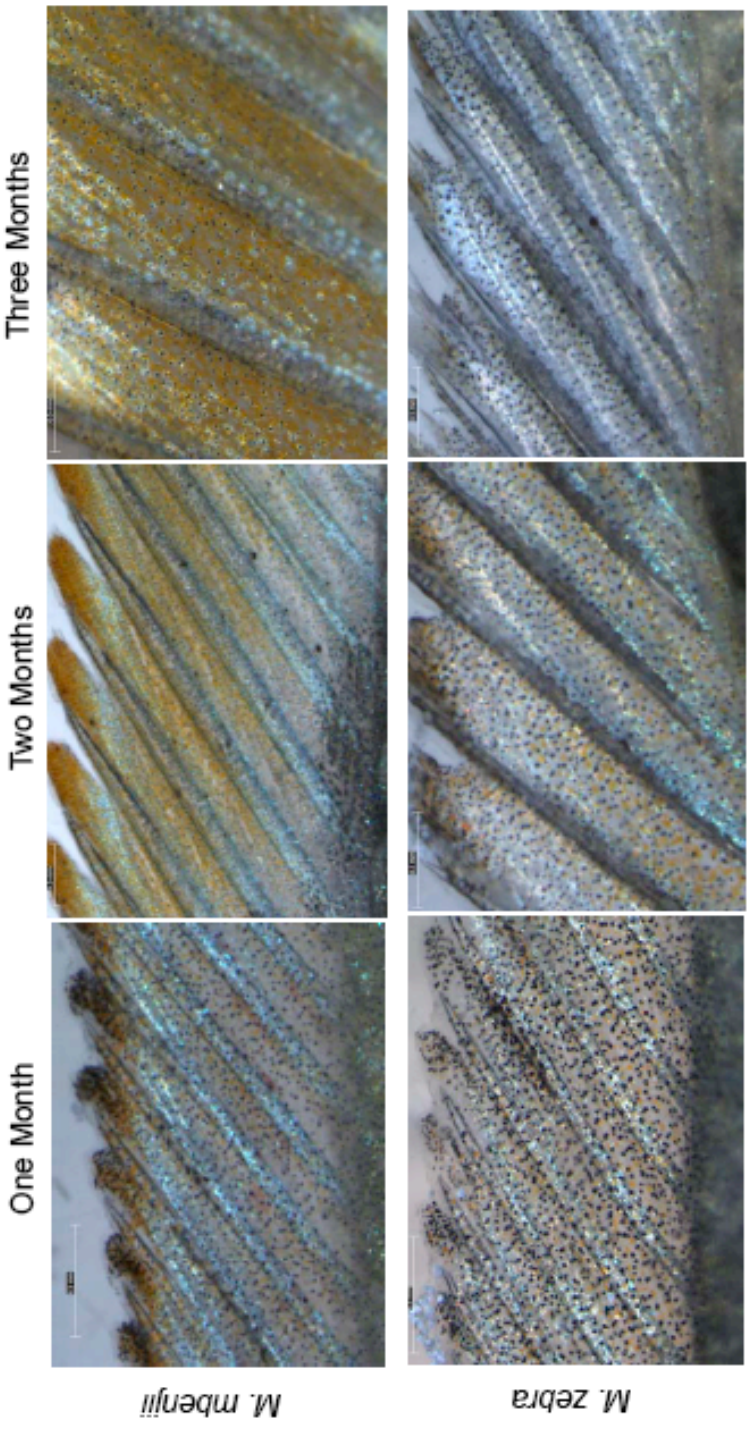


Figure 4-1. Images of the dorsal fin in *M. zebra* and *M. mbenjii* at one month, two months, and three months of age. At one month of age, the fins look very similar; however, the fins start to look different from one another at 2 months onward.

Figure 4-2. Single marker analysis curve for dorsal xanthophores. The marker at 4.19 Mb on Scaffold 9 had the highest F value.

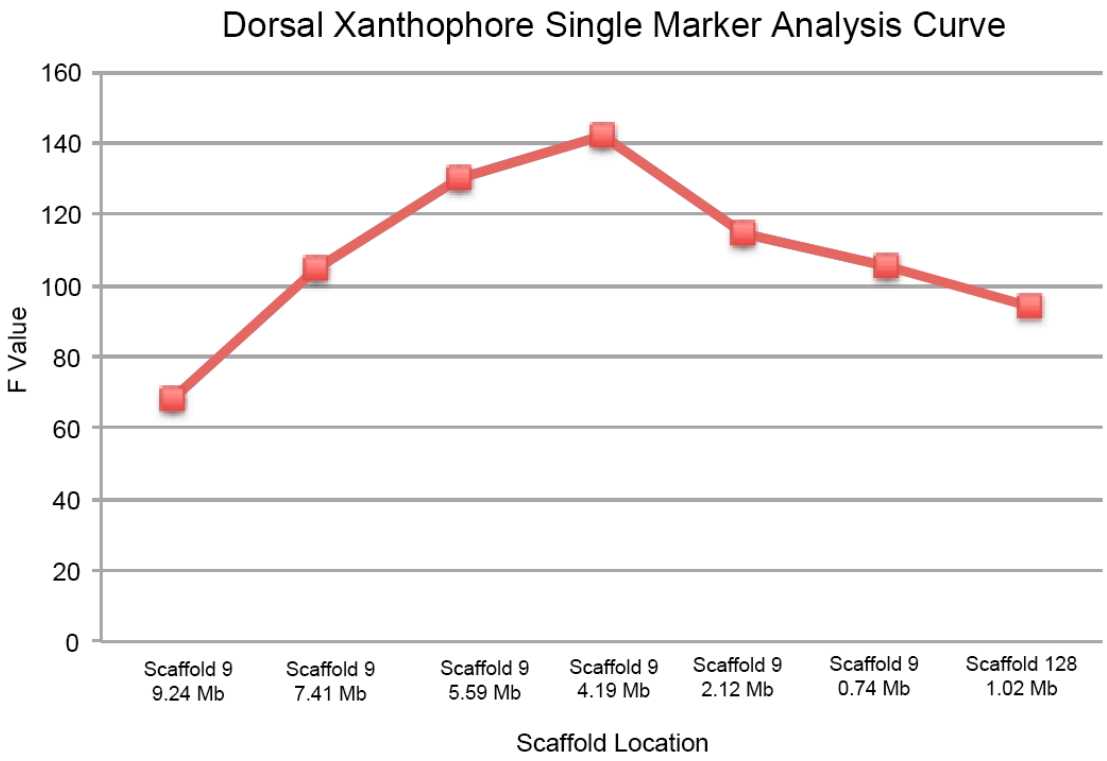


Figure 4-3. Image showing the location of assembly gaps 259, 260, and 261 located in the promoter and first intron of AAK1.

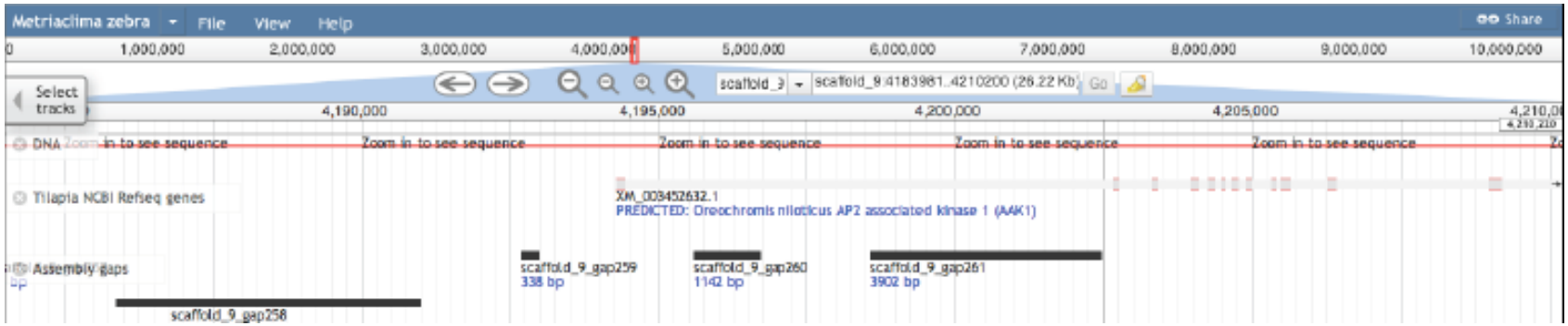


Figure 4-4. Example of PCR mismatch gels. (A) Gel showing PCR band differences in the grandfather and grandmother of my QTL cross. (B) Example of a mismatch gel for *M. zebra* males from natural populations.

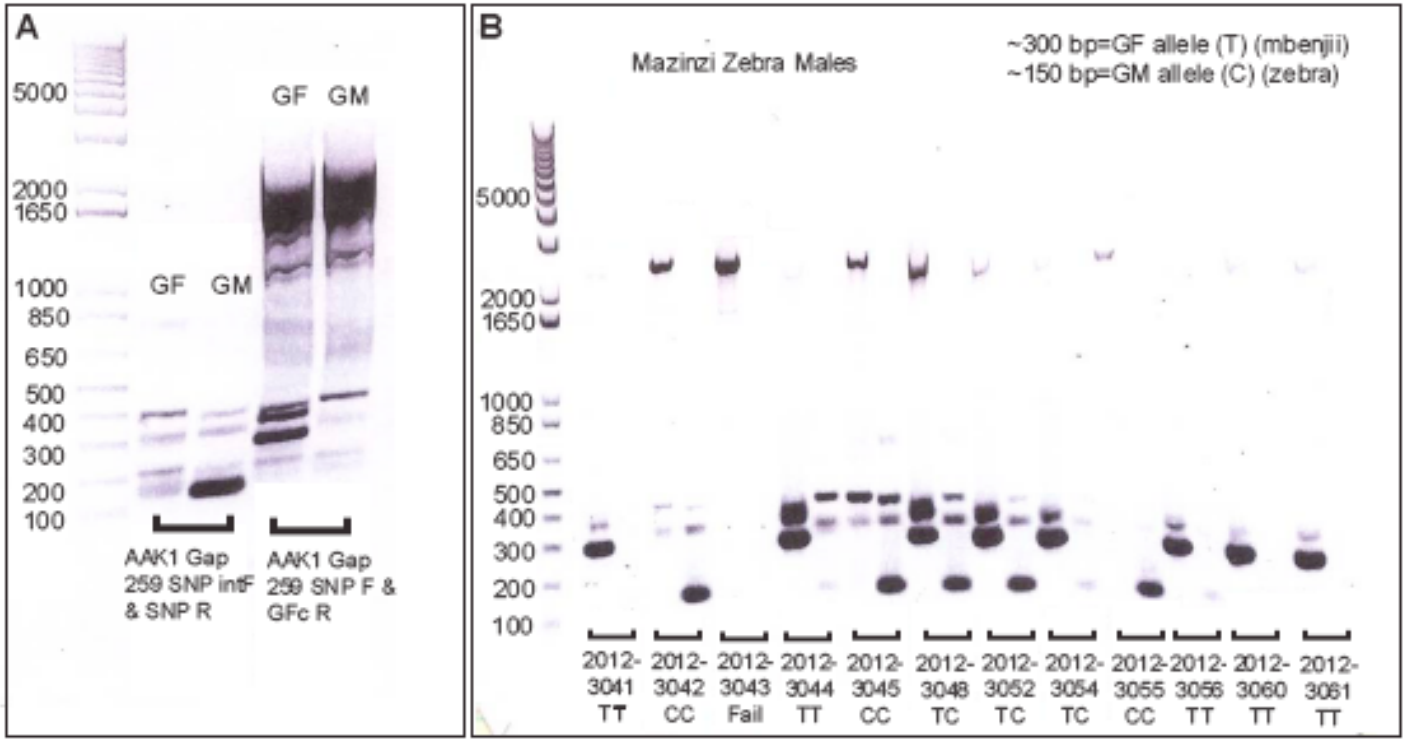


Figure 4-5. β-actin and AAK1 gels for cDNA made from adult fins. (A) β-actin controls for detecting the success of cDNA creation. Notice that the genomic fragments are all larger than the cDNA fragments. (B) Gel for amplification for cDNA in adult fins. Notice the presence of multiple bands for the genomic DNA and the lack of bands at the appropriate sizes in the cDNA.

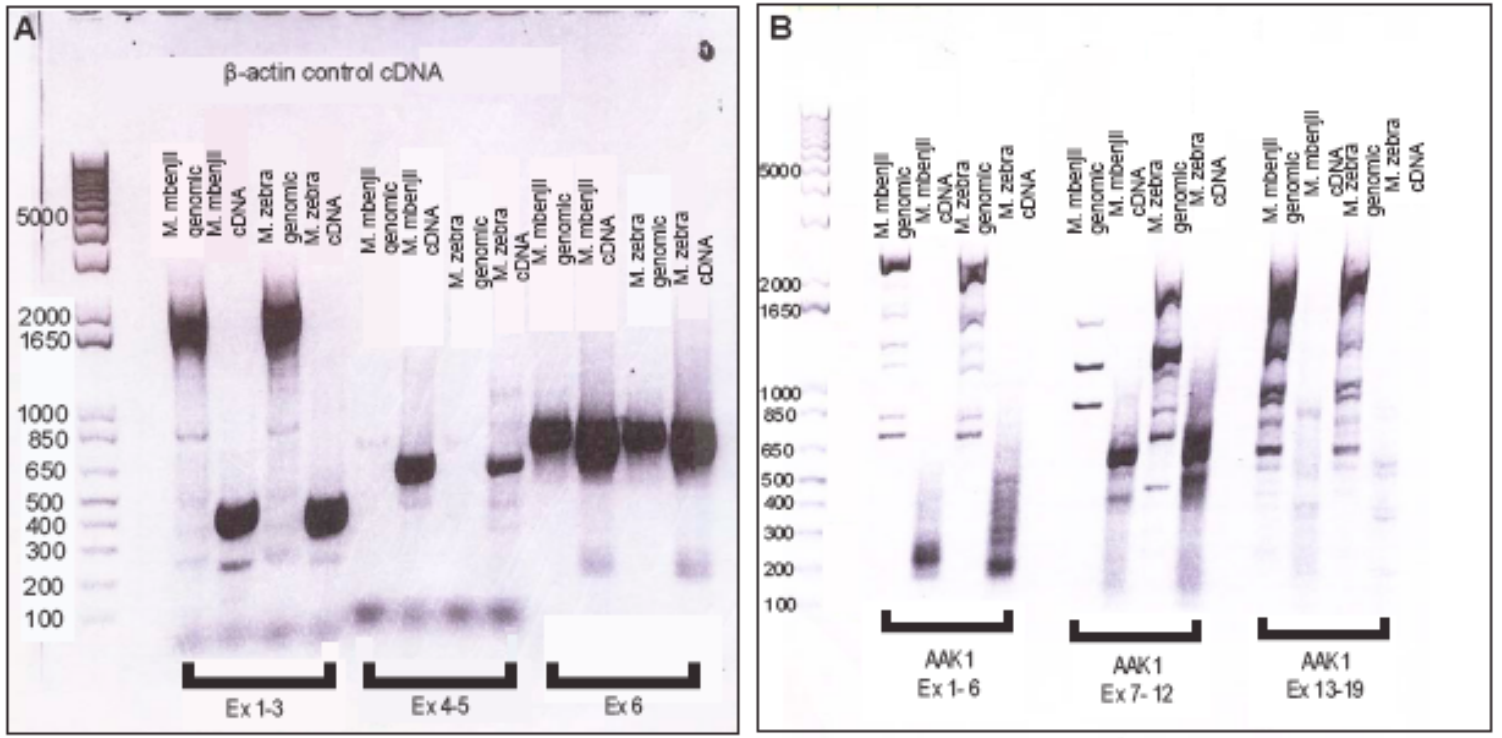


Figure 4-6. Dorsal fin cDNA gels for AAK1. (A) AAK1 cDNA from the dorsal fins of one month old *M. mbenjii* and *M. zebra*. Also included on this gel are β-actin controls and genomic DNA samples. (B) AAK1 cDNA from the dorsal fins of two month old and three month old *M. mbenjii* and *M. zebra*. Also included on this gel are β-actin controls. (C) AAK1 cDNA from the dorsal fins of adult *M. mbenjii* and *M. zebra*. Also included on this gel are β-actin controls.

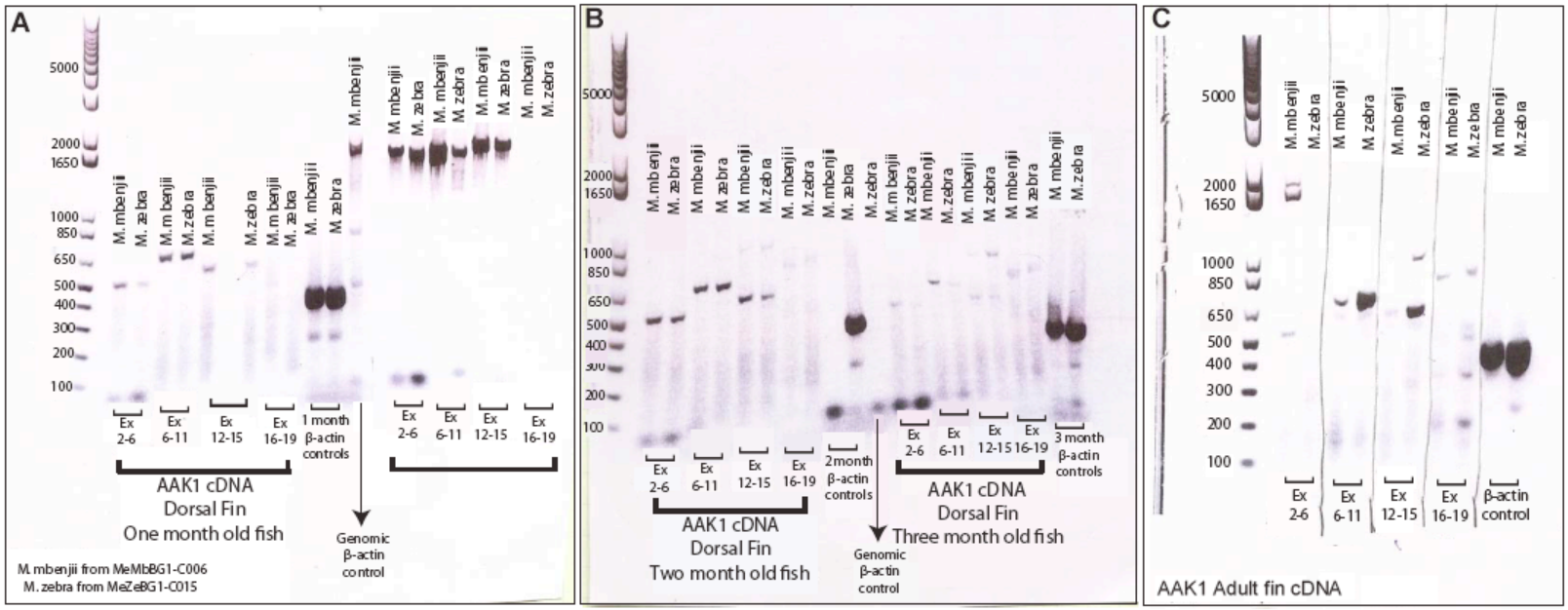


Figure 4-7. AAK1 gene structure predictions. *O. niloticus* (tilapia) is predicted to have 19 exons, while *M. zebra* is predicted to have two structures, one with 19 exons and one with twenty exons.

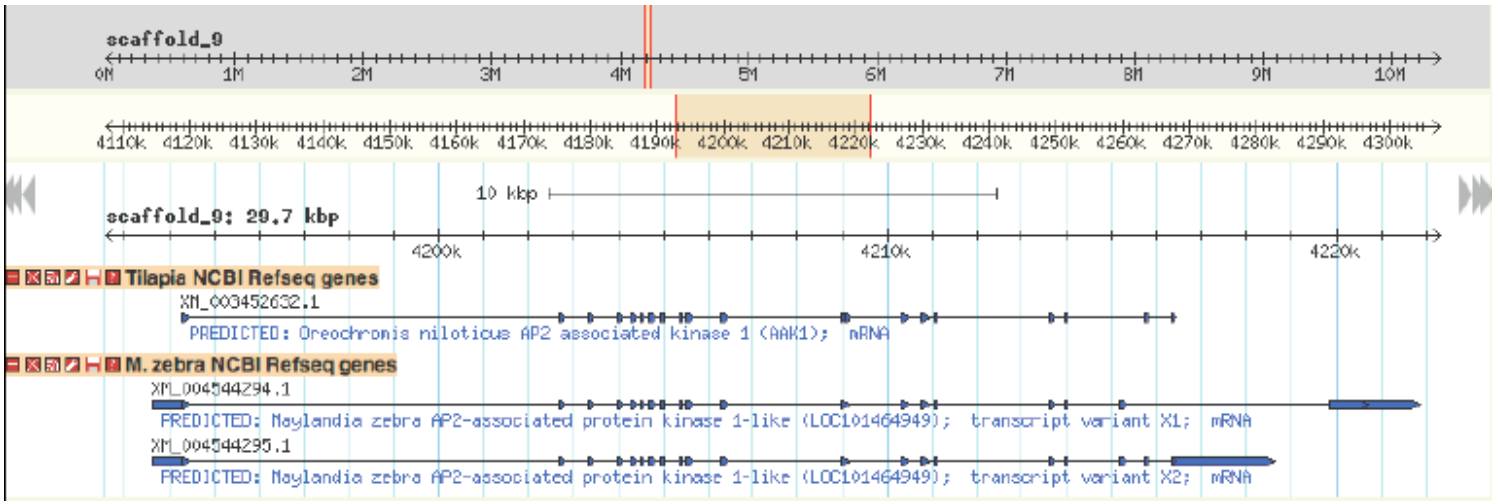
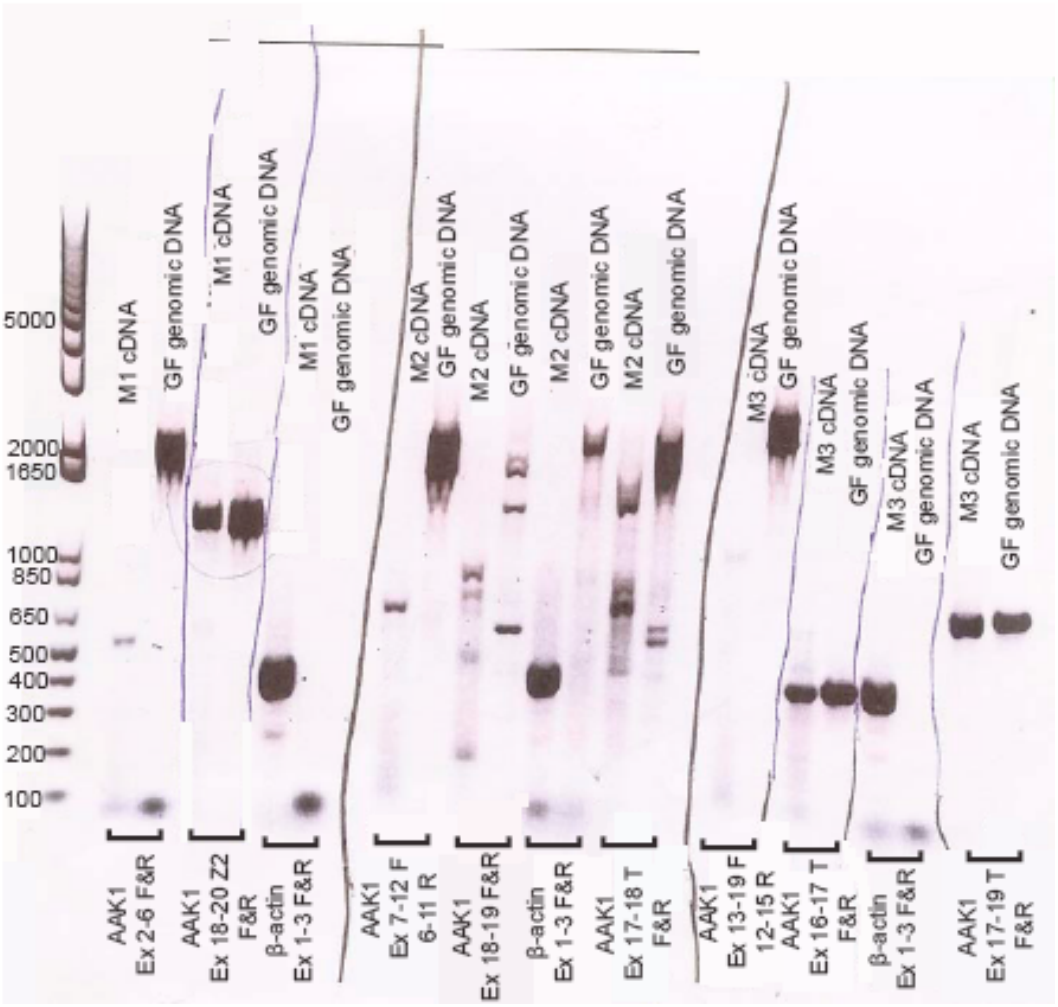


Figure 4-8. Gel showing PCR products in an attempt to determine the structure of AAK1. While the β-actin controls showed the presence of cDNA, I still obtained products where the cDNA and the genomic DNA fragments were the same size.



Tables

Table 4-1. Single marker analysis marker locations and F values.

Marker	Scaffold	Position	F value
MZ9388	9	9,239,908	68.22
MZ9245	9	7,411,658	105.02
UNH2012	9	5,592,622	130.34
MZ8983	9	4,192,304	142.17
MZ8826	9	2,199,167	114.72
MZ8730	9	742,447	105.47
UNH2162	128	1,015,293	93.81

Table 4-2. Genotype distribution for SNP in Gap 259.

Individual	Species	Sex	Genotype
2012-3041	Mazinzi *M. zebra*	Male	TT
2012-3042	Mazinzi *M. zebra*	Male	CC
2012-3043	Mazinzi *M. zebra*	Male	Failed PCR
2012-3044	Mazinzi *M. zebra*	Male	TT
2012-3045	Mazinzi *M. zebra*	Male	CC
2012-3048	Mazinzi *M. zebra*	Male	TC
2012-3052	Mazinzi *M. zebra*	Male	TC
2012-3054	Mazinzi *M. zebra*	Male	TT
2012-3055	Mazinzi *M. zebra*	Male	CC
2012-3056	Mazinzi *M. zebra*	Male	TT
2012-3060	Mazinzi *M. zebra*	Male	TT
2012-3061	Mazinzi *M. zebra*	Male	TT
2012-3064	Mazinzi *M. zebra*	Male	Failed PCR
2012-3066	Mazinzi *M. zebra*	Male	CC
2012-3067	Mazinzi *M. zebra*	Male	TC
2012-3257	Mazinzi *M. zebra*	Male	Failed PCR
2012-3258	Mazinzi *M. zebra*	Male	TC
2012-3259	Mazinzi *M. zebra*	Male	TC
2012-3260	Mazinzi *M. zebra*	Male	CC
2012-3297	Mazinzi *M. zebra*	Male	TC
2012-3046	Mazinzi *M. zebra*	Female	TT
2012-3047	Mazinzi *M. zebra*	Female	TT
2012-3049	Mazinzi *M. zebra*	Female	TT
2012-3050	Mazinzi *M. zebra*	Female	CC
2012-3051	Mazinzi *M. zebra*	Female	TT
2012-3053	Mazinzi *M. zebra*	Female	TT
2012-3057	Mazinzi *M. zebra*	Female	TC
2012-3058	Mazinzi *M. zebra*	Female	Failed PCR
2012-3059	Mazinzi *M. zebra*	Female	CC
2012-3062	Mazinzi *M. zebra*	Female	CC
2012-3063	Mazinzi *M. zebra*	Female	TT
2012-3065	Mazinzi *M. zebra*	Female	TC
2012-3255	Mazinzi *M. zebra*	Female	TC
2012-3256	Mazinzi *M. zebra*	Female	CC
2012-3261	Mazinzi *M. zebra*	Female	TT
2012-3262	Mazinzi *M. zebra*	Female	TC
2012-3263	Mazinzi *M. zebra*	Female	CC
2012-3264	Mazinzi *M. zebra*	Female	TC
2012-3322	Mazinzi *M. zebra*	Female	TC
2012-3347	Mazinzi *M. zebra*	Female	TT

Continued

Table 4-2 (Continued)

Individual	Species	Sex	Genotype
2012-3986	*M. mbenjii*	Male	CC
2012-3987	*M. mbenjii*	Male	Failed PCR
2012-3988	*M. mbenjii*	Male	TT
2012-3989	*M. mbenjii*	Male	TT
2012-3990	*M. mbenjii*	Male	TT
2012-3991	*M. mbenjii*	Male	TT
2012-3992	*M. mbenjii*	Male	TT
2012-3993	*M. mbenjii*	Male	TT
2012-3998	*M. mbenjii*	Male	TT
2012-4001	*M. mbenjii*	Male	TT
2012-4014	*M. mbenjii*	Male	TT
2012-3995	*M. mbenjii*	Female	TT
2012-3996	*M. mbenjii*	Female	TT
2012-3997	*M. mbenjii*	Female	TC
2012-3999	*M. mbenjii*	Female	TT
2012-4000	*M. mbenjii*	Female	Failed PCR
2012-4003	*M. mbenjii*	Female	TT
2012-4004	*M. mbenjii*	Female	TT
2012-4005	*M. mbenjii*	Female	TT
2012-4006	*M. mbenjii*	Female	TT
2012-4007	*M. mbenjii*	Female	TT
2012-4010	*M. mbenjii*	Female	TT
2012-4011	*M. mbenjii*	Female	Failed PCR
2012-4012	*M. mbenjii*	Female	Failed PCR
2012-4021	*M. mbenjii*	Female	TT
2012-4022	*M. mbenjii*	Female	TT
2012-4023	*M. mbenjii*	Female	TT
2012-4024	*M. mbenjii*	Female	TT

Table 4-3. Predicted fragment sizes for PCR amplification of AAK1

Fragment Being Amplified	**Genomic DNA Size (bp)**	**cDNA size (bp)**
Exon 1 to Exon 6	10,341	696
Exon 7 to Exon 12	6,199	996
Exon 13 to Exon 19	7,320	958
Exon 2 to Exon 6	1,848	525
Exon 6 to Exon 11	1,891	684
Exon 12 to Exon 15	2,054	600
Exon 16 to Exon 19	2,780	372

Table 4-4. Predicted exon sizes for the three current AAK1 models

	O. niloticus	***M. zebra Transcript 1***	***M. zebra Transcript 2***
Exon 1	160 bp	819 bp	818 bp
Exon 2	119 bp	119 bp	119 bp
Exon 3	109 bp	109 bp	109 bp
Exon 4	143 bp	143 bp	143 bp
Exon 5	122 bp	122 bp	122 bp
Exon 6	82 bp	82 bp	82 bp
Exon 7	133 bp	133 bp	133 bp
Exon 8	104 bp	104 bp	104 bp
Exon 9	80 bp	80 bp	80 bp
Exon 10	164 bp	164 bp	164 bp
Exon 11	173 bp	173 bp	173 bp
Exon 12	222 bp	222 bp	222 bp
Exon 13	188 bp	188 bp	188 bp
Exon 14	194 bp	194 bp	194 bp
Exon 15	69 bp	69 bp	69 bp
Exon 16	105 bp	105 bp	105 bp
Exon 17	102 bp	102 bp	102 bp
Exon 18	90 bp	159 bp	159 bp
Exon 19	95 bp	1991 bp	90 bp
Exon 20	N/A	N/A	2307 bp

Chapter 5: Identification of an XY sex determination region in Lake Malawi cichlid fishes

Claire O'Quin, Alexi Drilea, Ginni LaRosa, and Thomas Kocher

See Appendix 4 for supplementary table referenced in this chapter.

Abstract

Sex determination is the process by which a bipotential gonad becomes either male or female specific. Of the vertebrates, fishes display the largest variety of sex determination; with sex being determined genetically, environmentally, or by genotype-environment interactions. Cichlid fishes offer a unique opportunity to investigate the evolution of sex determination due to their recent evolutionary history and the role that sexual selection is thought to play in their rapid speciation. In this study, we analyze an F_2 hybrid cross between two cichlid species, *Metriaclima zebra* and *M. mbenjii*, in an effort to identify the sex determiner. Using a break point analysis, we were able to identify a 1.48 Mb region where an XY sex determiner was predicted to lie. We then identified and began to analyze a candidate gene, GSDF. We identified five polymorphisms in which the grandfather was heterozygous and the grandmother homozygous. Three of the polymorphisms in the promoter and the one in the first intron were analyzed for differential transcription factor binding, and a few were identified that could influence sex determination. More work is needed to definitively prove or disprove GSDF as the putative sex determiner in these fishes.

Introduction

The mechanism for sex determination, in which a bipotential gonad becomes male or female specific, is highly variable throughout the animal kingdom (Manolakou et. al, 2006). In vertebrates, sex can be determined genetically, environmentally, or a combination of both. In mammals, birds, and snakes, sex is determined genetically, with the presence of distinct sex chromosomes. Birds and snakes possess a female heterogametic system (WZ), while mammals possess a male heterogametic system (XY) (Ezaz et. al, 2006). However, in many reptiles, frogs, and fish, no differentiated sex chromosomes can be distinguished even though it is known that sex is determined genetically (Marshall Graves, 2008).

Vertebrates can also possess environmental sex determination. This is often observed in reptiles, amphibians, and fish, and can be observed as the sole method of sex determination or in conjunction with genetic sex determination (Ezaz et. al, 2006; Pieau, 1996). The most common environmental factor influencing sex determination is that of temperature. In many instances, sex is influenced by the temperature at which the egg is incubated, probably by alteration of hormone levels (Sarre et. al, 2004; Devlin and Nagahama, 2002).

Fishes display the largest variety of sex determination and sex chromosomal systems among vertebrates (Ezaz et. al, 2006). In species where sex is determined genetically, this can be accomplished via a single allele, multiple interacting loci, or segregation of morphologically distinct sex chromosomes. However, only about 10% of the fish studied karyologically have been found to possess differentiated sex chromosomes (Devlin and Nagahama, 2002). Sex determination dependent on

temperature or by genotype-environment interactions have also been established in fishes (Ezaz et. al, 2006).

Cichlid fishes offer a unique opportunity to investigate the evolution of sex determination due to their recent evolutionary history and the role that sexual selection is thought to play in their rapid speciation. Within the last 10 million years, over 2,000 cichlid species have evolved in Lakes Victoria, Tanganyika, and Malawi, with over 500 species being found in Lake Malawi alone (Genner et. al, 2007; Kocher, 2004). Ecological selection for habitat and feeding morphologies could be factor contributing to the emergence of new species within Lake Malawi (Albertson et. al 2005; Danley and Kocher, 2001). However, sexual selection is thought to also be a force in the speciation process. Sex roles are highly differentiated. Females are largely responsible for care of the offspring. Males compete for matings as evidenced by their pigmentation, with males displaying bright color patterns while females are typically drab and blend into the background habitat (Danley and Kocher, 2001).

Previous work has found several different sex-determining mechanisms in cichlids. Both XY and WZ sex determination systems have been found in the basal tilapiine species of cichlids (Lee et al., 2003; Lee et al., 2004). Both XY and WZ systems are also found in Lake Malawi cichlids, along with evidence of epistatic interactions between systems (Ser et al., 2010). Previous work in our lab indicates the presence of an XY sex determiner on LG7 in the genus *Metriaclima* (Ser et al., 2010). However, these studies analyzed small families, and were thus unable to map sex to a small region of the chromosome. We recently analyzed an F_2 hybrid cross between two *Metriclima* species, *M. zebra* and *M. mbenjii*, for variation in pigmentation traits

(O'Quin et al., 2010; O'Quin et. al, 2013). Using these same individuals, we were able to follow microsatellites on LG7 segregating with sex from the original *M. mbenjii* grandfather down through the F_2 lineage. This greatly increased the number of individuals we were able to utilize in our breakpoint analysis in order to narrow the genomic region associated with the XY sex determiner. Additionally, the release of a whole genome reference sequence for *M. zebra* allowed us to identify candidate genes within our region.

Materials and Methods

Hybrid Cross

One *M. mbenjii* male was crossed to one *M. zebra* female to produce a single F_1 family. This F_1 family consisted of a single male crossed to his sibling females to produce F_2 offspring. F_2 families were raised together until at least six months of age to allow them to reach sexual maturity. Fish were anesthetized with an overdose of MS-222 and sacrificed. The sex of each individual was determined by examination of gross gonadal anatomy. For this study, 221 F_2 males and 160 F_2 females were obtained. All animal procedures were approved by the University of Maryland IACUC (Protocol no. R-10-73).

Microsatellites

Fifteen microsatellite markers were identified and genotyped in the grandparents and 381 F_2 individuals (221 males and 160 females) along LG7. These included previously developed microsatellites and new microsatellites identified using the *M. zebra* genome assembly (Table S5-1). These microsatellites spanned a

region of 50 cM across six different scaffolds. Order of the scaffolds on the linkage group was previously determined (O'Quin et al., 2013).

PCR amplification was performed on DNA extracted from caudal fin clips. The size of the amplified microsatellite produced was determined on an ABI 3730 DNA sequencer using GeneMapper. Microsatellites were recorded and color-coded as being from the grandfather or grandmother at each genotyped site in each F_2 individual in order to perform a breakpoint analysis. Statistical analyses were performed using Fisher's exact test of independence.

Population Samples

20 individuals of each sex were sampled from the field for *M. zebra* at Mazinzi Reef and for 11 male individuals and 17 female individuals for *M. mbenjii* at Mbenjii Island. Individuals were dissected and sexed confirmed via inspection of the gonads. Fin clips were preserved in DMSO-salt buffer (Seutin et al., 1991). DNA was extracted from fin clips using standard phenol-chloroform/EtOH precipitation. DNA was quantified with Picogreen.

Candidate Gene Sequencing

NCBI Refseq and the *M. zebra* genome assembly were used to identify genes within the breakpoint region. We used a literature search to assess the viability of genes as candidates in playing a role in sex determination. Both *M. mbenjii* and *M. zebra* populations were examined since they are the grandparental species of the cross. Using the *M. zebra* genome assembly (www.bouillabase.org), primers were designed to amplify regions of interest in and around the candidate gene GSDF in the grandparents of the cross (Table S5-2). PCR products were gel purified for

sequencing. Sequences were aligned and polymorphisms identified using Sequencher. Identified polymorphisms were then sequenced in 11 males and 17 females of wild-caught *M. mbenjii* individuals. Statistical analyses were performed using Fisher's exact test of independence.

Transcription Factor Analysis

Approximately 20 bases surrounding each of the three identified polymorphisms in the promoter and first intron of GSDF were entered into JASPAR CORE Vertebrata database (http://jasper.genereg.net) in order to determine if any transcription factors might bind to that area. Both forms of the polymorphism were searched. All matrices were searched since we had no previous knowledge of transcription factors that could to bind these areas.

Results

Breakpoint Analysis

Of the 381 F_2 individuals genotyped, 121 were recombinant in the region studied. Fisher's exact test indicated the presence of an XY sex determiner in this region. Breakpoint analysis of the segregating alleles in these 121 individuals suggested the presence of a sex determiner located on Scaffold 21. Eight recombinant individuals helped to narrow this region, with five of them indicating the sex determiner lies approximately between 1.32Mb and 2.8Mb on Scaffold 21(Figure 5-1).

Candidate Gene Identification and Analysis

Within the recombinant region, 60 genes were predicted. The strongest candidate gene found within the recombinant region is that of growth and

differentiation factor 6B-like, found at approximately 1.95 Mb on Scaffold 21. This gene is a previously known sex determiner in fishes. Given the strong candidacy of GSDF (gonadal soma-derived factor) as a sex determiner, sequencing of this gene was undertaken in an effort to identify heterozygous polymorphisms between the *M. mbenjii* grandfather and *M. zebra* grandmother of the hybrid cross, which would be consistent with an XY sex determining system. All exons, introns, approximately 1 kb downstream of the gene, and the upstream region up to the next gene were sequenced for GSDF in both the grandmother and grandfather of our hybrid cross. While several polymorphisms were identified, only five were found to be heterozygous in the grandfather and homozygous in the grandmother. Three polymorphisms were located upstream of the gene at base pair positions 1,958,295 (Prom8-GF: G/T; GM: G/G), 1,958,086 (Prom7-8-GF: G/C; GM: C/C), and 1,955,171 (Prom3B-GF: A/C; GM: CC) on scaffold 21 (Figure 5-2). Two were located within the gene, one in intron one at base pair position 1,953,066 (Intron1-GF: A/G; GM: G/G) and one in exon 5 at base pair position 1,951,113 (Exon5-GF:A/T; GM: A/A) (Figure 5-2).

These five polymorphisms were then sequenced in wild-caught *M. mbenjii* male and female individuals in an effort to see if the polymorphisms were segregating in a sex specific manner in natural populations, with males being heterozygous and females being homozygous (Table 5-1). Statistical analysis revealed that none of these polymorphisms were segregating in a statistically different manner between males and females. Prom7-8 was close to segregating in a statistically different manner between males and females, with $p=0.06$. At this polymorphic site, 9 males

were heterozygous C/G, 1 was homozygous G/G, and one had poor sequencing results. For the females, there were 7 heterozygous C/G, 5 homozygous C/C, 4 homozygous G/G, and one with PCR that never worked.

A search for transcription factors that bound the three promoter polymorphisms and the polymorphism in the first intron found several different possibilities. We searched both forms of the polymorphism at each site. With each site, it was found that some transcription factors would bind both forms of the polymorphism, while some were only predicted to bind one of the polymorphisms. Table 5-2 lists the transcription factors predicted to bind each site and the sequence they are predicted to bind.

Discussion

Previous work by Ser et al. found an XY sex determination system on LG7 for several *Metriaclima* species (2010). The two species that were the grandparents of our cross were included in this study. While the *M. mbenjii* results were inconclusive, *M. zebra* had results indicating the presence of an XY sex determination system. Knowing this, we decided to focus on this linkage group when investigating sex in our cross. Thanks to access to the *M. zebra* genome assembly (www.bouillabase.org) and previous work in our lab (O'Quin et al., 2013), we had a better idea about the order of scaffolds on linkage group 7, which allowed us to add more markers for our breakpoint analysis. This, combined with a large F_2 family size, allowed us to narrow sex down to a 1.48 Mb region.

We decided to focus on GSDF as our primary candidate gene due to previous work linking it to sex determination in other fishes. GSDF is a novel member of the

TGF-β family found only in teleosts (Sawatari et al., 2007). It was first identified in trout and it was shown that expression is limited to the gonads (Mazurais et al., 2005; Sawatari et al., 2007). Expression studies in trout, salmon, and medaka has shown that GSDF is expressed in the Sertoli cells in the developing testes and the granulosa cells of the ovaries (Sawatari et al, 2007; Luckenbach et al., 2008; Gautier et al., 2011a).

Sequencing of the coding region of GSDF, as well as the upstream and downstream regions was undertaken due to evidence found by Myosho et al. 2012. They determined that mutations existed in both of these regions in GSDF in *O. luzonensis*, in addition to the coding region. However, the mutations in the coding regions did not change the amino acid. In fact, it was determined that mutations in the upstream region appeared responsible for upregulation of GSDF expression in males of *O. luzonensis*. In a recent study by Rondeau et al., it was found that between males and females of the sablefish, there existed male and female specific insertions within the upstream region of GSDF (2013). While they do not have expression data to confirm that GSDF is the sex determiner in this fish, their evidence is very compelling.

In our study, it was one of the heterozygous polymorphisms in the upstream region, (Prom7-8) which was close to being statistically significant in being found more in males than females, which could be an indication that the upstream region is important in cichlids as well. Unfortunately, none of the polymorphisms we identified showed a strict segregation with all males being heterozygous and all females being homozygous. This could indicate the presence of another sex determiner, which has

been suggested previously as a possibility in cichlids (Ser et al., 2007). More samples should be examined to see if this segregation continues to hold true.

Due to four out of the five heterozygous polymorphisms being in regions where transcription factors could bind and regulate gene expression, we decided to investigate what transcription factors were predicted to bind to those specifically sites, with they hypothesis that differential binding of transcription factors could help in making males versus females. Also, it was previously shown that the proximal promoter region of GSDF contains evolutionarily conserved *cis* regulatory elements in teleosts (Gautier et al., 2011b). However, all of the polymorphisms we investigated were much further upstream than where these conserved elements were predicted to bind (Gautier et al., 2011b).

While we found several different predicted transcription factors for each site, we found only a few that appeared to be connected to sex determination. It was predicted that Sox9 and Sox17 could bind the alternate form of the Prom3B polymorphism. Sox9 has been previously shown to play a role in male-specific sex-determination in both mouse and chicken (Kent et al., 1996). In tilapia, *sox9* was expressed at similar levels in both XX and XY gonads during early development, with higher levels of expression in XY gonads at 35-70 days after hatching (Ijiri et al., 2008). This is an interesting result because the alternate form of the Prom3B polymorphism, which Sox9 binds, is the heterozygous form you would expect to find in an XY male.

Sox17 was also predicted to bind the alternate form of the Prom3B polymorphism. Sox17 has been shown to be involved in frog and mouse

spermatogenesis (Navarro-Martin et al., 2009). However, in the sea bass, expression was found to be higher in females than males during sex differentiation. This would not match the same model predicted for Sox9. However, this expression pattern may be limited to sea bass, especially since this fish displays growth-dependent sex differentiation, in which larger fish become females (Navarro-Martin et al., 2009). More work would be needed to see if Sox17 expression might be different in fish with a strict genetic basis for sex determination.

USF1 was predicted to bind both forms of the polymorphism found in intron 1. This factor was found to interact with SF1 and GATA factors in order to regulate the transcription of *Fshr* in Sertoli cells of the testes (Gautier et al., 2011b). While we only focused on transcription factors that bound the polymorphisms we identified, it would be interesting to look at the rest of the promoter to see if the SF1 and GATA binding sites are also conserved in cichlids.

BRCA1 was predicted to bind the shared form of Prom 8 and Prom7-8. While, BRCA1 is most well known for its role in breast cancer, one of its normal functions appears to be control of the cell cycle and repression of cell proliferation (Zhang et al., 1998). While not directly linked to sex determination, BRCA1 might be able to influence it indirectly. It is thought that cell proliferation in a bipotential gonad can push it towards a male fate. In mice it was shown that blocking cell proliferation during a specific 8 hour period of gonad development blocked the development of testis cords and caused reduced or blocked expression of male specific genes in the XY gonads (Schmahl and Capel, 2003). Therefore, if BRCA1 can repress cell proliferation, it is conceivable that since it binds the shared form of the

polymorphism, which would be homozygous in females, that it blocks cell proliferation in genetic females, pushing the gonad to the female fate.

While we did not definitively identify GSDF as the sex determiner, we were able to preliminarily analyze the gene for its role in sex determination. Future work could include looking at expression differences in male versus female gonads, expanding the number of population samples sequenced for the polymorphisms, and investigating if any of the polymorphisms specifically drive gene expression by using reporter assays.

Figure 5-1. Breakpoint analysis of recombinant individuals. The individuals whose names are highlighted in purple are females and those highlighted in green are males. The top panel shows the chromosome inherited from the F_1 mother for each individual, with the dark pink representing one of her chromosomes and the light pink representing her other chromosome. The bottom panel shows the chromosome inherited from the F_1 father. The dark blue represents one chromosome while the light blue represents the other. We believe the dark blue chromosome to be the Y chromosome, since this is the chromosome the F_1 father inherited from the F_0 grandfather. The females all have the light blue chromosome between marker MZ15764 and Labeo2016, while all males have the dark blue chromosome between those markers, indicating the gen for sex lies between those markers. It should be noted that the scaffold and base for each marker is listed below the marker names.

	UNH2092 93:1616439	UNH2199 99:932675	UNH973 0:15726268	MZ1012 0:13037865	MZ371 0:5124400	MZ15715 21:725939	MZ15761 21:1159614	MZ15764 21:1328483	MZ15782 21:1697011	MZ15808 21:2201516	MZ15818 21:2381267	Labeo2016 21:2801036	MZ15984 21:5091978	UNH896 100:1113309	UNH2044 115:1910371
F2-006	174	176	164	203	227	262	246	217	250	244	273	186	162	136	160
F2-063	182	170	133	239	227	262	246	217	250	244	273	186	162	136	160
F2-190	226	142	146	207	248	250	301	223	259	239	225	223	179	183	167
F2-260	162	136	140	233	218	288	301	217	238	239	225	223	179	183	167
F2-313	182	170	133	239	227	262	246	217	250	244	273	186	162	188	170
F2-026	182	170	133	239	227	262	246	217	250	244	273	186	162	136	160
F2-306	182	170	133	239	227	262	246	217	250	244	273	186	162	188	170
F2-244	174	176	140	233	248	250	NA	223	259	239	225	223	179	183	167

	UNH2092 93:1616439	UNH2199 99:932675	UNH973 0:15726268	MZ1012 0:13037865	MZ371 0:5124400	MZ15715 21:725939	MZ15761 21:1159614	MZ15764 21:1328483	MZ15782 21:1697011	MZ15808 21:2201516	MZ15818 21:2381267	Labeo2016 21:2801036	MZ15984 21:5091978	UNH896 100:1113309	UNH2044 115:1910371
F2-006	162	136	140	233	218	288	301	221	246	244	208	200	196	188	170
F2-063	182	170	133	239	227	252	263	221	246	244	208	208	139	170	165
F2-190	162	136	140	233	218	288	301	217	246	244	208	200	196	188	170
F2-260	162	136	140	233	218	288	301	221	246	244	208	200	196	188	170
F2-313	162	136	140	233	218	288	263	221	246	244	208	200	196	188	170
F2-026	182	170	133	239	227	252	263	221	238	217	225	208	139	170	165
F2-306	162	136	140	233	218	288	301	217	238	217	225	200	196	188	170
F2-244	182	170	133	239	227	252	NA	221	238	217	225	208	139	170	165

Figure 5-2. Gene model of GSDF. Numbers indicate the number of base pairs. Stars indicate the approximate location of the five heterozygous polymorphisms. The genotypes of the grandfather (GF) and grandmother (GM) at these polymorphic sites are found under the name assigned to each polymorphism.

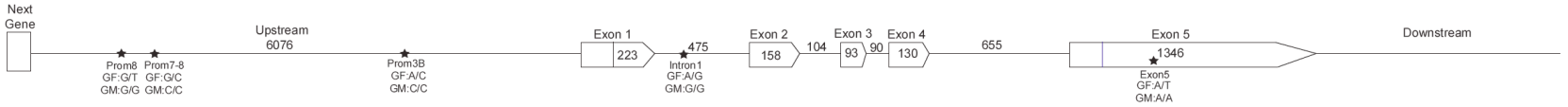


Table 5-1. Polymorphism sequencing results in population samples. Fisher's exact for males versus females at each site.

	Prom8	Prom7-8	Prom3B	Intron1	Exon5
Mbenjii-male-2012-3986	G/G	C/G	C/C	A/G	Bad Sequence
Mbenjii-male-2012-3987	G/G	C/G	A/C	A/G	Bad Sequence
Mbenjii-male-2012-3988	G/G	C/G	A/C	A/G	A/T
Mbenjii-male-2012-3989	G/G	C/G	C/C	A/A	T/T
Mbenjii-male-2012-3990	G/G	C/G	A/C	A/G	A/T
Mbenjii-male-2012-3991	G/T	C/G	A/C	A/G	Bad Sequence
Mbenjii-male-2012-3992	G/G	G/G	C/C	A/G	T/T
Mbenjii-male-2012-3993	G/T	C/G	C/C	A/G	Bad Sequence
Mbenjii-male-2012-3998	G/G	C/G	A/C	G/G	A/A
Mbenjii-male-2012-4001	G/G	Bad Sequence	A/C	A/G	A/T
Mbenjii-male-2012-4014	G/G	C/G	C/C	A/A	A/T

Continued

Table 5-1 (Continued)

	Prom8	Prom7-8	Prom3B	Intron1	Exon5
Mbenjii-female-2012-3995	G/G	G/G	C/C	G/G	Bad Sequence
Mbenjii-female-2012-3996	G/G	C/C	A/A	G/G	A/A
Mbenjii-female-2012-3997	G/G	C/G	A/C	G/G	A/A
Mbenjii-female-2012-3999	G/G	C/G	A/C	G/G	A/A
Mbenjii-female-2012-4000	G/G	G/G	C/C	A/A	T/T
Mbenjii-female-2012-4003	G/G	C/G	A/C	A/G	Bad Sequence
Mbenjii-female-2012-4004	G/G	C/C	A/C	A/G	A/T
Mbenjii-female-2012-4005	G/G	C/G	A/A	G/G	A/T
Mbenjii-female-2012-4006	G/G	C/C	A/A	G/G	A/T
Mbenjii-female-2012-4007	G/G	C/G	A/C	A/G	Bad Sequence
Mbenjii-female-2012-4010	G/G	G/G	A/C	G/G	A/T
Mbenjii-female-2012-4011	G/T	C/G	C/C	A/G	A/T

Continued

Table 5-1 (Continued)

	Prom8	Prom7-8	Prom3B	Intron1	Exon5
Mbenjii-female-2012-4012	G/G	G/G	A/C	A/A	A/T
Mbenjii-female-2012-4021	Failed PCR	C/C	C/C	A/A	T/T
Mbenjii-female-2012-4022	G/G	Failed PCR	Bad Sequence	G/G	A/T
Mbenjii-female-2012-4023	G/G	C/G	C/C	A/G	A/T
Mbenjii-female-2012-4024	G/G	C/C	C/C	A/G	A/A
Fisher's Exact p-value	0.55	0.06	0.35	0.08	0.69

Table 5-2. Polymorphic sites and the predicted transcription factors that bind them.

Polymorphism	Site and Surrounding Sequence	Transcription Factor Name	Transcription Factor Binding Sequence	JASPAR Score
Prom8	TCTAAACGTTT**G**TCCGTGTATA	Foxd3*	AAACGTTT**G**TCC (+)	0.82/0.81
Prom8 Alternate	TCTAAACGTTT**T**TCCGTGTATA	Hltf°	AAACGTTT**G**T (+)	0.82
		BRCA1°	A**C**AAACG (-)	0.87
		SOX10°	GTTT**G**T (+)	0.82
		ELK1*	TACACGGA**C**A (-)	0.81/0.87
		ETS1*	T**G**TCCG (+)	0.82/0.99
		HIF1A::ARNT °	**G**TCCGTGT (+)	0.82
		ELF5†	TTT**T**TCCGT (+)	0.88
		NFATC2†	TT**T**TCCG (+)	0.89
Prom7-8	GTCCGCGAA**C**ACACGGAAAAATCC	BRCA1°	CGAA**C**AC (+)	0.82
Prom7-8 Alternate	GTCCGCGAA**G**ACACGGAAAAATCC	Myc°	AA**C**ACACGGA (+)	0.81
		Arnt::Ahr°	CGTGT**G** (-)	0.84
		ELK1*	**C**ACACGGAAA (+)	0.89/0.90
		Pax2†	A**G**ACACGG (+)	0.90

Continued

Table 5-2 (Continued)

Polymorphism	Site and Surrounding Sequence	Transcription Factor Name	Transcription Factor Binding Sequence	JASPAR Score
Prom3B	GGATATTTAAAAGC**C**GAATGGGAAG	TBP*	ATATTTAAAAGC**C**GA (+)	0.84/0.84
Prom3B Alternate	GGATATTTAAAAGC**A**GAATGGGAAG	SPIB*	AGC**C**GAA (+)	0.82/0.86
		Mafb°	GC**C**GAATG (+)	0.85
		TFAP2A°	GC**C**GAATGG (+)	0.83
		Sox2†	CCATTC**T**GCTTTTAA (-)	0.82
		Sox9†	GC**A**GAATGG (+)	0.80
		Sox17†	CCCATTC**T**G (-)	0.82
		Sox10†	CATTC**T** (-)	0.89
Intron 1	GTGCTGATGCAC**G**TGCGGTGAGGTC	MAX°	ATGCAC**G**TGC (+)	0.81
Intron 1 Alternate	GTGCTGATGCAC**A**TGCGGTGAGGTC	Myc*	TGCAC**G**TGCG (+)	0.85/0.82
		Mycn°	TGCAC**G**TGCG (+)	0.86
		HIF1A::ARNT*	GCAC**G**TGC (+)	0.99/0.85
		USF1*	CAC**G**TGC (+)	0.92/0.82
		Arnt*	CAC**G**TG (+)	0.99/0.83

Continued

Table 5-2 (Continued)

Polymorphism	Site and Surrounding Sequence	Transcription Factor Name	Transcription Factor Binding Sequence	JASPAR Score
		Arnt::Ahr*	CAC**G**TG (+)	0.80/0.80
		ZEB1*	CAC**G**TG (+)	0.88/0.82
		ELK4†	ACCGCA**T**GT(-)	0.83

For transcription factors where they bind both forms of the polymorphism, the JASPAR score for binding the shared form is listed first and the alternate form second.
*=binds both forms of polymorphism
°=binds shared form
†=binds alternate form
(+)=positive strand
(-)=negative strand

Appendices

Appendix 1:

Supplemental Tables and Figures from Chapter 2

Supplementary Table S2-1. Comparison of mean scale melanophore number of bars and spaces at $p<0.05$

	Bar	**Space**	**p value**
M. zebra	1.91	1.74	0.00013
M. mbenjii	1.51	1.52	0.79
F_1 hybrid	1.61	1.56	0.08
F_2 hybrid	1.67	1.59	$8.005e^{-06}$

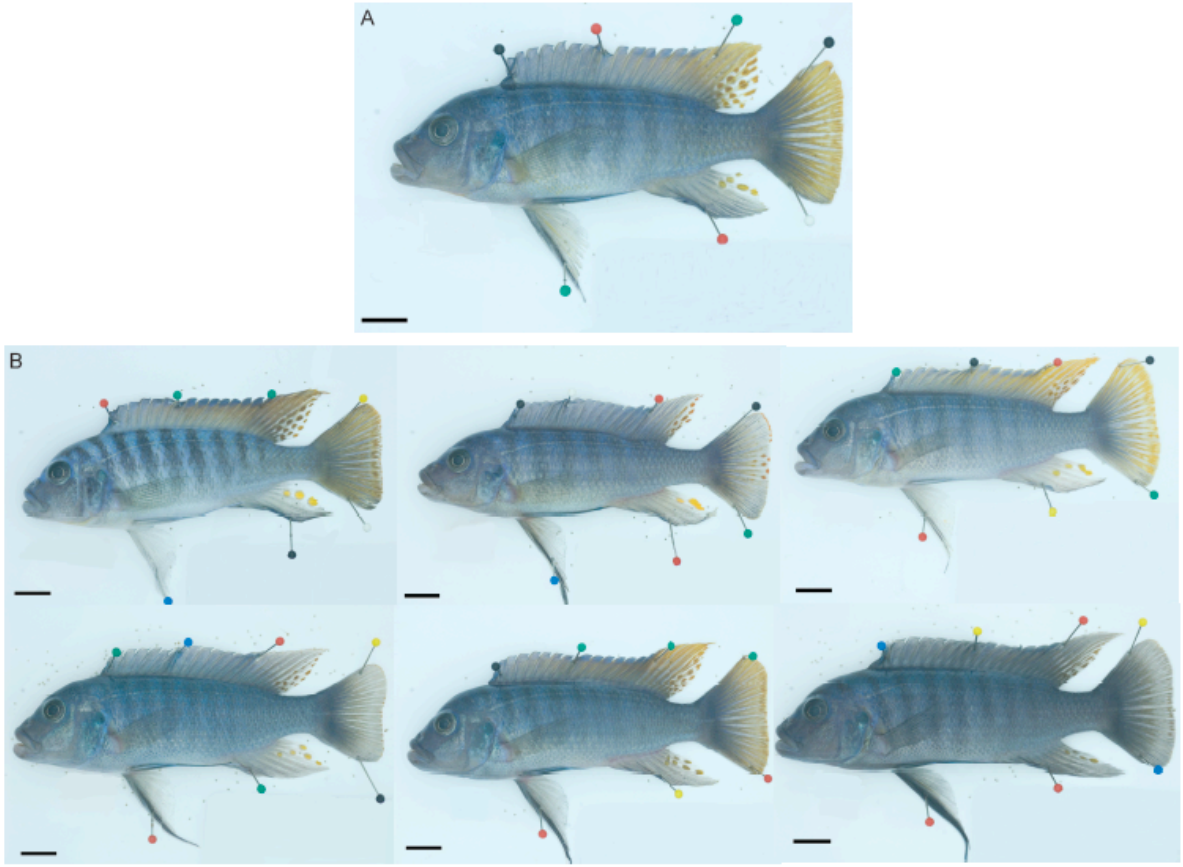
Supplementary Table S2-2. Loading scores for principal components and percent variance explained.

	Principal Component 1	Principal Component 2	Principal Component 3	Principal Component 4	Principal Component 5
Bar	-0.41	0.08	-0.16	0.11	-0.03
Space	-0.36	0.18	-0.29	0.20	-0.07
Dorsal Melano	-0.36	0.08	-0.24	0.18	-0.23
Caudal Melano	-0.36	-0.29	-0.17	-0.11	-0.06
Pelvic Melano	-0.39	0.03	0.08	-0.04	0.22
Cheek Melano	-0.32	-0.31	-0.21	-0.22	0.11
Dorsal Xantho	0.27	-0.30	-0.42	0.40	-0.06
Caudal Xantho	0.26	-0.41	-0.48	0.10	0.20
Pelvic Xantho	0.08	0.40	-0.40	-0.37	0.66
Gular	0.14	0.11	-0.35	-0.60	-0.61
Bar Spacing	0.12	0.59	-0.26	-0.37	-0.16
% of Variance Explained	40.1%	13.8%	11.9%	9.3%	7.2%
Castle-Wright Estimate	1.56	1.06	0.92	0.05	0.21
Standard Error	1.84	13.50	19.92	0.01	2.64

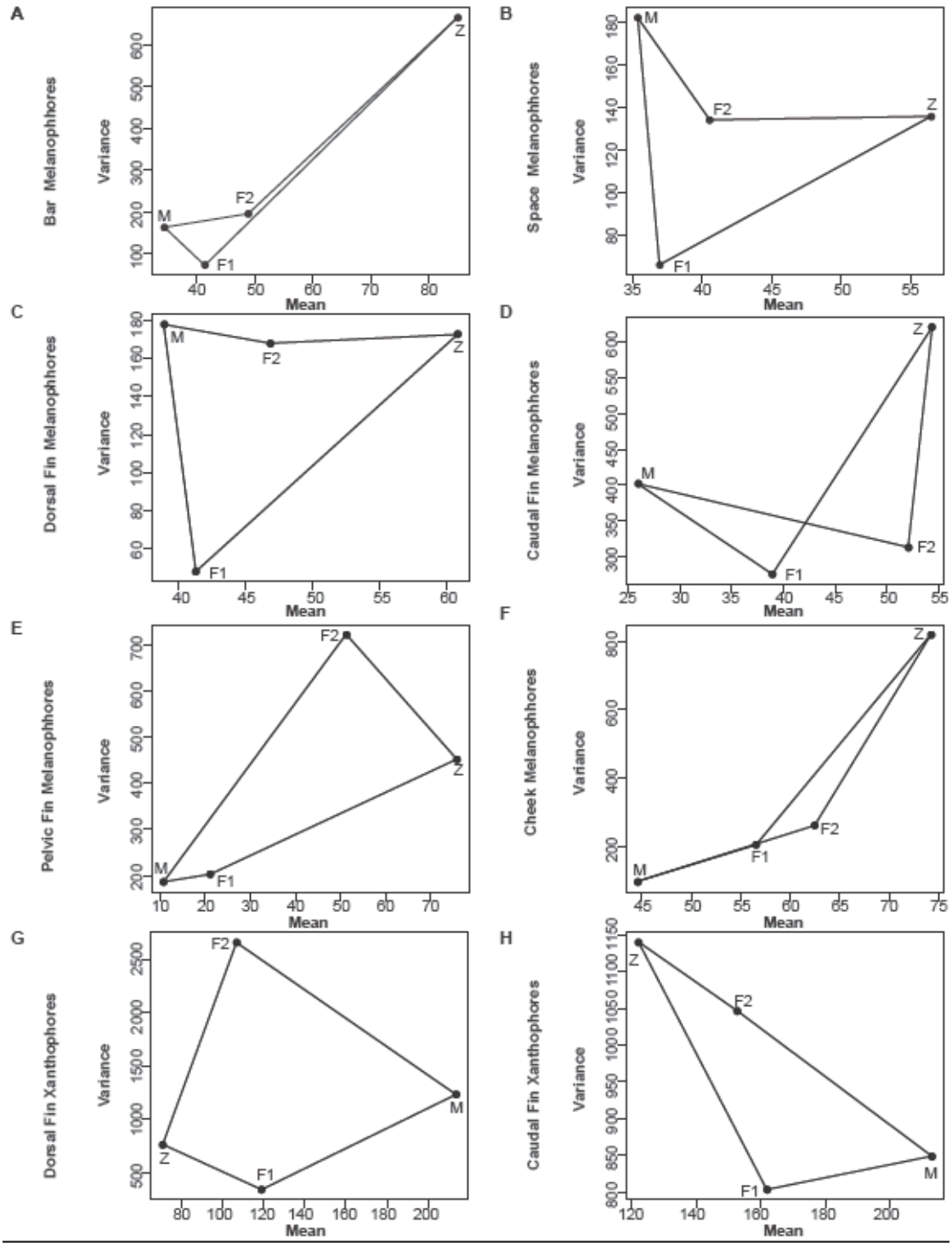
Supplementary Table S2-3. Dominance calculations for each phenotype.

Phenotype	Additive genotypic value (a)	Dominance genotypic value (d)	Dominant genotype	Degree of dominance (d/a)
Bar Scale Melanophores	25.3	-18.3	M. mbenjii	-0.72
Space Scale Melanophores	10.55	-0.85	M. mbenjii	-0.85
Dorsal Fin Melanophores	10.95	-8.6	M. mbenjii	-0.79
Caudal Fin Melanophors	14.18	-1.23	M. mbenjii	-0.09
Pelvic Fin Melanophores	32.5	-11.35	M. mbenjii	-0.35
Cheek Melanophores	14.83	-2.88	M. mbenjii	-0.19
Dorsal Fin Xanthophores	71.44	-23.35	M. mbenjii	-0.33
Caudal Fin Xanthophores	45.47	-5.63	M. mbenjii	-0.12

Supplementary Figure S2-1. Male hybrid phenotypes. A) Example of a typical F_1 hybrid male. B) Examples of the range of phenotypes observed in F_2 hybrid males.



Supplementary Figure S2-2. Triangle plots of phenotypic means plotted against the variance. A) Bar scale melanophores B) Space scale melanophores C) Dorsal fin melanophores D) Caudal fin melanophores E) Pelvic fin melanophores F) Cheek melanophores G) Dorsal fin xanthophores H) Caudal fin xanthophores. M=*M. mbenjii*, Z= *M. zebra*, F1=F_1 hybrids, F2=F_2 hybirds.



Appendix 2:

Supplemental Tables from Chapter 3

Supplementary Table S3-1. Non-RAD primer sequences

Primer Name	Forward Primer Sequence	Reverse Primer Sequence	Genomic Location-Scaffold:Base
Spindly	CAGGTGGTTGGTTTTCTTGG	TCTCTCAGACGCTCTCTGCATC	34:228462
CSF1R	TGAACAGGTTGCTTACGCTCT	TGCTGTCTCTCTTGGTGGTGTT	198:829661
MZ28416	CCATCAGCCAATCAGCTACA	GGAAGCCCCAGTGTCAGATA	58:1973203
usat14168	TGCTCAACACAGCTAAGCAAA	AGAGGGATTAAGGCGGATGT	18:1030085
usat9538	CACAAGGTCCAATTCCATCC	TTTTGCTTCCACCACGATAA	10:1042581
MZ10190	CTGTCTTCGCGTGGTGACTA	CGGTGTAATAAACACCCACCA	10:9603154
CADH	GGACGGTTCTTTCTTGGTCA	ACTTCCTGCCAGTGTTGGAC	273:205421
usat49371	GGGATTAACGCTGCCTCATA	AGCCTCATTCGCAGAGTCAT	202:914425
MITF	TGCCGACTCTAAGAGCACCT	GGGCAAGCCATTGTTCAGTA	130:258851
MZ29472	GCCCATTGATGTGGCTTAAA	GCTGCCTTTCCAGCTATTCA	63:2335426
MZ1012	GAGCACATTGACTGGTGTGC	ATTCTGTAGGCCCATTGCAG	0:13037865
MZ371	AAGGTAATTTGCAGGGAGCA	CCAGGTTGATGGATGGAAGT	0:5124400
Labeo2016	CCCAAATTGTTCATCACCAT	CCCTGATCTTTGTCAAACCAG	21:2801036
MZ33953	CTATTGAGCCCCACATTGCT	TGTGCGTGGAGCTTTGTTAG	84:1877802
MZ31651	TTAACAGTGTAAGCGCAAACG	AAACTCCACCGCTGTCAAAC	73:2547017

Continued

Supplementary Table S3-1 (Continued)

Primer Name	Forward Primer Sequence	Reverse Primer Sequence	Genomic Location-Scaffold:Base
Somato17	AGAGCCATGTGCAATCAACA	TGACTTGAAAGATGGGCTGA	8:9213226
Somato20	GGTCTGGTGGCCAGGTAGTA	CCTCCTGCTCTGTTCCTCAG	8:9261286
MZ27256	ATACATGCAGCGAGCACAAA	TGGGCTGGTATGAAACACAA	54:1524636
MZ30207	TTTTGGTGTCTGTGCTTTGG	TAATTCCTCGCCCAATGTTC	67:547222
MZ52420	CCCACAGTTCTGGACAACCT	GAGCTGGCCAACTACTACGC	272:154037
MZ37607	GCACACGTCCTCTCTTCCTC	AAGAGGAAAGCACTGGGAGA	103:1963623
MZ13482	TGTCACACTGAGAGAAGGAAGAGA	TAAGAACGGTGCCTGTGTTG	16:5485770
MZ13525	TCCGTCCCTCTCTGTCTACC	GACTGAAAGCAGGAGGCTGT	16:6011743
MAP	AACCAATCAATGAGCTGGAA	GTGGAACTCACTGGCACACA	49:532113
SL9A7	TCAAGAGTCCAAGTGACACCA	AGCCTCAGCAGAAGAACCAG	49:462833
Syntaxin	AGCAGTGTGAGGACGGAGTT	TTCGTTAGTCAACGTGTGTGTG	49:732084
Bloc1S2	TCCACTGATGAGGTGAGGTT	ACAATGGCCAAAGGACAAAA	49:2903493

Supplementary Table S3-2. Candidate genes for identified QTL regions.

QTL Region	Gene	Function	Reference
Dorsal and Caudal Xanthophores	GAPDV1-GTPase-activating protein and VPS9 domain-containing protein 1	Interacts with Rab5 as a GEF and coordinates vesicle uncoating.	Stenmark H: **Rab GTPases as coordinators of vesicle traffic**. Nat *Rev Mol Cell Bio* 2009, **10**:513-525.
Dorsal and Caudal Xanthophores	STAR-Steroidgenic acute regulatory protein	Thought to be important in carotenoid binding and deposition.	Walsh N, Dale J, McGraw KJ, Pointer MA, Mundy, NI: **Candidate genes for carotenoid coloration in vertebrates and their expression profiles in the carotenoid-containing plumage and bill of a wild bird.** *Proc R Soc B* 2012, **279**: 58-66.
Dorsal and Caudal Xanthophores	TRPM6/-Transient receptor potential cation channel subfamily M member 6/1	Member of a gene family that has been previously linked to melanophore mutations in zebrafish.	Iuga AO, Lerner EA: **TRP-ing up melanophores: TRPM7, melanin synthesis, and pigment cell survival**. *J Invest Dermatol* 2007, **127**: 1855-1856.
Dorsal and Caudal Xanthophores	VP33A-Vacuolar protein sorting-associated protein 33A	Thought to be involved in melanosome biogenesis; causes pigment dilution mutant in mice.	Suzuki T, Oiso N, Gautam R, Novak EK, Panthier JJ, Suprabha PG, Vida T, Swank RT, Spritz RA: **The mouse organellar biogenesis mutant buff results from a mutation in *Vps33a*, a homologue of yeast *vps33* and *Drosophilia* carnation.** *PNAS* 2003, **100**: 1146-1150.

Continued

Supplementary Table S3-2 (Continuted)

QTL Region	Gene	Function	Reference
Dorsal and Caudal Xanthophores	RP3A- Rabphilin-3A	Effector of Rab27. Rab27a mutations in humans and mice results in melanosome transport defects.	Chavas LMG, Ihara K, Kawasaki M, Torii S, Uejima T, Kato R, Izumi T, Wakatsuki S: **Elucidation of Rab27 recruitment by its effectors: structure of Rab27a bound to Exophlin4/Slp2-a.** *Structure* 2008, **16**: 1468-1477.
Dorsal and Caudal Xanthophores	ADRB2- Beta-2-adernergic receptor	Causes pigment dispersion of melanophores.	Morishita F, Katayama H, Yamada K: **Subtypes of beta adrenergic receptors mediating pigment dispersion in chromatophores of the medaka, *Oryzias latipes***. *Comp Biochem Physiol* 1985, **81C:** 279-285.
Dorsal and Caudal Xanthophores	AAK1-AP2-associated protein kinase 1	Enhances stabilization of AP2 with plasma membrane and its derived vesicles.	Stenmark H: **Rab GTPases as coordinators of vesicle traffic**. *Nat Rev Mol Cell Bio* 2009, **10**:513-525.
Dorsal and Caudal Xanthophores	BCOD	Cleaves colorful carotenoids to colorless apocarotenoids; reduction of expression causes yellow skin in chickens.	Erikkson J, Larson G, Gunnarsson U, Bed'hom B, Tixier-Boichard M, Stromstedt L, Wright D, Jungerius A, Vereijken A, Randi E, Jensen P, Andersson L: **Identification of the *Yellow Skin* gene reveals a hybrid origin of the domestic chicken**. *PLoS Genet* 2008, **4**: e1000010.

Continued

Supplementary Table S3-2 (Continued)

QTL Region	Gene	Function	Reference
Dorsal and Caudal Xanthophores	ARRD3- Arrestin domain-containing protein 3	Can regulateβ2-adernergic receptors, which in turn can cause pigment dispersion.	Nabhan JF, Pan H, Lu Q: **Arrestin domain-containing protein 3 recruits the NEDD4 E3 ligase to mediate ubiquitination of the β2-adernergic receptor.** *EMBO reports* 2010, **11:** 605-611 Morishita F, Katayama H, Yamada K: **Subtypes of beta adrenergic receptors mediating pigment dispersion in chromatophores of the medaka, *Oryzias latipes***. *Comp Biochem Phys C* 1985, **81**: 279-285.
Dorsal and Caudal Xanthophores	IQGA--RAS GTPase-activating like protein IQGAP1	Expressed by human keratinocytes.	Press lauer S, Hinterhuber G, Cauza K, Horvat R, Rappersberger K, Wolff K, Foedinger D: **RasGAP-like Protein IQGAP1 is expressed by human keratinocytes and recognized by autoantibodies in association with bullous skin disease.** *J Invest Dermatol* 2003, **120**: 365-371.
Pelvic Melanophores	POU domain-Class3; transcription factor 1	Also known as BRN2. Thought to be involved in melanocyte differentiation and growth.	Cook AL, Boyle GM, Leonard JH, Parsons PG, Sturm RA: **BRN2 in melanocytic cell development, differentiation, and transformation.** In *From Melanocytes to Melanoma: the Progression to Malignancy.* Edited by Hearing VJ and Leon SPL. New Jersey: Humana Press; 2006: 149-167.

Continued

Supplementary Table S3-2 (Continued)

QTL Region	Gene	Function	Reference
Pelvic Melanophores	CXB5- Gap junction-beta 5 protein	Also known as Connexin 31.1. Shown to be expressed in human skin.	Aasen T, Kelsell DP: **Connexins in skin biology.** In *Connexins*. Edited by Harris AL and Locke D. New Jersey: Human Press; 2009: 307-321. Scemes E, Spray DC, Meda P: **Connexins, pannexins, innexins: novel roles of "hemi-channels".** *Pflugers Arch* 2009, **457**: 1207-1226.
Pelvic Melanophores	MEOX2- Homeobox protein MOX-2	Blocks Pax3-DNA interactions; Pax3 involved in melanogenesis.	Kubid JD, Young KP, Plummer RS, Ludvik AE, Lang D: **Pigmentation PAX-ways: the role of Pax3 in melanogenesis, melanocyte stem cell maintenance, and disease.** *Pigment Cell Melanoma Res* 2008, **21**: 627-645.
Pelvic Melanophores	DGKB- Diacylglycerol kinase beta	Can regulate levels of DAG; increased DAG increased melanin levels in cultured human melanocytes.	Martelli AM, Bortul R, Tabellini G, Bareggi R, Manzoli L, Narducci, P, Cocco L: **Diacylglycerol kinases in nuclear lipid-dependent signal transduction pathways.** *Cell Mol Life Sci* 2002, **59**: 1129-137. Park HY, Lee J, Gonzalez S, Middlekamp-Hup MA, Kapasi S, Peterson S, Gilchrest BA: **Topical application of a Protein Kinase C inhibitor reduces skin and hair pigmentation.** *J Invest Dermatol* 2004, **122:** 159-166.
Pelvic Melanophores	ARF3- ADP-ribosylation factor 3	Could possibly be involved in vesicle formation pathway.	Boman AL, Zhang C, Zhu X, Kahn, RA: **A family of ADP-ribosylation factor effectors that can alter membrane transport through the *trans*-Golgi**. *Mol Biol Cell* 2000, **11**: 1241-1255.

Appendix 3

Supplemental Table from Chapter 4

Supplementary Table S4-1. Primer sequences for all primers used in study.

Primer Name	Forward Primer Sequence	Reverse Primer Sequence
MZ9388	GAGCTATGCAAGTGCCGTTT	CTGGCTCCCAGCATAAAAAG
MZ9245	CTTCCAGTCTGGTGCCAATA	AGCTGAAGCGCTGTCTTGAT
UNH2012	AAGAGGGTGCTGCTTTTGAA	AGGCAACACCAACATCACTG
MZ8983	ACACAAGAGGGGATTTTGGA	CCCCCTATCTCAAGGACCAT
MZ8826	GGGCTATGACAGCTTCACAAA	TTCTGCAACATTGTCTTGACG
MZ8730	TGTGCTTATGCTTGCTCACC	ATCAATCTGTCTGCCCGAAC
UNH2162	CCCACGCCATGTCTAACAGT	AGCACCCTTTTGTTGTTTCG
AAK1_SNP1	GAGAAAGCACTTGGGGTCTG	AGCAAATTTCTGGCAGGAGA
AAK1_Gap259	CAAGAATGCAACTAAGAATATGAATG	CCAAAGTGAAACAAGCAAACC
AAK1_Gap260	TGGCCTTGCTGCAGATAATA	CAGGTGTATTGCCAGTAGCTGT
AAK1_Gap261	TTGCTGTACTCTATGAACGTCTTT	TAAGTGGCACTGCATTGGTG
AAK1_Gap259B	GCTGCACTCAAATGAATGTCC	AGTCCCACGACATGGTAAGG
AAK1_Gap259C	CAGTTTTCAGCACAGTTGGTG	CCCCACAGAATTCATTTGCT
AAK1_Gap261B	TTGCAAGGACACTCAAAAGC	*Use AAK1_Gap261 Reverse
AAK1_Gap261B	*Use AAK1_Gap261 Forward	TGGTCGCTGATTTGTTTTTG

Continued

Supplementary Table S4-1 (Continued)

Primer Name	Forward Primer Sequence	Reverse Primer Sequence
AAK1_Exon1-6	GCAGCTTCATCGGACGAG	ATGGTGAAACTGCCATCACA
AAK1_Exon7-12	TACGCTGTCATACCGTGCTC	TACGTCTGTGGCCACTCTTG
AAK1_Exon13-19	GCATCCAGTCTGCTGCAATA	TCGATGAGCTGATCCACCA
AAK1_Gap260Cl	GTGGCCTTGCTGCAGATAATA	*Use AAK1_Gap260B Reverse
AAK1_Gap261C	TAAAGCGATTCGTTGCACAT	*Use AAK1_Gap261 Reverse
AAK1_Gap261C	*Use AAK1_Gap261 Forward	AGACAACCTTGGCAGATCAAA
AAK1_Gap261D	CATTTTGCCAAAGGAAAGGA	*Use AAK1_Gap261 Reverse
AAK1_Gap261E	AATAAACTCAAACCAGACTGAGCA	*Use AAK1_Gap261 Reverse
AAK1_Gap261E	*Use AAK1_Gap261 Forward	ACTGCCTTGCTCAGTCTGGT
AAK1_Gap259_SNP	GAAGTGCTGAAGTGGACTGG	CCTCAACTGTTGCCTCAACC
AAK1_Gap259_SNPintb	TTGTCCTCTCTCCCTCCC	*Use AAK1_Gap259_SNP Reverse
AAK1_Gap259_GFc	*Use AAK1_Gap259_SNP Forward	AGCAACTTGTTGGGATCAGA
β-actin Exon 1-3	GCCGTCACACTCACAGCTT	TCTTCTCCCTGTTGGCTTTG
β-actin Exon 4-5	TGTTCGAGACCTTCAACACC	TTGATCTTCATGGTGGATGG
β-actin Exon 6	CCCCACCTGAGCGTAAATACT	TCGGAACACATGTGCACTTTA

Continued

Supplementary Table S4-1 (Continued)

Primer Name	Forward Primer Sequence	Reverse Primer Sequence
AAK1_Gap261F	TGTGAGCAACGGCACATTTA	*Use AAK1_Gap261 Reverse
AAK1_Gap261F	*Use AAK1_Gap261 Forward	ACAAAAGGCAGCCGCTAAA
AAK1_Gap261G	GAACTGAATGTTGACCCGAAG	*Use AAK1_Gap261 Reverse
AAK1_Gap261Cl	GTTGCTGTACTCTATGAACGTCTTT	GTAAGTGGCACTGCATTGGTG
AAK1_Exon2-6	GGTTTTGCCATAGTCTTTTTGG	CCATCTCTGGAGCACGGTAT
AAK1_Exon6-11	*Use AAK1_Exon7-12 Forward	CTGCTGCTGCTGCTTCATAA
AAK1_Exon12-15	*Use AAK1_Exon13-19 Forward	TGAGGCTGGGAATGAGATTT
AAK1_Exon16-19	AGGGACAGAAGTCCCCTGAT	*Use AAK1_Exon13-19 Reverse
AAK1_Gap261Cl_int	GAAGGAAAGGACGTTGCCTAA	*Use AAK1_Gap261Cl Reverse
AAK1_Exon1-3Z	TGGTGGCTGTGTTAGCTCTG	TCAGGACTTCCCACACATCA
AAK1_Exon1-3T	AGCTTCATCGGACGAGTCTT	*Use AAK1_Exon1-3Z Reverse
AAK1_Exon16-18Z2	AGGGACAGAAGTCCCCTGAT	GGAGCAGAGGAGCAAGAAGA
AAK1_Exon18-20Z2	CCTGTCATGGAAACCCATCT	AGCTGATCCACCAGCAGAAG
AAK1_Exon18-19Z1	TTCTTGCTCCTCTGCTCCTC	GGGGGCTATGTGACTCTCAA

Continued

Supplementary Table S4-1 (Continued)

Primer Name	Forward Primer Sequence	Reverse Primer Sequence
AAK1_Exon16-17T	GAAGCTGATTGAGGGACAGAA	ACTGGCAGGAATGAGCTCTG
AAK1_Exon17-18T	GCAGAGGTGTGTGTGGACTC	TTTTTGGAGATGAGCACTGG
AAK1_Exon18-19T	TCTGCTTTCCTCATGTCGTG	GAGCTGATCCACCAGCAGA
AAK1_Gap260_int	*Use AAK1_Gap260Cl Forward	GAACCGAGTCGTAAGTGTGGA

Appendix 4

Supplemental Tables from Chapter 5

Supplementary Table S5-1. Microsatellite primer sequences

Primer Name	Forward Primer Sequence	Reverse Primer Sequence	Genomic Location-Scaffold:Base
UNH2092	ACTACCAAGGTCTCCCGTTG	GGACTGCAATGTTGGTGCTA	93:1616439
UNH2199	GATCACTCGCAGCTAGGACA	TGAGTGCAAAAGACTGCTTGA	99:932675
UNH973	CACCTTCAAAGCAGCTGGTAA	AGTGCTCGGGGATAAAGTCA	0:15726268
MZ1012	GAGCACATTGACTGGTGTGC	ATTCTGTAGGCCCATTGCAG	0:13037865
MZ371	AAGGTAATTTGCAGGGAGCA	CCAGGTTGATGGATGGAAGT	0:5124400
MZ15715	GTGATGGCCTCCTGACATTT	CTTTTCATCAGCCAGGCTCT	21:725939
MZ15761	TCTCTTCACTGGCACCACAG	CAGGTCCTGAGCAGCCTAAG	21:1159614
MZ15764	GATTTCCCAGGCATTTGTTG	AGTGGAGACCAGGCTGAAAA	21:1328483
MZ15782	ACCAAGAACACCTTGGCATC	TTTTCTCCACCACGTTCCTC	21:1697011
MZ15808	CTCCAGCTGAACTAGCAGCA	TTTTGGAGCTACTGGTGAGGA	21:2201516
MZ15818	TACTCATTGGCCCTGACTGG	CATGTCTGCATGGAGCAACT	21:2381267
Labeo2016	CCCAAATTGTTCATCACCAT	CCCTGATCTTTTGTCAAACCAG	21:2801036
MZ15984	CGTCACTTGAGCACTTTTTCC	CCTTCTGATGCGTTCAGTTG	21:5091978
UNH896	CCTCTGTCCCTCCATGTGTT	AGCCTGGCTTTAGAGGCAAT	100:1113309
UNH2044	CTGCTGTCGGGGAAATTTTA	GGAAAGAATTGTGGCAGAGG	115:1910371

Supplementary Table S5-2. Gene sequencing primer sequences

Primer Name	Forward Primer Sequence	Reverse Primer Sequence
GSDF_Exon1_2	CAAGTTCAGAACCCCAAGGA	ACCAGACGTTCGAAACCAAC
GSDF_Exon2_4	AAGACTCAAGGGCTGGATCA	TCATGGCAGTCCATCAAAAA
GSDF_Exon3_4	TGTGTTGGCATCCTGGATATT	TCAAAACAGAGCGAGCTTGA
GSDF_Exon4_5	GTCAGCTGCATGGAGGAAAT	TGATCCAGCCCTTGAGTCTT
GSDF_Exon 5	AAAGCATGTCAGTCATATCACCA	TATGGATGAATCCAGCACCA
GSDF_EndEx5	GCTTTATTGTATAGCACTGTAAGCA	GGCATGACTTTCCATTGGTT
GSDF_Down1	GTTGAACGGCTGCTGTCTTA	TGTTTGTTTGCAAATGTATGCTT
GSDF_Down2	CCGACAGCCAGCAGAAGT	GTCACAAGTATGCGCCACTG
GSDF_Ex1Up	CCAAGACAAATGCCATCATC	ACTGCCCTCAGAGTGAGCTG
GSDF_Prom1	TGCAGTCCATTAAACATCCTG	AAACTGTTTTCTACTTCAATTTGTTTG
GSDF_Prom2	GGCAGCAGAATTTTAAGCACA	TTTAAAAGCCGAATGGGAAG
GSDF_Prom4	TGGATCAAATGATGGCTTTTC	AAATGAGGCCTCTTCTTCACA
GSDF_Prom5	CTGTTGTGTGTGTTTATCTGGAA	TCGCATGACCCACAGAACT
GSDF_Prom6	GCTGCAGGTGATGTTAATATCTGA	AAATTATGTTCACAGACCTTTGG

Continued

Supplementary Table S5-2 (Continued)

Primer Name	Forward Primer Sequence	Reverse Primer Sequence
GSDF_Prom7	TTTTAATTTGTTGGGGCAAA	CCGCCATGTGGCAATACA
GSDF_Prom8	TGGGTGTTGTAACAGTTGCTG	TTTGTAACACTATTAATGTCTCTCAGG
GSDF_Prom9	TGCAATCCAATAGTCTTGCTG	GACAAAGATACTTTCACGCAGAA
GSDF_Prom3(b)	CAGGCTGTCCTCAAAACTCC	GGCTATTGCAGCATCACTCA
GSDF_Prom8-9	AGGTGTGGCAGCATTTAACA	CACAACAACACACCAGCAAG
GSDF_Prom7-8	CACGAGAAGCATTTTTGTATGC	CATCAGTGCGCTGGACTTATTA
GSDF_Prom6-7	CCTCAGTTCACAGTTTTGGAA	CACAGAAAGTACTTGTTTGCACAT
GSDF_Prom5-6	CCATTGACAGACCCCAGTTT	AGACAGCCATTTAACTTTGTGC

Bibliography

Alberson RC, Streelman JT, Kocher TD. 1999. Directional selection has shaped the oral jaws of Lake Malawi cichlid fishes. *PNAS* 100:5252-5257.

Allender CJ, Seehausen O, Knight ME, Turner GF, Maclean N. 2003. Divergent selection during speciation of Lake Malawi cichlid fishes inferred from parallel radiations in nuptial coloration. *PNAS* 100:14074–14079.

Arduini BL, Henion PD. 2004. Melanophore sublineage-specific requirement for zebrafish *touchtone* during neural crest development. *Mech of Develop* 121:1353-1364.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3(10):e3376.

Barson NJ, Knight ME, Turner GF. 2007. The genetic architecture of male colour differences between a sympatric Lake Malawi cichlid species pair. *J Evol Biol* 20:45–53.

Beldade P, Brakefield PM. 2002. The genetics and evo-devo of butterfly wing patterns. *Nature Rev Gen* 3:442–452.

Bennett DC, Lamoreux ML. 2003. The color loci of mice-a genetic century. *Pigment Cell Res* 16:333-344.

Braash I, Schartl M, and Volff J. 2007. Evolution of pigment synthesis pathways by gene and genome duplication in fish. *BMC Evol Biol* 7:74.

Broman KW, Wu H, Sen S, Churchill GA. 2003. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19:889-890.

Budi EH, Patterson LB, Parichy DM. 2008. Embryonic requirements of ErbB signaling in neural crest development and adult pigmentation pattern formation. *Development* 135:2603-2614.

Budi EH, Patterson LB, Parichy DM. 2011. Post-embryonic nerve-associated precursors to adult pigment cells: genetic requirements and dynamics of morphogenesis and differentiation. *PLOS Genet* 7:e1002044.

Candolin U. 2003. The use of multiple cues in mate choice. *Biol Rev* 78:575-595.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. 2011. Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics* 1:171-182

Cheverud JM. 2000. Detecting epistasis among quantitative trait loc. In: Wolf JB, Brodie ED III, Wade MJ, editors. Epistasis and the evolutionary process. New York: Oxford University Press. p 58–81.

Conner SD, Schmid. 2002. Identification of an adaptor-associated kinase, AAK1, as a regulator of clathrin-mediated endocytosis. *J Cell Biol* 156:921-929.

Curran K, Lister JA, Kunkel GR, Prendergast A, Parichy DM, Raible DW. 2010. Interplay between Foxd3 and Mitf regulates cell fate plasticity in the zebrafish neural crest. *Dev Biol* 344:107-118.

Danley PD, Kocher TD. 2001. Speciation in rapidly diverging systems: lessons from Lake Malawi. *Mol Ecol* 10:1075-1086.

Devlin RH, Nagahama Y. 2002. Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. *Aquaculture* 208:191-364.

Dieffenbach CW, Dveksler GS. 2003. PCR primer: a laboratory manual. Cold Spring Harbor Laboratory Press.

Dominey WJ. 1984. Effects of sexual selection and life history on speciation: species flocks in African cichlids and Hawaiian Drosophila. In: Echelle A, Kornfield I, editors. In evolution of fish species flocks. Orono: University of Maine at Orono Press. p 231–249.

Ezaz T, Stiglec R, Veyrunes F, Marshall Graves JA. 2006. Relationships between vertebrate ZW and XY sex chromosome. *Curr Biol* 16:736-743.

Falconer DS, Mackay TFC. 1996. Introduction to quantitative genetics. Burnt Mill, Harlow, UK: Longman.

Fedorova L, Fedorov A. 2003. Introns in gene evolution. *Genetica* 118:123-131.

Frohnhofer HG, Krauss J, Maischein HM, Nusslein-Volhard C. 2013. Iridophores and their interactions with other chromatophores are required for stripe formation in zebrafish. *Development* 140:2997-3007.

Fukamachi S, Sugimoto M, Mitani H, Shima A. 2004. Somatolactin selectively regulates proliferation and morphogenesis of neural-crest derived pigment cells in medaka. *PNAS* 101:10661-10666.

Gautier A, Le Gac F, Lareyre JJ. 2011a. The *gsdf* gene locus harbors evolutionary conserved and clustered genes preferentially expressed in fish previtellogenic oocytes. *Gene* 472:7-17.

Gautier A, Sohm F, Joly JS, Le Gac F, Lareyre JJ. 2011b. The proximal promoter region of the zebrafish *gsdf* gene is sufficient to mimic the spatio-temporal expression pattern of the endogenous gene in sertoli and granulosa cells. *Biol Reprod* 85:1240-1251.

Genner MJ, Seehausen O, Lunt DH, Joyce DA, Shaw PW, Carvalho GR, Turner GF. 2007. Age of cichlids: new dates for ancient lake fish radiations. *Mol Biol Evol* 24:1269–82.

Guyon R, Rakotomanga M, Azzouzi N, Coutanceau JP, Bonillo C, D'Cotta H, Pepey E, Soler L, Rodier-Goud M, D'Hont A, Conte MA, van Bers N, Penman DJ, Hitte C, Crooijmans R, Kocher TD, Ozouf-Costaz C, Baroiller JF, Gailbert F. 2012. A high-resolution map of the Nile tilapia genome: a resource for studying cichlids and other percomorphs. *BMC Genomics* 13:222.

Hearing VJ. 2005. Biogenesis of pigment granules: a sensitive way to regulate melanocyte function. *J Dermatol Sci* 37:3-14.

Hoekstra HE. 2006. Genetics, development and evolution of adaptive pigmentation in vertebrates. *Heredity* 97:222-234.

Ijiri S, Kaneko H, Kobayashi T, Want DS, Sakai F, Paul-Prasanth B, Nakamura M, Nagahama Y. 2008. Sexual dimorphic expression of genes in gonads during early differentiation of a teleost fish, the Nile tilapia *Oreochromis niloticus*. *Biol Reprod* 78:333-341.

Iuga AO, Lerner EA. 2007. TRP-ing up melanophores: TRPM7, melanin synthesis, and pigment cell survival. *J Invest Dermatol* 127: 1855-1856

Iwashita M, Watanabe M, Ishii M, Chen T, Johnson SL, Kurachi Y, Okada N, Kondo S. 2006. Pigment pattern in *jaguar*/*obelix* zebrafish is caused by a Kir7.1 mutation: Implications of the regulation of melanosome movement. *PLoS Genet* 2:1861-1870.

Kelsh RN, Brand M, Jiang YJ, Heisenberg CP, Lin S, Haffter P, Odenthal J, Mullins MC, van Eeden FJ, Furutani-Seiki M, Granato M, Hammerschmidt M, Kane DA, Warga RM, Beuchle D, Vogelsang L, N¨usslein-Volhard C. 1996. Zebrafish pigmentation mutations and the processes of neural crest development. *Development* 123:369–389.

Kelsh, RN. 2004. Genetics and Evolution of Pigment Patterns in Fish. *Pigment Cell Res* 17: 326-336.

Kelsh, RN, Harris, ML, Colanesi, S, Erickson, CA. 2009. Stripes and belly-spots—A review of pigment cell morphogenesis in vertebrates. *Seminars in Cell & Dev Biol* 20:90-104.

Kent J, Wheatley SC, Andrews JE, Sinclair AH, Koopman P. 1996. A male-specific role for *SOX9* in vertebrate sex determination. *Development* 122:2813-2822.

Kidd MR, Danley PD, Kocher TD. 2006. A direct assay of female choice in cichlid fishes: all the eggs in one basket. *J Fish Biol* 68:373–384.

Knight RD, Nair S, Nelson SS, Afshar A, Javidan Y, Geisler R, Rauch GJ, Schilling TF. 2004. *Lockjaw* encodes a zebrafish *tfap2a* required for early neural crest development. *Development* 130:5755-5768.

Kocher TD. 2004. Adaptive evolution and explosive speciation: the cichlid fish model. *Nature Rev Gen* 5:288–298.

Konings A. 2007. Malawi cichlids in their natural habitats. El Paso, TX: Cichlid Press.

Lee BY, Penman DJ, Kocher TD. 2003. Identification of a sex-determining region in Nile tilapia (*Oreochromis niloticus*) using bulked segregant analysis. *Anim Genet* 34:379-383.

Lee BY, Hulata G, Kocher TD. 2004. Two unlinked loci controlling the sex of blue tilapia (*Oreochromis aureus*). *Heredity* 92:543-549.

Lee BY, Lee WJ, Streelman JT, Carleton KC, Howe AE, Hulata G, Slettan A, Stern JE, Terai Y, Kocher TD. 2005. A second-generation genetic linkage map of tilapia (*Oreochromis* spp.). *Genetics* 170:237-244.

Lister JA, Robertson CP, Lepage T, Johnson SL, Raible DW. 1999. *Nacre* encodes a zebrafish microphthalmia-related protein that regulates neural-crest derived pigment cell fate. *Development* 126:3757–3767.

Lopes SS, Yang X, Muller J, Carney TJ, McAdow, AR, Rauch G, Jacoby AS, Hurst JD, Delfino-Machin M, Haffter P, Geisler R, Johnson SL, Ward A, Kelsh RN. 2008. Leukocyte tyrosine kinase functions in pigment cell development. *PLoS Genet* 4: e1000026.

Luckenbach JA, Iliev DB, Goetz FW, Swanson P. 2008. Identification of differentially expressed ovarian genes during primary and early secondary oocyte growth in coho salmon, *Oncorhynchus kisutch*. *Reprod Biol Endocrin* 6:2.

Lynch M, Walsh B. 1998. Genetics and analysis of quantitative traits. Sunderland, MA: Sinauer.

Magalhaes IS, Seehausen O. 2010. Genetics of male nuptial colour divergence between sympatric sister species of a Lake Victoria cichlid fish. *J Evol Biol* 23:914–924.

Manolakou P, Lavranos G, Angelopoulou R. 2006. Molecular patterns of sex determination in the animal kingdom: a comparative study of the biology of reproduction. *Reprod Biol Endocrin* 4:59.

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TFC, McCarroll SA, Visscher PM. 2009. Find the missing heritability of complex diseases. *Nature* 461:747-753.

Marks MS, Seabra MC. 2001. The melanosome: membrane dynamics in black and white. *Nat Rev Mol Cell Bio* 2:1-11.

Marshall Graves JA. 2008. Weird animal genomes and the evolution of vertebrate sex and sex chromosomes. *Annu Rev Genet* 42:565-586.

Mazurais D, Montfort J, Delalande C, Le Gac F. 2005. Transcriptional analysis of testis maturation using trout cDNA macroarrays. *Gen Comp Endocr* 142:143-154.

McNeill MS, Paulsen J, Bonde G, Burnight E, Hsu MY, Cornell RA. 2007. Cell death of melanophores in zebrafish *trpm7* mutant embryos depends on melanin synthesis. *J Invest Dermatol* 127:2020-2030.

Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA. 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res* 17:240-248.

Mills MG, Nuckels RJ, Parichy DM. 2007. Deconstructing evolution of adult phenotypes: genetic analyses of *kit* reveal homology and evolutionary novelty during adult pigment pattern development of *Danio* fishes. *Development* 134:1081–1090.

Mills MG, Patterson LB. 2009. Not just black and white: Pigment pattern development and evolution in vertebrates. *Semin Cell Dev Biol* 20:72–81.

Minchin JEN, Hughes SM. 2008. Sequential actions of Pax3 and Pax7 drive xanthophores development in zebrafish neural crest. *Dev Biol* 317:508-522.

Muske LE, Fernald RD. 1987. Control of a teleost social signal II. Anatomical and physiological specializations of chromatophores. *J Comp Physiol A* 160:99–107.

Myosho T, Otake H, Masuyama H, Matsuda M, Kuroki Y, Fujiyama A, Naruse K, Hamaguchi S, Sakaizumi M. 2012. Tracing the emergence of a novel sex-determining gene in Medaka, *Oryzias luzonensis*. *Genetics* 191:163-170.

Nakamasu A, Takahaski G, Kanbe A, Kondo S. 2009. Interactions between zebrafish pigment cells responsible for the generation of Turing patterns. *PNAS* 106:8429-8434.

Navarro RE, Ramos-Balderas JL, Guerrero I, Pelcastre V, Maldonado E. 2008. Pigment dilution mutants from fish models with connection to lysosome-related organelles and vesicular traffic genes. *Zebrafish* 5:309-318.

Navarro-Martin L, Galay-Burgos M, Sweeney G, Piferrer F. 2009. Different *sox17* transcripts during sex differentiation in sea bass, *Dicentrarchus labrax*. *Mol Cell Endocrinol* 299:240-251.

Nery LM, Castrucci AMD. 1997. Pigment cell signaling for physiological color change. *Comp Biochem Physiol* 118:1135-1144.

O'Quin CT, Drilea AC, Roberts RB, Kocher TD. 2012. A small number of genes underlie male pigmentation traits in Lake Malawi cichlid fishes. *J Exp Zool B* 318:199-208.

O'Quin CT, Drilea AC, Conte MA, Kocher TD. 2013. Mapping of pigmentation QTL on an anchored genome assembly of the cichlid fish, *Metriaclima zebra*. *BMC Genomics* 14:287.

Odenthal J, Rossnagel K, Haffter P, Kelsh RN, Vogelsang L, Brand M, van Eeden FJM, Furutani-Seiki M, Granato M, Hammerschmidt M, Heisenberg C, Jiang Y, Kane DA, Mullins MC, Nusslein-Volhard C. 1996. Mutations affecting xanthophore pigmentation in the zebrafish, *Danio rerio*. *Development* 123:391–398.

Parichy DM, Rawls JF, Pratt SJ, Whitfield TT, Johnson SL. 1999. Zebrafish *sparse* corresponds to an orthologue of c-*kit* and is required for the morphogenesis of a subpopulation of melanocytes, but is not essential for hematopoiesis or primordial germ cell development. *Development* 126:3425–3436.

Parichy DM, Ransom DG, Paw B, Zon LI, Johnson SL. 2000a. An orthologue of the *kit*-related gene *fms* is required for development of neural crest-derived xanthophores and a subpopulation of adult melanocytes in the zebrafish, *Danio rerio*. *Development* 127:3031–3044.

Parichy DM, Mellgren EM, Rawls JF, Lopes SS, Kelsh RN, Johnson SL. 2000b. Mutational analysis of *Endothelin receptor b1* (rose) during neural crest and pigment pattern development in the zebrafish *Danio rerio*. *Dev Biol* 227:294–306.

Parichy DM, Johnson SL. 2001. Zebrafish hybrids suggest genetic mechanisms for pigment pattern diversification in Danio. *Dev Genes Evol* 211: 319-328.

Parichy DM, Turner JM. 2003a. Temporal and cellular requirements for Fms signaling during zebrafish adult pigment pattern development. *Development* 130:817-833.

Parichy DM, Turner JM, Parker NB. 2003b. Essential role for *puma* in development of post-embryonic neural crest-derived cells in zebrafish. *Dev Biol* 256:221–241.

Parichy DM, Reedy MV, Erickson CA. 2006a. "Regulation of melanoblast migration and differentiation." Chapter 5, in : The Pigmentary System and its Disorders, 2nd Edtion (Nordland, J.J., Boissy, R.E., Hearing, V.J., King, R.A., and Ortonne, J.P., Eds.). Oxford.

Parichy DM. 2006b. Evolution of danio pigment pattern development. *Heredity* 97:200–210.

Parichy DM, 2007. Homology and evolution of novelty during *Danio* adult pigment pattern development. *J Exp Zool (Mol Dev Evol)* 308b:578-590.

Paripatananont T, Tangtrongpairoj J, Sailasuta A, Chansue, N. 1999. Effect of astaxanthin on the pigmentation of goldfish *Carassius auratus*. *J World Aquac Soc* 30:454-460.

Patterson LB, Parichy DM. 2013. Interactions with iridophores and the tissue environment required for patterning melanophores and xanthophores during zebrafish adult pigment stripe formation. *PLOS Genet* 9: e1003561.

Pieau C. 1996. Temperature variation and sex determination in replites. *Bioessays* 18:19-26.

Pierce B. 2010. Genetics: a conceptual approach. New York, NY:W.H. Freeman and Company.

Quigley IK, Manuel JL, Roberts RA, Nuckels RJ, Herrington ER, Mac-Donald, EL, Parichy DM. 2005. Evolutionary diversification of pigment pattern in *Danio* fishes: differential *fms* dependence and stripe loss in *D. albolineatus*. *Development* 132:89–104.

R Development Core Team 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. See http://www.R-project.org

Rasband WS. 2009. Image J. U. S. National Institutes of Health, Bethesda, MD. See http://rsb.info.nih.gov/ij/.

Roberts RB, Ser J, Kocher TD. 2009. Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlids. *Science* 326:998–1001.

Rondeau EB, Messmer AM, Sanderson DS, Jantzen SG, von Schalburg KR, Minkley DR, Leong JS, Macdonald GM, Davidsen AE, Parker WA, Mazzola RSA, Campbell B, Koop BF. 2013. Genomics of sablefish (*Anoplopoma fimbria*): expressed genes, mitochondrial phylogeny, linkage map and identification of a putative sex gene. *BMC Genomics* 14:452.

Sarre SD, Georges A, Quinn A. 2004. The ends of a continuum: genetic and temperature dependent sex determination in reptiles. *Bioessays* 26:639-645.

Sauka-Spengler T, Bronner-Fraser M. 2008. A gene regulatory network orchestrates neural crest formation. *Nat Rev Mol Cell Bio* 9:557-566.

Sawatari E, Shikina S, Takeuchi T, Yoshizaki. 2007. A novel transforming growth factor-β superfamily member expressed in gonadal somatic cells enhances primordial germ cell and spermatogonial proliferation in rainbow trout (*Oncorhynchus mykiss*). *Dev Biol* 301:266-275.

Schmahl J, Capel B. 2003. Cell proliferation is necessary for the determination of male fate in the gonad. *Dev Biol* 258:264-276.

Seehausen O, van Alphen JJM. 1998. The effect of male coloration on female mate choice in closely related Lake Victoria cichlids (*Haplochromis nyererei*) complex. *Behav Ecol Sociobiol* 42:1–8.

Ser JR, Roberts RG, Kocher TD. 2010. Multiple interacting loci control sex determination in Lake Malawi cichlid fish. *Evolution* 64:486-501.

Seutin G, White BN, Boat PT. 1991. Preservation of avian blood and tissue samples for DNA analysis. *Can J Zoolog* 69:82-90.

Stuart-Fox D, Moussalli A, Whiting MJ. 2008. Predator-specific camouflage in chameleons. *Biol Lett* 4:326-329.

Sugimoto M. 2002. Morphological color changes in fish: regulation of pigment cell density and morphology. *Microsc Res Techniq* 58:496-503.

Summers K, Symula R, Clough M, Cronin T. 1999. Visual mate choice in poison frogs. *Proc. R. Soc. Lond. B.* 266:2141-2145.

Svetic V, Hollway GE, Elworthy S, Chipperfield TR, Davison C, Adams RJ, Eisen JS, Ingham PW, Currie PD, Kelsh RN. 2007. Sdf1a patterns zebrafish melanophores and links the somite and melanophore pattern defects in *choker* mutants. *Development* 134:1011-1022.

Takahashi T, Sota T, Hori M. 2013. Genetic basis of male colour dimorphism in a Lake Tanganyika cichlid fish. *Mol Ecol* 22:3049-3060.

Takimoto G, Higashi M, Yamamura N. 2000. A deterministic genetic model for sympatric speciation by sexual selection. *Evolution* 54:1870–1881.

Turner GF, Burrows MT. 1995. A model of sympatric speciation by sexual selection. *Proc R Soc London Sect B* 260:287–292.

Turner GF, Seehausen O, Knight ME, Allender CJ, Robinson RL. 2001. How many species of cichlid fishes are there in African lakes? *Mol Ecol* 10:793–806.

Van Ooijen JW, Voorrips RE. 2001. JOINMAP 3.0: Software for the calculation of genetic linkage maps. Wageningen, The Netherlands: Plant Research International.

Van Oppen MJH, Turner GF, Robinson CRRL, Deutsch JC, Genner MJ, Hewitt GM. 1998. Assortative mating among rock-dwelling cichlid fishes supports high estimates of species richness from Lake Malawi. *Mol Ecol* 7:991–1001.

Watanabe M, Iwashita M, Ishii M, Kurachi Y, Kawakami A, Kondo S, Okada N. 2006. Spot pattern of leopard *Danio* is caused by mutation in the zebrafish *connexin41.8* gene. *EMBO Rep* 7:893-897.

Zeng Z-B. 1992. Correcting the bias of Wrights's estimates of the number of genes affecting a quantitative character: A further improved method. *Genetics* 131:987–1001.

Zhang H, Somasundaram K, Peng Y, Tian H, Zhang H, Bi D, Weber BL, El-Deiry WS. 1998. BRCA1 physically associates with p53 and stimulates its transcriptional activity. *Oncogene* 16:1713-1721.