NATURAL HYBRIDIZATION AND MELANOMA IN SWORDTAIL FISH: GENETIC BASIS AND SELECTIVE MECHANISMS

A Dissertation

by

MATEO GARCIA OLAZABAL

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Gil G. Rosenthal
Co-Chair of Committee,	Manfred Schartl
Committee Members,	Kirk Winemiller
	Heath Blackmon
Head of Department,	Alex Keene

December 2021

Major Subject: Ecology and Evolutionary Biology

ABSTRACT

Understanding how genotype maps to phenotype is critical to understanding how evolution can generate and maintain biological diversity. Hybridization provides a great resource to pinpoint genes of interest, disentangle multivariate effects on fitness, and measure evolutionary change in real time. New mutations arising in diverging species can interact negatively in hybrids, generating what is known as hybrid incompatibilities. A famous example of hybrid incompatibility comes from distantly related laboratory hybrids between *Xiphophorus* species that cause malignant melanoma. This study utilizes an ongoing hybridization process in natural conditions in other *Xiphophorus* species to study the genetic basis and evolutionary persistence of melanoma in the wild.

Xiphophorus birchmanni is polymorphic for a coloration pattern on its caudal fin called spotted caudal (Sc). *X. malinche* lacks this pattern. *X. birchmanni – X. malinche* hybrids are also polymorphic regarding Sc and its expression can vary from a few black spots to extremely malignant melanoma. A study of juvenile vs adult Sc frequencies suggests that the phenotype is under natural selection in high incidence hybrid populations and therefore I tested whether sexual selection contributes to its maintenance. Visual mate choice trials showed that neither *X. birchmanni* nor hybrid females prefer spotted over non-spotted individuals. Future studies should further characterize environmental factors or other traits associated with Sc that might be favored by sexual selection via mate choice or intrasexual competition.

To identify the genetic basis of the trait, I performed a Genome Wide Association Study in a *X. birchmanni* population and determined that a previously known potent oncogene (*xmrk*) is responsible for driving the expression of the pattern. This was followed by a population ancestry

and admixture mapping study that proposed *adgre5* and *xmrk* as the genes responsible for the hybrid incompatibility causing melanoma. I performed functional cell culture and transgenic experiments to determine that *adgre5* acts as a tumor suppressor gene.

As far as I know, this is the only study that combines behavior, genomics and molecular biology techniques in an integrative approach to identify and functionally test a hybrid incompatibility (melanoma) to the single gene level in naturally occurring hybridizing species.

DEDICATION

Para Clementina, mi persona preferida en el mundo

ACKNOWLEDGEMENTS

I would like to thank Dr. Gil Rosenthal and Dr. Manfred Schartl for all the interesting science discussions and lessons, for showing me how to lead a group, and for all their guidance and support during my doctorate. I thank my committee members Dr. Kirk Winemiller and Dr. Heath Blackmon for their constructive feedback and ideas for my dissertation. I thank Dr. Molly Schumer for all her bioinformatics guidance and overall mentorship. I thank Dr. Svenja Meierjohann for welcoming me to her lab and all her cell culture support.

Thanks to my fellow Rosenthal lab members, Dr. Pablo Delclos, Dr. Chris Holland, Dr. Gaston Jofre, Dr. Daniel Powell, Carolyn Henderson, Megan Exnicios, Stephen Bovio, Amanda Beckman, Owen Dorsey, Emma Lehmberg, Paola Fascinetto-Zago, Faith Hardin, Giorgia Rando, Irene Andermarcher, for making field and lab work much more interesting and fun. I am especially grateful for the help I got from undergrad students Jeffery White, Nicole Guentzel, Will Ledbetter, and Anabel Perry, and all the help from the CICHAZ staff members: Don Clemente and family. Special thanks to the members of the Schartl lab and the Meierjohann Lab at Wurzburg University, Dr. Mateus Adolfi, Dr. Du Kang, Brigitta, Anita, Conny, Jenny, Petra, Ivan, Georg, Petra, Rupesh, Taf, Christina, for welcoming and teaching me all the nuances of cell culture, cloning and transgenics. I would also like to thank the administrative staff at Texas A&M, Jennifer Bradford, Will Bailey and Jason Martina for helping me navigate the endless bureaucracy.

This work would have not been possible without all the support from several groups of friends that always listened and cared about my work. Thanks to the great Old School gang: Pulga, Javi, Fede, Igna, for always being there even from a faraway distance. Thanks to the amazing Los Pajeros gang: Rui, James, Ross, Braeden, Rhys, Alex, Gilberto, Matt, and Phil. I am super grateful

this doctorate brought me to Texas and we all met. Special thanks to Gilberto “for checking my thesis”. I have to thank all my Germany friends: Jolle, Volker, Denis, Nico, Nadine, Carolina, Falk, Phillip, Marie, Mateus, Kang, Rupesh, Mboromosoro, Torben and specially Rodrigo and Santiago. I will never forget our glorious times in the bunker with Rodri and Santi. Thanks to the Ball Fondlers: Pablo, Luke, David and Andrew for all the good times. I also have to thank Texas A&M, Austin Huns, WRK and Neckarsulmer, Rugby Clubs for accepting me and looking after me. You were the greatest escape and work life balancers I could ask for. Bettina Tassino, thank you for always being my trustworthy counselor, as well as my Uruguayan scientific friends: Matias, Passer, Caco, Pato, Masche, Manta, Pupi.

This work would have been impossible without my and Clemen’s (which I consider my own) family. Regardless of the tough times we had to go through you were always a lighthouse to which I turned to for guidance, support and comfort. I am so proud of my new nephews Lorenzo and Geronimo that taught me how undeniable evolution is: holding a subset of your genes in the next generation is definitely a mind-blowing experience.

Finally, this work would have not been possible without the help of my dearest person in the world: Clementina. It is not even the endless hours of work you put into my research project, the travel all around the world to visit me, the long-distance relationship we had to endure. It is the lovely sensation that, together, anything seems possible; and the beautiful excitement I feel when I think about our life together ahead of us. Thank you for being a role model, for pushing me to be a better person and for being the greatest companion I could ask for!

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of my advisor, Professor Gil Rosenthal of the Department of Biology, my co-advisor Professor Manfred Schartl of the University of Wurzburg, Professor Heath Blackmon of the Department of Biology, and Professor Kirk Winemiller of the Department of Wildlife and Fisheries Sciences.

The genotyping libraries used in Chapter II were provided by Dr. Molly Schumer of the Department of Biology at Stanford University. All other work conducted for the dissertation was completed by the student independently.

Funding sources

Graduate study was supported by the Fulbright- ANII doctorate scholarship from Uruguay, and the HEEP Graduate Student Fellowship from the Hagler Institute at Texas A&M

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

Understanding the ultimate and proximate mechanisms that maintain genetic variation for discrete traits in a population remains a challenge for evolutionary biology. In particular, the evolutionary mechanisms underlying the maintenance of color polymorphism have been studied for a long time, and yet color polymorphism still remains to be explained in most studied groups (Gray and McKinnon, 2007). Because variation in color can affect both natural and sexual selection, it is of particular evolutionary importance. The study of the causes and consequences of color variation transcends behavioral, developmental, physiological, and ecological biology. Most importantly, since pigmentation traits have historically been the most tractable at the genetic level, they provide an excellent opportunity to study molecular evolution (Kronforst et al., 2012). Most of our knowledge comes from well-established animal models such as *Drosophila*, butterflies, spiders, sticklebacks, cave fish and lizards, but there are plenty of emerging systems. *Xiphophorus* is one of them.

Drosophila pigments consist of dark melanins and light sclerotins. Different physiological studies of their pigmentation patterns and their association with thermoregulation (Brisson et al., 2006; Gibert et al., 1999), resistance to desiccation (Jacobs, 1985; Wittkopp et al., 2011) and susceptibility to nematode infection (Dombeck and Jaenike, 2004; Jacobs, 1985) have provided conflicting observations suggesting the relation between pigmentation and physiology is complex and that it may depend not on the color itself but on their role as structural components of

the insect cuticle. *Drosophila*'s pigmentation genes were discovered thanks to decades of classical *Drosophila* genetics and biochemical approaches (Lindsley et al., 1968) (reviewed in (Wright, 1987)). In order to produce any pigment, first, tyrosine must be converted to dihydroxyphenylalanine (Dopa) by the tyrosine hydroxylase enzyme encoded by the *pale* gene. Dopa can be then converted to black melanin by extracellular enzymes encoded by the *yellow* gene family or into dopamine which will then become brown melanin (*yellow* gene family) or used to produce light pigments. The *ebony* gene converts dopamine to N-β-alanyldopamine (NBAD), the precursor of yellow sclerotin, while the *tan* gene reverses this synthesis. Colorless sclerotins can be produced when the family of dopamine-acetyl-transferases (DATS) convert dopamine to N-acetyl dopamine (NADA). This nonlinear (reversible) pathway implies that multiple genetic paths can result in the same phenotype. For example, since dopamine is the precursor to both brown melanin and yellow sclerotin, darker pigmentation can be produced either by increasing the synthesis of dark pigments or by decreasing the synthesis of light pigments. Moreover, because *Drosophila* pigments are secreted by epithelial cells in a cell-autonomous manner, color patterns also depend on the spatial regulation of the pigment enzymes. Indeed, many evolutionary independent correlations between gene expression and pigmentation pattern exist (Koopp, 2009). In the Kronforst et al. (2012) review the authors summarize and compare the many diverse studies concerning genetic changes responsible for the different color pattern variation across different species and populations to conclude that (i) there is overlap in the genes that control color patterns but sets of genes are not always the same across species, (ii) not all pigmentation genes contribute to pigmentation differences in every species, and (iii) major evolutionary changes can be caused by loci unknown to the pigmentation pathway. However, they also state

that in general pigmentation differences are controlled by multiple loci, and that multiple mutations contribute to the overall effect of each locus, suggesting that pigmentation evolution in *Drosophila* lacks major mutations. Instead, the accumulation of many subtle mutations is responsible for most differences within and between species. In fact, the evolution of *Drosophila* pigmentation is mostly associated with cis-regulatory changes (Jeong et al., 2008; Wittkopp et al., 2009). This could be related to the fact that many intermediate metabolites in the pigmentation pathway are also neurotransmitters (True, 2003; True et al., 2005), and therefore variations in the activity of these enzymes have pleiotropic effects on nervous system function and behavior. It is interesting how even among closely related species there exists a great heterogeneity among their color determination mechanisms, indicating very little evolutionary constraint.

Pigment variation in lizards has a tremendous ecological importance: it is imperative for camouflage (Rosemblum, 2006), can serve as antipredator warnings (Savage and Slowinski, 1992), it can play a role in intra specific communication (Chan et al., 2009) and thermoregulation (Clusella-Trullas et al., 2007). This has led to a wide range of variation in color across individuals, populations and species, which can help us understand genetic mechanisms behind ecologically important color traits. Reptile color patches are determined by the interaction of three types of cells containing different pigments (Bagnara and Hadley, 1973): yellow and orange coloration in the xanthophores, blue and green colors depending on the reflective properties of the iridophores and darker color patches coming from melanophores containing melanin. The study of the genetic basis of the blanched coloration of the White sand lizards provides an interesting example of rapid genetic convergent evolution. Unlike other examples of convergent evolution studies (Nosil and Sandoval, 2008; Steiner et al., 2009), this case involves three distantly related species that are exposed to the same, recently formed, selective environment (Rosemblum and Harmon,

2011). Each species presents a particular *Mc1r* mutation, originally believed to be associated with the blanched phenotype (Rosenblum et al., 2004) but later determined by functional studies that this is only the case in two of the three species. Moreover, in each of those species, the mutation acts differently: in one case, *Mc1r* is compromised in its ability to integrate into the cell membrane, resulting in a dominant blanched phenotype, and in the other case, *Mc1r* is unable to transmit the signal, resulting in a recessive blanched phenotype (Rosenblum et al., 2010). Allelic dominance is going to affect the ability of natural selection to act upon them (Orr, 2010) and thus can affect the distribution of adaptive alleles in nature. This is an interesting example of how understanding the genetic architecture of coloration highlights important similarities and differences across species during a rapid convergent evolution scenario.

Drosophila and lizards are good systems to study the genetic basis of coloration, not only because their coloration patterns have been widely studied, but they also benefit from a wide range of genetic resources too. The genus *Xiphophorus* of livebearing fishes presents a striking variation in pigmentation patterns, especially melanin-based ones, which are polymorphic within and between species (Basolo, 2006; Culumber, 2014). Accurate breeding experiments can be designed since these species reproduce well in aquaria which allow precise quantitative trait locus mapping experiments (Powell et al., 2021). Most importantly, the genus benefits from recently developed genomic resources (Cui et al., 2013; Powell et al., 2020; Schartl et al.,2013; Schumer et al., 2012, 2014, 2016) and even though gene knock out techniques are currently not possible on *Xiphophorus*, they are readily available in the closely related medaka (Wittbrodt et al., 2002). The study of melanin-based pigmentation patterns in *Xiphophorus* becomes even more interesting when it intricated history of hybridization is considered.

A growing body of literature has called attention to the importance of hybridization to the evolutionary process of speciation; genetic exchange is a pervasive feature of the evolutionary history of many if not most organisms, and is fundamental to generating phenotypic diversity and new evolutionary lineages (Abbott et al., 2013). A critical question in speciation is whether, when barriers to gene flow are established, gradual accumulation of independent changes will suffice to establish reproductive barriers and if not, what are the consequences of secondary contact. In the context of speciation, hybridization can act as a homogenizing force between parental species if hybrids do not suffer a fitness loss (Barton and Hewitt, 1985; Nosil et al., 2009). On the contrary, it may strengthen reproductive barriers promoting speciation via reinforcement to avoid a reduced hybrid fitness (Wu, 2001). Finally, hybrid speciation might also result, when the new populations of mixed ancestry remain reproductively isolated from their parentals (Abbott et al., 2013; Schumer et al., 2014).

A genetic model for the evolution of isolation mechanisms in hybrids is known as the Bateson-Dobzhansky-Muller model of hybrid incompatibility (BDMI). In 1922, JBS Haldane observed that: “When in the offspring of two different animal races one sex is absent, rare, or sterile, that sex is the heterozygous [heterogametic or XY] sex” (Haldane, 1922). Haldane and several subsequent surveys showed that this rule is obeyed in the most varied taxa and that it holds for all animals that have sex chromosomes. The significance of this rule therefore relies on its universality, which implies the existence of some common genetic mechanism during (postzygotic isolation) speciation among very different kinds of animals. In other words, there has to be a shared genetic process that explains this pattern. The problem in trying to understand the evolution of hybrid sterility is: how could natural selection allow the evolution of progeny that are infertile or inviable? (Orr, 1997). Hybrids can be pictured as a fitness valley between the

two adaptive peaks the parental species are. So, if these species have a common ancestor, how is it possible to go through the valley, from one peak to another? Selection would never allow it. Imagine one species has genotype *AA*, the other *aa*, and the sterile or inviable hybrid *Aa*. How is it possible to go from *A* to *a*? The *a* mutation necessarily arises in heterozygous state (*Aa*) and, therefore, it would be eliminated by selection (Orr, 1997).

Bateson was the first to propose an explanation of how sterility could arise in hybrids (Orr, 1996). He said that if sterility is the consequence of a combination of two factors, each factor has to be acquired by each species separately, on which its effect would be imperceptible until lineages hybridize (Bateson, 1909). These discussions appear in a "forgotten essay" (Orr, 1996) and where only resurfaced much later by Orr (Orr, 1996). Therefore, it can be said that by the time Dobzhansky published his studies on hybrid sterility, it was widely known that hybrids between different species suffered partial or complete sterility, yet, the causes of that sterility still remained obscure (Dobzhansky, 1934). Dobzhansky used *Drosophila* hybrids to discover that the problem that caused hybrids to be sterile was related to a failure of chromosome pairing. This failure could be due to structural dissimilarities between chromosomes or disturbances in cell physiology which prevented chromosomes from pairing. Dobzhansky hypothesized that the "disturbance of the gametogenesis may be due to the actions of complementary genetic factors contributed by both parents" (Dobzhansky, 1934). The key here is that Dobzhansky realized that there needed to be at least two independently evolved factors, contributed by each parental species, whose incompatibility is only exposed when they hybridize.

Independently, Muller also developed framework to explain why fitness of hybrids is reduced and focused on that there necessarily needs to be more than one single genetic change in

order to occur. Muller discusses that, since in the first hybrids mutations must exist in heterozygous conditions at first, no single mutation should be lethal. If mutations were lethal, hybrids would be incapable of reproducing at all and the mutation would never become established. It is therefore inevitably, that hybrid fitness reduction is a consequence of the interaction of at least two mutations (Jhonson, 2002). A model for hybrid incompatibilities via two genetic changes supposes that *A* and *a* are alleles at one locus, *B* and *b* are alleles at another locus, and that *a* and *b* together are incompatible (reduced fitness). If an *AABB* population gets split into two isolated populations, hybrid incompatibility could evolve if one population evolves to *aaBB* and the other one to *AAbb*. Hybrids between these populations would be *AaBb* and thus unfit. Maladaptive hybrid genotypes can therefore arise without the need to travel through an adaptive valley because they are not under natural selection (Jhonson, 2002). Furthermore, Muller realized that this model implies that hybrid incompatibilities must be initially asymmetric: asymmetric in the sense that it depends on the direction of the introgression and that only certain specific combination of alleles cause incompatibility. If alleles a and b are incompatible, *A* and *B* cannot be incompatible since *AABB* is the ancestral genotype, therefore the initial incompatibilities must be asymmetric. Work on *Drosophila* hybrids has supported this claim. Hybrid sterility factors are not the same when *D. simulans* is introgressed into *D. mauritiana* as those found in the reciprocal introgressions of *D. mauritiana* into *D. simulans* (Palopoli and Wu, 1994).

The logic behind BDMI is pretty simple: consider two populations evolving and accumulating different substitutions independently. If an allele from one population is introgressed into the genome of the other, we would expect it to vary in its functionality from either functioning reasonably well to not at all. Genomes are combination of alleles that have been tested together thus it is hard to predict that an introduced allele will perform better in a new genetic background

than in its own. Therefore, we expect a mixture of genes from two species to be less fit than those from a single species. Hybrid sterility or unviability or reduced fitness could, consequently, be a simple by-product of the divergence of two genomes (Orr, 1997). BDMI models are generally accepted nowadays and yet, we lack specific genetic mechanisms which explain how mutations with neutral effects within a particular species yield negative effects when expressed together in hybrids. Lynch and Force (2000) suggest that duplicate genes, and more generally any kind of genomic redundancies, "provide a powerful substrate for the origin of genomic incompatibilities" (Lynch and Force, 2000). Whenever there is a gene duplication, there are known cases of neofunctionalization of the duplicated genes. But the second copies of duplicated genes most commonly become silenced because they tend to accumulate negative mutations. Consider a pair of duplicate genes *A* and *B* in an ancestral species, fixed respectively for alleles *A* and *B* and that part of this population gets isolated. We can expect that each population, will randomly lose function at one of the two loci. There is a 50% probability that one population becomes fixed for *A* and *b* alleles and the other for *a* and *B* alleles. Hybridization between them would lead to *AaBb* progeny which would produce one fourth of nonfunctional ab gametes.

Interestingly, after branching off from the tetrapods, fishes underwent a whole genome duplication resulting in them often having two copies of a gene (Wittbrodt et al., 1998; Meyer and Van-de-Peer, 2005). Some of the melanin-based pigmentation patterns in *Xiphophorus* can result in malignant melanoma in both hybrid and non-hybrid situations (Schartl, 2008). Almost a century ago (Gordon, 1931), it was found that certain hybrids of platyfish (*Xiphophorus maculatus*) and swordtails (*Xiphophorus hellerii*) develop a highly malignant melanoma and since then *Xiphophorus* has been a dominant system in the study of melanoma. After decades of classical crossing experiments (Anders, 1991), the genetic theory of melanoma formation in these fish is

that in X. maculatus, there is an oncogene controlled by a tumor suppressor and that in *X. hellerii*, these genes are both lacking. When both the oncogene and the tumor suppressor are present the phenotype results in benign pigmentation. However, in the hybrids, because their genome is a combination of both parental species' genomes, it sometimes happens that the oncogene is expressed without the control of the tumor suppressor and therefore they develop malignant melanoma (Meierjohann and Schartl, 2006; Schartl, 2008). Thirty years ago, the oncogene was isolated: *Xiphophorus* melanoma receptor kinase (*xmrk*) (Wittbrodt et al., 1992), a modified *egfrb* duplicated receptor that is constitutively active and leads to a ligand-independent activation of the intracellular signal transduction network causing an unlimited cell proliferation. After many years of knowing the oncogene, we still haven't been able to identify the tumor modifier gene.

The *Xiphophorus birchmanni – Xiphophorus malinche* natural hybrid zone (Rosenthal et al., 2003) provides an excellent opportunity to identify the tumor modifier. *X. birchmanni* and *X. malinche* produce viable hybrids along several populations in the Sierra Madre (Hidalgo, Mexico) that appeared 35-56 generations ago (Schumer et al., 2014) likely due to human-mediated disruption of pheromonal communication (Fisher et al., 2006). Some individuals of *X. birchmanni* and *X. birchmanni - X. malinche* hybrids present a macromelanophore pigment pattern called Spotted Caudal (Sc), which varies in its expression from a few black spots to extreme melanosis and eventually malignant melanoma (Fig. 1). Recently formed hybrid zones constitute "natural laboratories for evolutionary studies" (Hewitt, 1998), which can be used to pinpoint genes of interest, disentangle multivariate effects on fitness, and measure evolutionary change in real time (Lexer et al., 2010). Natural variation provides an underutilized resource for insight into the molecular mechanisms underlying disease. While most research on this topic relies on

carefully controlled domesticated models, this study benefits from an ongoing hybridization process under natural conditions of a cancer research model organism, that provides an excellent opportunity to study the genetic and evolutionary causes and consequences of melanoma.

CHAPTER II

NATURAL HYBRIDIZATION AND THE SEARCH FOR INCOMPATIBLE ALLELES THAT CAUSE MELANOMA IN *XIPHOPHORUS*

Introduction

Ever since Mendel postulated that "internal factors" are passed from one generation to the next, the relationship between genetic polymorphisms and phenotype has been subject of interest. The basic approach has been to screen many individuals of different genotype, obtained either by mutagenesis or natural variation in the wild, to connect their genetic variation to an interested phenotype (Korte and Farlow, 2013). This method is nowadays known as population genomics, but the field of study is as old as genetics itself. As expected, discrete traits were the first to be understood. Such is the case with the *ABO* blood groups of humans (Bernstein, 1925), Batesian mimicry in Papilio butterflies (Punnet, 1915) and heterostyly in Primula (Bateson and Gregory, 1905). Subsequent quantitative studies (Muller, 1928; Wright, 1922) provided evidence of the immense amount of "hidden" genetic variability and confirmed Darwin's view that there is plenty of genetic variability for selection to act (Darwin, 1868). However, important questions arouse: what is the average amount of genetic variation in natural populations and to which extent did selection control that variation (Lewontin, 1974).

John Hubby, Richard Lewontin, and Harry Harris, were the first to estimate genome-wide levels of variation in populations of *Drosophila pseudoobscura* and humans respectively (Harris, 1966; Hubby and Lewontin, 1966). They used gel electrophoresis to screen for protein variations according to their migration rates on the gel. Even though they found that a large number of loci are polymorphic, the technique they used limited their ability to detect any amino acid change that did not affect the protein's charge or any silent DNA changes (Lewontin, 1985). The first

studies of DNA sequence variation relied on restriction enzymes to detect variation in the presence of their cutting sites (Avise, 1983; Nei and Li, 1979). Impressive results, which still stand today, were achieved with this method. For example, transposable elements contributed low-frequency polymorphisms in *Drosophila* but almost nothing to human variability while in both *Drosophila* and humans single nucleotide polymorphisms contribute the most to variability (Kan and Dozy, 1978; Langley and Aquadro, 1987). The survey of DNA variation improved considerably with the invention of PCR amplification. PCR and sequencing machines provided the ability to sequence multiple copies of a specific region of the genome pretty easily. The main result from this kind of work was that most of the variability is neutral: it either does not affect protein structure or was found in non-coding regions (Kimura, 1983; Kreitman, 1983). Since then, the methods for surveying DNA variation have only advanced, becoming more and more affordable and therefore causing a flourishing of population genomic studies.

Population genomic studies can be defined as the study of numerous loci or genome regions with the objective of understanding evolutionary processes. Black et al., 2001 define it more precisely as "the use of genome-wide sampling to identify and to separate locus-specific effects (such as selection, mutation, assortative mating and recombination) from genome-wide effects (such as drift or bottlenecks, gene flow and inbreeding) to improve our understanding of microevolution" (Luikart et al., 2003). Population genomics assume that neutral loci across the genome should be equally affected by the evolutionary history and demography of a population whereas loci under selection will reveal unique, and therefore identifiable, patterns of variation (Luikart et al., 2003). Quantitative Trait Locus (QTL) mapping has proven to be a powerful method to identify genome regions associated to a particular trait either in F2 populations or in-

bred lines. This has a limited mapping resolution because of the reduced amount of recombination of those populations and because the allele frequencies and combinations present in lab populations will differ from natural populations underrepresenting the functional diversity present (Korte and Farlow, 2013). Genome wide association studies (GWAS), overcome these limitations by testing thousands to millions of genetic variants across the genomes of many individuals either to study the genetic architecture (basically, the number of loci that contribute and their respective contribution to the phenotype) of a trait or to identify causative/predictive factors for a given trait (Korte and Farlow, 2013; Tam et al., 2019). The first GWAS published was in 2005 for the age-related macular degeneration (Klein et al., 2005) and since then, according to the NHGRI-EBI Catalog of genome-wide association studies (MacArthur et al. 2017), as of September 15th 2021, there has been 5329 publications and more than 5700 according to Uffelmann et al., 2021. GWAS have successfully identified risk loci in a great number of disease and especially in cancer (Liang et al.,2020; Luikart et al., 2003; Uffelmann et al., 2021).

Hybridization between closely related species, from either lab reared hybrids or naturally occurring populations, is also a useful tool for understanding the genetic basis of the phenotypic traits that distinguish species. *Xiphophorus* has an extensive history of hybridization which has recently been exploited to understand the genetic basis of sexually selected traits (Powell et al., 2021; Schartl et al., 2021) and melanoma (Lu et al., 2020; Powell et al., 2020). The *X. birchmanni – X. malinche* natural hybrids are particularly interesting to study because different populations differ in their ancestry patterns. In certain hybrid populations, separate genetic clusters have formed and maintained in near-perfect isolation for over 30 generations through strong assortative mating whereas in other populations ancestry is randomly distributed (Culumber and Rosenthal, 2013; Schumer et al., 2017). *Xiphophorus* birchmanni is polymorphic for a coloration

pattern called Spotted Caudal (Sc) "which typically consists of one or more irregular, elongated patches of black pigmentation commencing close to the base of the middle or lower caudal fin rays and extending posteriorly for roughly one-third of the fin length" (Sedarti et al., 1995). This macromelanophore formation is particularly interesting because it can turn into malignant melanoma in the *X. birchmanni – X. malinche* natural hybrids (Powell et al., 2020; Culumber, 2014). It is proposed to be a typical case of a hybrid incompatibility in which a negative epistatic interaction between genes of both species cause the malignant melanoma (see Chapter 1). The *X. birchmanni – X. malinche* hybrids provide an excellent opportunity to test this hypothesis. If melanoma is indeed caused by a negative epistatic interaction between *X. birchmanni* and *X. malinche* genes, I expect that populations with different ancestry patterns differ in their Sc frequency and malignancy since in structured populations the probability of having a negative epistatic interaction is lower. Therefore, to understand the genetic basis of this trigger between benign coloration (Sc) to malignant pigmentation, I performed two population genomics studies: a GWAS in a pure parental population of *X. birchmanni* to identify the genetic driver of Sc, and a comparative study of the presence and malignancy of Sc between structured and unstructured populations of *X. birchmanni – X. malinche* hybrids to determine whether an epistatic interaction might be responsible for the melanoma.

Methods

Collection and Phenotyping

The spotted caudal trait is polymorphic within *X. birchmanni*. Thus, it is possible to search for variants associated with the phenotype. In order to do so, I collected 232 unspotted controls and 159 spotted cases (males and juvenile males) from the *X. birchmanni* Coacuilco population using baited minnow traps. Fish were anesthetized in MS-222 (Texas A&M IACUC protocol #2016-0190) following a standardized morphometric protocol which consisted of using a background grid and spreading the caudal and dorsal fin. A Nikon d90 DSLR digital camera equipped with a macro lens was used. Fish were scored for presence/absence of the spotted caudal phenotype. A small caudal fin clip was taken and fish were kept in fresh water until full recovery from the anesthesia before returning them back to the river. Fin clips were preserved in ethanol until DNA extraction. Area of the spot, body and dorsal fin standard length and depth were measured from photographs using ImageJ (Schneider et al, 2012).

DNA extraction

DNA extractions were made following the DNeasy Quiagen Kit instructions. Briefly, 20ul of Poteinase K were added to each fin clip and incubated overnight at 56 oC. Samples were vortexed and then 200ul of Buffer AL and ethanol were added to the sample before vortexing again. The mixture was then transferred to a DNeasy Mini spin column placed in a 2ml collection tube and centrifugated at 8000 rpm for 1 minute. Flow through was discarded and 500ul of Buffer AW1 were added before centrifuging at 8000 rpm for 1 minute. Flow through was discarded and 500ul of Buffer AW2 were added. Samples were then centrifuged 3 minutes at 14000

rpm, flow through was discarded and then samples were centrifuged 1 minute at 14000 rpm. The DNeasy Mini spin column was then transferred to a new 1.5ml Eppendorf and 50ul of Buffer AE were added. After 1 minute incubation at room temperature, samples were centrifuged for 1 minute at 8000 rpm to elute the DNA. All centrifugations were performed at room temperature. Extracted DNA was quantified using the Nanodrop 8000 (Thermo Fisher Scientific, Waltham, MA).

Library Preparation

In order to perform low coverage whole genome sequencing (1X average coverage) ten to thirty nanograms of DNA were used to construct Tn5 libraries. This basically consists of combining each sample with a unique combination of i5 and i7 index in order to identify it later. For this, first the Tn5 first has to be precharged by combining two Tn5 adapters by mixing 15 uL of Tn5 (100 ng/uL), 122 uL Reassociation Buffer/Gycerol 1/1 and 3 uL of each adaptor and incubate in a thermal cycler at 37 ℃ for 30 minutes. (Reassociation Buffer: 10 mM Tris pH 8.0, 50 mM NaCl, 1 mM EDTA; each adaptor was previously prepared by mixing 10 uL of forward and reverse oligonucleotides with 80 uL of reassociation buffer and annealed for 10 min at 95 ℃, 1 min at 90 ℃, Reduce temp by 1°C/cycle 60 times in thermocycler and then hold at 4 ℃). Afterwards, for 96 samples a master mix of 120 uL of the pre-charged Tn5, 240 uL of TAPS buffer (500 mM) and 600 uL of water was made. (5x TAPS buffer: 50 mM TAPS-NaOH, 25 mM MgCl2, 50% v/v DMF (pH 8.5)). 8 uL of the master mix were added to each well of a plate, combined with 2 uL of DNA (3-10 ng/uL) and incubated at 55 ℃ for 7 minutes. This shears the DNA. The Tn5 remains bound to the DNA unless stopped so 2.5 uL 0.2% SDS were added to each reaction and then incubated at 55 ℃ for 7 minutes to stop it. 3 ul of this sheared DNA were

combined with 12 uL of a master mix containing the i5 (Master mix for 96 samples: 120 uL 10 uM i5, 900 uL OneTaq HS Quick-Load 2X, 420 uL water) and 1 uL of 10 uM i7. Samples were run on a thermal cycler at 68 ℃ for 3 min, 95 ℃ for 30 seconds, 95 ℃ for 10 seconds, 55 ℃ for 30 seconds and 68 ℃ for 30 seconds. The last three steps were repeated 12 times and then samples were held at 4℃. Reactions were cleaned using magnetic beads. 10 uL of each library was pooled with other libraries sharing an i5 index and then 0.6 volumes of beads were added, mixed thoroughly and incubated at room temperature for 10 min. Tubes were transferred to a magnet for 5 minutes and the supernatant was carefully removed to a clean tube. 0.2 volumes of beads were added and the previous steps were repeated but the supernatant discarded this time. Then 200 uL of freshly prepared 70% ethanol were added, and tubes were let stand in the magnet for 30 seconds and then remove the ethanol. The ethanol wash was repeated again and the beads were allowed to air dry for 10-15 minutes. The tubes were removed from the magnet and 13 uL of Tris (10 mM) were added, mixed thoroughly, let stand for 5 min, and then the tubes were transferred to the magnet for 5 minutes. Finally, 11 uL of the eluate were carefully removed. Libraries were quantified using the Qubit fluorimeter (Thermo Scientific, Wilmington, DE). Library size distribution and quality were visualized using an Agilent Tapestation (Agilent, Santa Clara, CA). Libraries were sequenced on the Illumina NextSeq 4000 or the Illumina HiSeq 4000 across four lanes to collect paired-end 75 basepair reads and 150 basepair reads respectively.

Bioinformatic Analysis

The following bioinformatics analyses were performed in collaboration with Dr. Molly Schumer. To estimate allele frequency differences between Sc and non-Sc groups we used an

older version of samtools-legacy program (https://github.com/lh3/samtools-legacy/blob/master/samtools.1) because it allows to test for differences between the two groups assuming a one-degree χ^2 distribution (Heng Li, personal communication to Molly Schumer, (Li, 2011)). To evaluate the accuracy of this method, we performed simulations using macs (Chen et al., 2009) to generate diploid sequences of 1 Mb for 391 individuals, assuming $\theta = 0.0012$ per site and previously inferred recombination rate of *X. birchmanni* (Schumer et al., 2018). We used this to generate nucleotide sequences with the observed base composition of *X. birchmanni* using seq-gen (Rambaut and Grassly, 1997) and then paired haploid sequences were converted to generate diploid individual fastq files for each individual using the program wgsim (https://github.com/lh3/wgsim). With a sequencing error rate of 0.01 and 150 bp paired-end reads, we set the number of reads generated to 3330 to match the average 1X coverage of our data. Then, we assigned randomly 232 unspotted controls and 159 spotted cases, mapped the reads with bwa-mem and ran samtools-legacy mpileup and bcftools-legacy to quantify allele frequency differences in the simulated spotted and unspotted groups. To evaluate the accuracy of the pipeline, we compared the simulated allele frequency difference to the true allele frequency difference at each SNP, repeating the procedure 100 times. The average difference between true and inferred allele frequencies was 1.2%. This indicates that samtools-legacy has excellent accuracy in quantifying allele frequency differences between groups given coverage and sample sizes matching our data.

For the real data, individual we mapped reads with bwa-mem, sorted bam files with samtools and removed reads with mapping quality scores less than 30. Reads were mapped to the newly assembled genome developed by Powell et al., 2020. Using these bam files we ran samtools-legacy mpileup and bcftools-legacy to estimate allele frequency differences between

spotted and unspotted individuals. For the allele frequency analysis, we only used high quality SNPs (1,254,071 SNPs genome-wide) ascertained by high coverage whole genome sequence data from 26 individuals from the *X. birchmanni* Coacuilco population (Schumer et al., 2018). The appropriate genome-wide significance threshold for allele frequency differences between spotted and unspotted individuals, was determined by shuffling spotting phenotypes and repeating the genome-wide scan for allele frequency differences between cases and controls 500 times. Based on these permutations, we set our genome-wide significance threshold at 1-7 since fewer than 5% of permutations had associations at or lower than this threshold.

Population structure and spotted caudal

I sampled 6 hybrid populations along a less than 3km stretch of the Calnali River. Three of those populations (CalnaliMid, Aguazarca, Piloncillo) are located upstream the city of Calnali. The other three are either in the stretch of the river that runs through the city (Plaza, Peatonal) or slightly downstream (CalnaliLow). Phenotyping, DNA extraction, and library prep for low sequencing methods were the same as the ones described above for the GWAS. The following bioinformatics analyses were performed in collaboration with Dr. Molly Schumer and are well described in Powell et al. (Powell et al., 2020). We used ancestry informative sites across the genome to calculate an individual hybrid index that is then used to determine population structure. We used a recently developed hidden Markov model-based approach, AncestryHMM (Corbett-Detig and Nielsen, 2017) to call ancestry. We used this approach over multiplexed-sho*tg*un genotyping (Schumer et al., 2016) because of its reduced computational times and increased efficiency for large datasets. Briefly, reads are first mapped to the parental genomes using bwa mem (Li and Durbin, 2009) and bam files are sorted with samtools (Li, 2011). Then, reads that did not

map to either parental genome are identified and filtered with ngsutils (Breese and Liu, 2013). With the remaining reads, bcftools mpileup and custom scripts are used to count ancestry informative sites for each allele. Only individuals with more than 300,000 reads were used (Schumer et al., 2016) as input for the AncestryHMM pipeline and only sites with posterior probability of 90% or higher we considered. The AncestryHMM pipeline is available on github (https://github.com/Schumerlab/ancestryinfer).

Results

GWAS

The GWAS identified a strong association between two loci in chromosome 21 and the spotted caudal phenotype (estimated false discovery rate of 5%). The two hits identified by the GWAS are located 5mb apart from each other and are: *xmrk* and the melanosome transporter genemyosin VIIA and Rab interacting protein (*myrip*) (Figure 1).

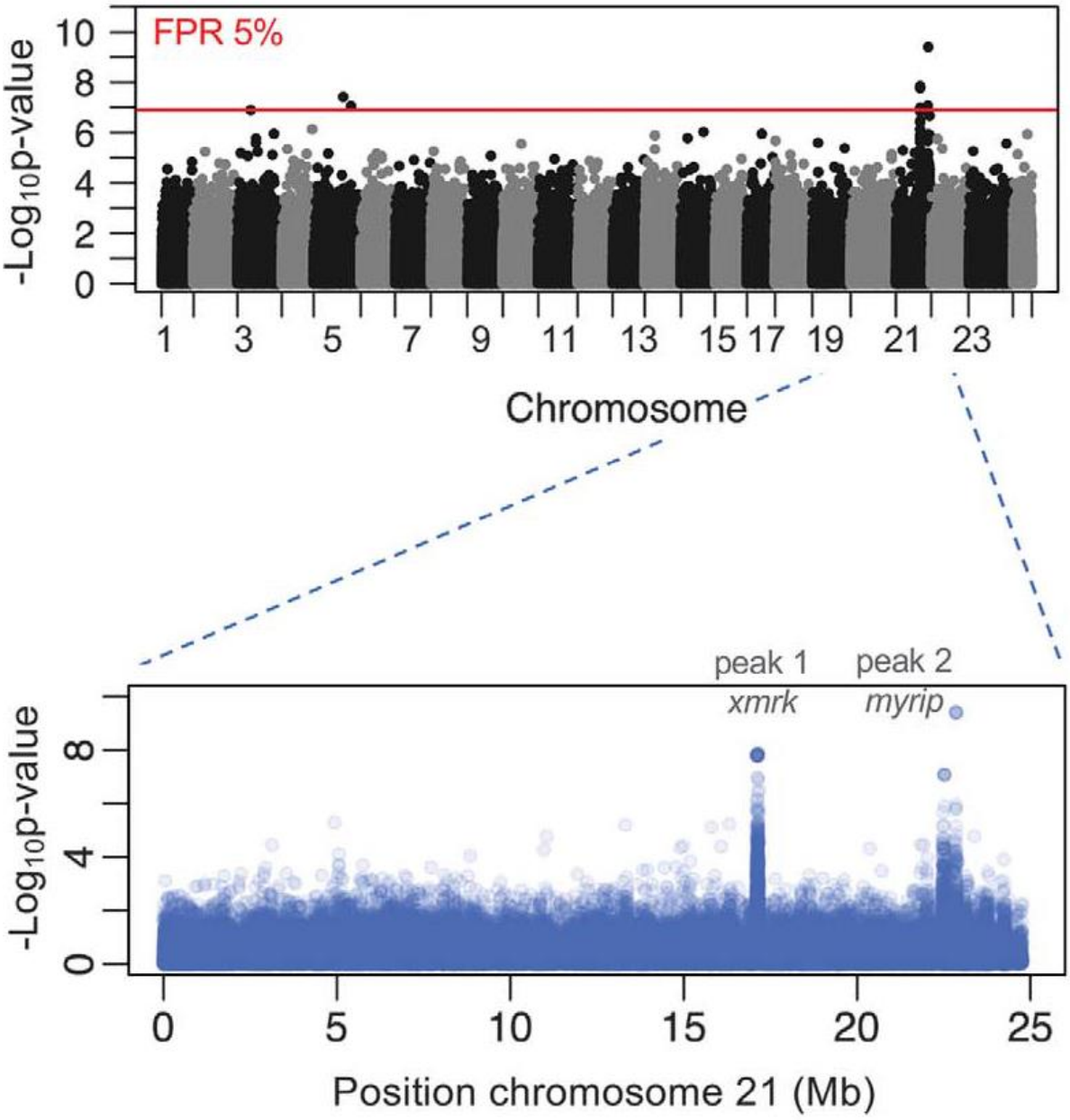


Figure 1. (Top) Results of genome-wide association scan for allele frequency differences between spotted cases and unspotted controls for every chromosome. Red line indicates the genome-wide significance threshold. (Bottom) Zoom into chromosome 21, where two distinct regions are strongly associated with spotting. Published in Powell et al., 2020.

Population structure can generate allele frequency differences between spotted and unspotted individuals by chance (Price et al., 2006). To evaluate that, and since low coverage data does not permit calling diploid genotypes with confidence, we generated pseudo-haploid calls for every individual at each known SNPs by randomly picking a sequencing read and assigning the allele supported by that read to represent a haploid genotype of that individual. With these pseudo-haploid calls we performed a PCA of genome-wide SNPs in all individuals using plink (Purcell et al., 2007) and checked whether any of the first 20 PCs had a correlation with the spotting phenotype. Only PC1 and PC2 had correlations with the spotting phenotype ($p<0.001$) and together they explain 1.5% of the observed genome-wide SNP variation. However, this could be due to the fact that we used pseudo-haploid calls which can weaken the association between the spotted trait and the focal SNP. However, if we analyze p-values with (range of p-values for peak 1 from 100 replicates: 1-8 - 5-6; peak 2: 3-10 - 2-13) and without (range of p-values for peak 1 from 100 replicates: 2-7 - 3-6; peak 2: 2-11 - 1-8) adding PC1 and PC2 as covariates, they are similar. This suggests that the population structure observed is subtle and not affecting the GWAS results.

Xmrk is known to be a neofunctionalization of a duplicated *egfrb*, with identified mutations that cause constitutive activation (Schartl, 2008). Nevertheless, both genes are closely related and their sequences are very similar. To determine whether the GWAS peak contained *egfrb* or *xmrk* we built a likelihood phylogeny of *egfrb* and *xmrk* sequences from X. maculatus using RAxML (Stamatakis, 2006). The gene in the GWAS peak is most closely related to *xmrk*.

Population structure and spotted caudal

Populations upstream the city of Calnali (CalnaliMid, Aguazarca, Piloncillo) are structured according to their ancestry into two separate genetic clusters: birchmanni-like hybrids and malinche-like hybrids, whereas populations located within the city of Calnali or downstream of it (Plaza, Peatonal, CalnaliLow) did not show this pattern (Figure 2 b). Moreover, ancestry unstructured populations showed increased Sc frequencies (ANOVA: p value 0.00188, Figure 2 a) and higher Sc index (calculated as the Sc area/standard length, ANOVA: p value< 0.001, Figure 3)

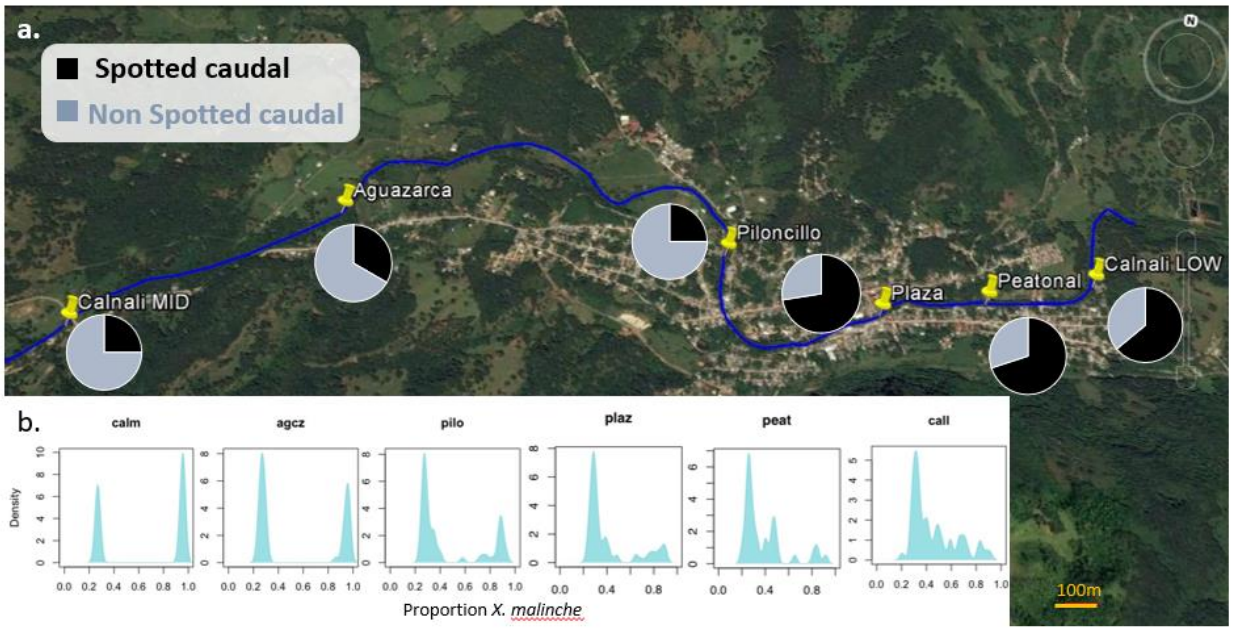


Figure 2. a) Sc frequencies along the Calnali River b) HMM data for each population's ancestry distribution.

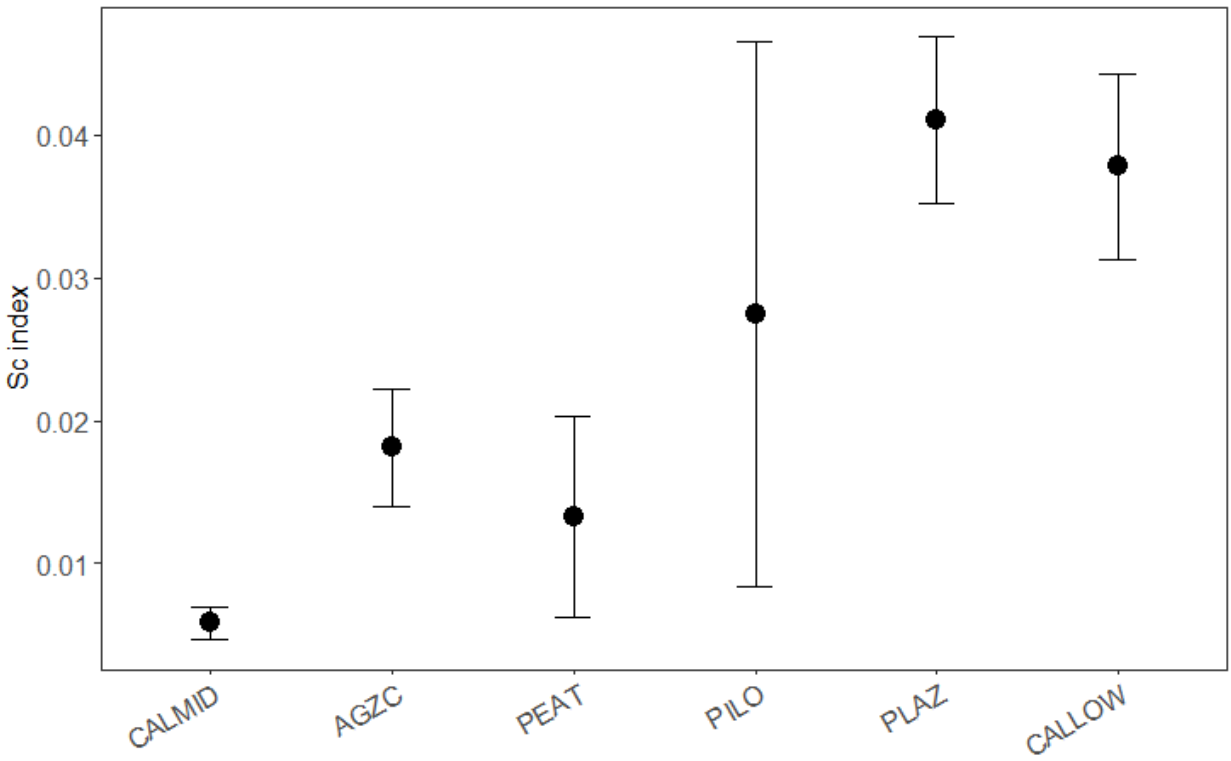


Figure 3. Sc index (Sc area/standard length) for each population (CALMID = CalnaliMid, AGZC = Aguazarca, PEAT= Peatonal, PILO= Piloncillo, PLAZ= Plaza, CALLOW= CalnaliLow). Dots show the mean and error bars indicate one standard error of the mean.

Discussion

The GWAS identified associations between spotted caudal and both *xmrk* and *myrip* (Figure 1) and therefore either one of them or both could potentially be the driver of the hybrid incompatibility. However, Powell et al., 2020 determined that *myrip* is not expressed in the adult caudal tissue nor in the melanoma (Figure 4 a). On the contrary, *xmrk* is more expressed in spots and melanoma compared to unspotted (Figure 4 a) and has been previously proven to be responsible for melanoma development in other *Xiphophorus* species (Meierjohann and Schartl, 2006; Schartl, 2008). Individuals of X. maculatus- X. helleri hybrids that either have a deletion of the *xmrk* gene (Wittbrodt et al., 1989) or whose *xmrk* reading frame is disrupted by a transposable element (Schartl et al., 1999) are loss-of-function mutants for melanoma development. Moreover, we know that *xmrk* is not only necessary but sufficient for tumor development because transgenic expression of X. maculatus *xmrk* in Oryzias latipes (the closely related medaka fish) leads to tumor development at a high frequency (Schartl et al., 2010; Winkler et al., 1994). Unlike *egfrb*, which requires a ligand and signals only after the growth factor has bound to the extracellular domain, *xmrk* is constitutively active (Wittbrodt et al.,1992). This happens because *xmrk* has two exchanged amino acids in the extracellular domain that destroy the formation of intramolecular disulfide bonds and instead intermolecular cysteine bridges between two receptor monomers form which mimics the effect of ligand binding (Gomez et al., 2001). *Xmrk* causes a constant activation of the ras/raf/MAPkinase mitogenic cascade as well as anti-apoptotic pathways (Schartl, 2008). *Xmrk*'s behavior therefore, is consistent with a one of the main characteristics of cancer cells: unlimited cell proliferation, independent of growth regulation factors. Powell et al. 2020

found that *X. birchmanni* also has the two mutations (G364R and C582S) responsible for the constitutively activation of *xmrk* in *X. maculatus* (Figure 4 b (Powell et al., 2020))

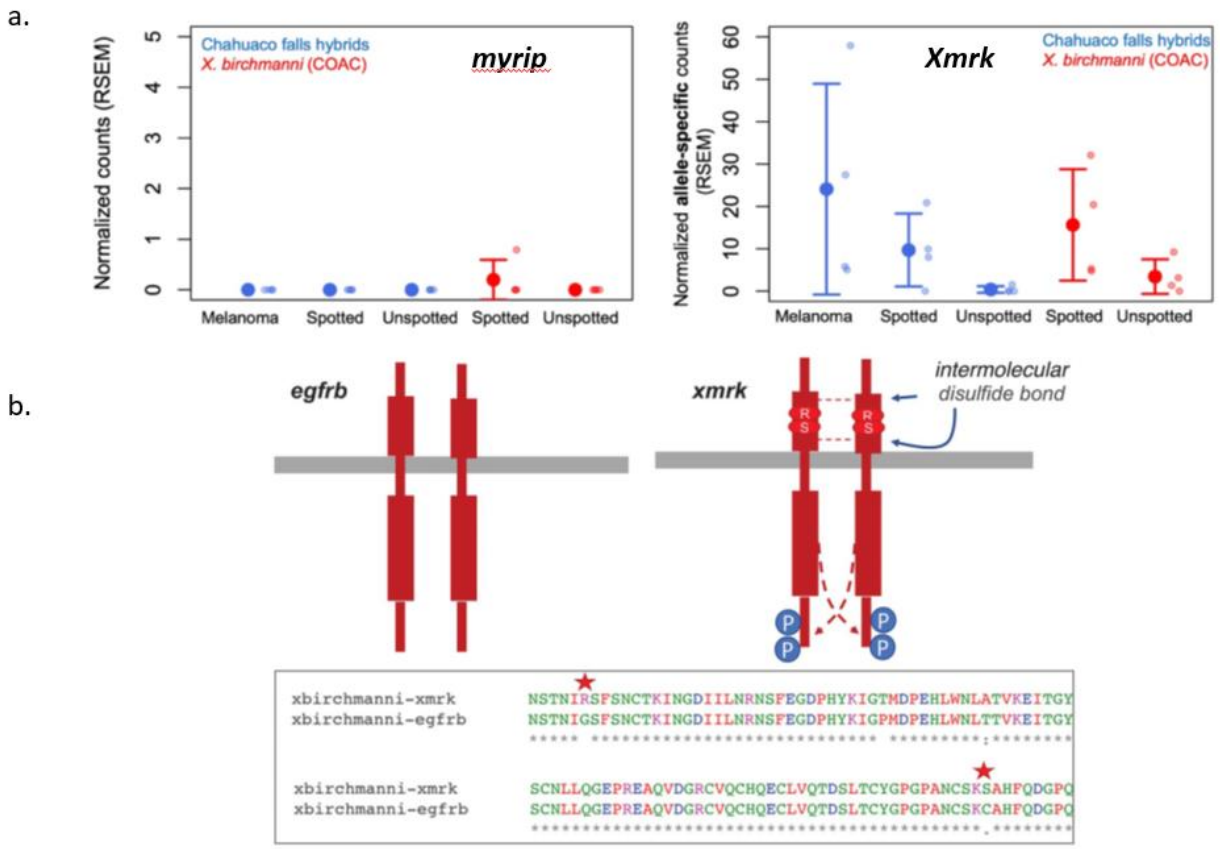


Figure 4. a) Expression of GWAS hits *myrip* and *xmrk* in caudal tissue of *X. birchmanni* and natural hybrids. b) The schematic compares the ancestral form of the protein (*egfrb*) to the predicted structure of *xmrk* in *X. birchmanni*. Proteins are shown in red, and the cell membrane is shown in gray. Blue circles show phosphorylation. (Inset) A partial clustal alignment of *X. birchmanni egfrb* and *xmrk* with these substitutions highlighted. Colors indicate properties of the amino acid, and asterisks indicate locations where the amino acid sequences are identical. Published in Powell et al. (2020).

myrip is a protein that interacts with the molecular motor Myosin VIIa and is involved in the transport of melanosomes in the retinal pigment epithelium. Furthermore, *myrip* is related to skin melanosome transport via regulating the activity of Myosin Va, another molecular motor (Ramalho et al., 2009). The lack of expression of myrip in adult tissue suggests that *myrip* may not be directly involved in the development of melanoma. However, there has been plenty of discussions (reviewed in (Culumber, 2014)) regarding the existence of a macromelanophore determining locus (*Mdl*) (Weis and Schartl, 1998; Schartl et al., 2013). *Mdl* is supposed to determine when and where and how many macromelanophore spots appear in the body or fin compartments. In fact, it is pattern determining locus in the sense of developmental biology since pigmentation patterns occur in the first several weeks of life. From phylogenetics distributions of macromelanophore patterns and *xmrk* we know that mdl and *xmrk* are tightly linked but separated loci (Weis and Schartl, 1998). A feasible study to test if *myrip* can play a role in the pigmentation patterning would be to generate transgenic medaka lines with *Xiphophorus* genes of *xmrk* and *myrip* and see if there are changes in pigmentation patterns compared to the already existing *xmrk* medaka lines (Schartl et al., 2010).

Hybrid ancestry-structured populations showed Sc frequencies and levels of expression similar to our reference parental population whereas ancestry-unstructured populations showed increased frequencies and levels of expression (Figure 2). The association between the collapse in population structure and increased frequency and magnitude of expression of Sc is an indicator that hybrid melanoma is caused by a negative interaction between *X. birchmanni* and *X. malinche* genes. The rational is simple: in structured populations we find individuals whose genome

is either entirely more similar to *X. birchmanni* or *X. malinche* (Schumer et al., 2017) and therefore it is less likely that the genes from each species responsible for the incompatibility are present in a given individual. In ancestry-unstructured populations however, individuals' genomes are an admixed composition of both parental species' genomes and therefore it is more likely that the negatively interacting genes from each species responsible for the incompatibility are present. This was later confirmed by an admixture mapping study on a high incidence Sc hybrid population of *X. birchmanni – X. malinche* hybrids in which the interaction between *X. birchmanni xmrk* with *X. malinche* allele of a gene called *cd97* (Powell et al., 2020) (To be discussed further in the introduction of Chapter 4).

A question that arises looking as this data is why is there a marked collapse in the ancestry population structure of the hybrids? Schumer et al. (2017) (Schumer et al., 2017) propose that the ancestry population structure is maintained by strong assortative mating and therefore the question becomes: why is there such a marked collapse in assortative mating? Even though it is beyond the scope of this study, I hypothesize that it is related to environmental disturbance and its relationship to mate choice in *Xiphophorus*. *Xiphophorus* females show mating preferences for visual cues but even more robust preferences for chemical cues and given the often co-existence of several closely related species of *Xiphophorus,* it is suggested to play an important role in maintaining reproductive isolation (Fisher et al., 2006). Indeed, it has been demonstrated that increased contents of humic acid (organic material contamination) block their olfactory bulbs making them unable to discriminate between each other, promoting hybridization (Fisher et al., 2006). Although not measured, water quality drops as the river runs through the town (personal observation: more trash in the river, turbidity and bad smell) and this coincides with the collapse in ancestry population structure/assortative mating: ancestry structure is maintained in hybrid

populations found upstream the city of Calnali while those within the city or downstream of it lack this ancestry structure (there is another population further downstream the city called CHAF that shows this too (Powell et al., 2020)). Further experiments should properly categorize this water quality droppage.

CHAPTER III

PHENOTYPIC CORRELATES OF A MELANOTIC PHENOTYPE IN HYBRIDIZING SWORDTAILS (TELEOSTEI: *XIPHOPHORUS)*

Introduction

Cancer is arguably one of the worst consequences of multicellularity. It is considered as an evolutionary process that occurs at two levels: within individuals there is somatic selection of cancer genes and cell lines, whereas at the population level natural selection predisposes or prevents cancer (Frank and Nowak, 2004; Greaves, 2000). There are several hypotheses that address how cancer might have originated (Blanpain, 2013; Graham, 1992) but they do not explain how it can persist. Although counterintuitive, some types of cancer cell lineages are under positive or purifying selection. It has been suggested that this is driven by evolutionary conflicts (Haig, 1993; Kleene, 2005) which "led to rapid evolution in genetic, developmental and physiological systems of control over cellular resources, which creates evolutionary disequilibrium and organism-level maladaptation, manifested as increased cancer risk" (Crespi and Summers, 2006). For example, *SPANX* is a family of X-linked genes whose expression is exclusively limited to normal testis and melanoma tumor cells. These genes are inferred to be under strong positive selection since they contribute to spermatozoa fitness, evolving to achieve high translation rates (Kouprina et al., 2004). These genetic pathways, positively evolved in a sexual conflict context, are apparently co-opted by cancer cells to thrive, differentiate and avoid apoptosis (Kleene, 2005).

Another example is the case of the CAG repeat region of the human androgen receptor gene. Short repeats are associated with increased risk of prostate cancer, and with more aggressive forms of the disease while at the same time they are also associated with increased fertility at the phenotypic level (via increased transactivation of the androgen receptor at the molecular level) (Summers and Crespi, 2008).

These signatures of positive selection on oncogenes are examples of the widely known tug of war between sexual and natural selection. Oncogenes could even provide a net benefit, if their expression somehow enhances a phenotype under sexual selection. By definition, sexual selection is going to favor traits that increase mate acquisition and fertilization, regardless of their effect on survival and fecundity (Safran et al., 2013). In fact, when sexual and natural selection act on the same trait they can act in opposite directions. An elegant example of this is the wild Soay sheep. Large horns, determined by the *Ho+* allele (of the relaxin-like receptor 2 (*RXFP2*)), confer an advantage in strong intra-sexual competition and therefore is associated with higher reproductive success. Smaller horns however, determined by *HoP*, confer increased survival, resulting in a trade-off effect of overdominance for fitness at *RXFP2*, which maintains the polymorphism (Johnston et al., 2013). (As stressed in the introduction of this thesis, this is another example that highlights the importance of identifying the genetic architecture of trait variation).

Coloration patterns are sexually selected signals in a wide variety of taxa (Andersson, 1994). Usually, the most desirable mate has the most pronounced visual trait (reviewed in (Ryan and Keddy-Hector, 1992)) simply because they elicit more stimulation in choosers (Ryan, 1990a). If an oncogene can enhance pigmentation patterns, which confer a sexual advantage, regardless of how detrimental for survival a cancer gene can be, sexual selection could contribute

to its maintenance. This chapter of my dissertation is based on that. As mentioned before, the oncogene *xmrk* is associated with the expression of the polymorphic Spotted caudal (Sc) pigment trait in *X. birchmanni* and melanoma formation in the *X. birchmanni – X. malinche* hybrids. Analyses of the ratio of synonymous to non-synonymous substitutions of *xmrk* show that, rather than evolving as a pseudo-gene, *xmrk* has a positively selected function still unknown (Schartl, 2008; Volff and Schartl, 2003). A recent study (Powell et al., 2020) observed shifts in the frequency of the spotted caudal trait between juvenile and adult males in a *X. birchmanni – X. malinche* hybrid population with high incidence of melanoma. Spotted juveniles suffered an increased mortality compared to unspotted juveniles. These differences were observed only in hybrid populations with high incidence of melanoma, but not in hybrid populations with lower frequencies or in a pure *X. birchmanni* population (Figure 5 a). Melanoma, when advanced, can invade surrounding tissues and histology showed degradation of muscular tissue that connects to the caudal fin (Figure 5 b) that causes them to escape more slowly when startled (Figure 5 c). If we consider that fish with expanded Sc are more conspicuous to avian and piscine predators, a reduced ability to escape from predators could explain this observed differential mortality of spotted individuals (Powell et al., 2020). This raises the question of how is this trait still segregating in some hybrid populations and therefore this chapter explores mating preference for Sc in parental and hybrid populations to determine whether large spots, or any other traits associated with large spots, confer a mating advantage.

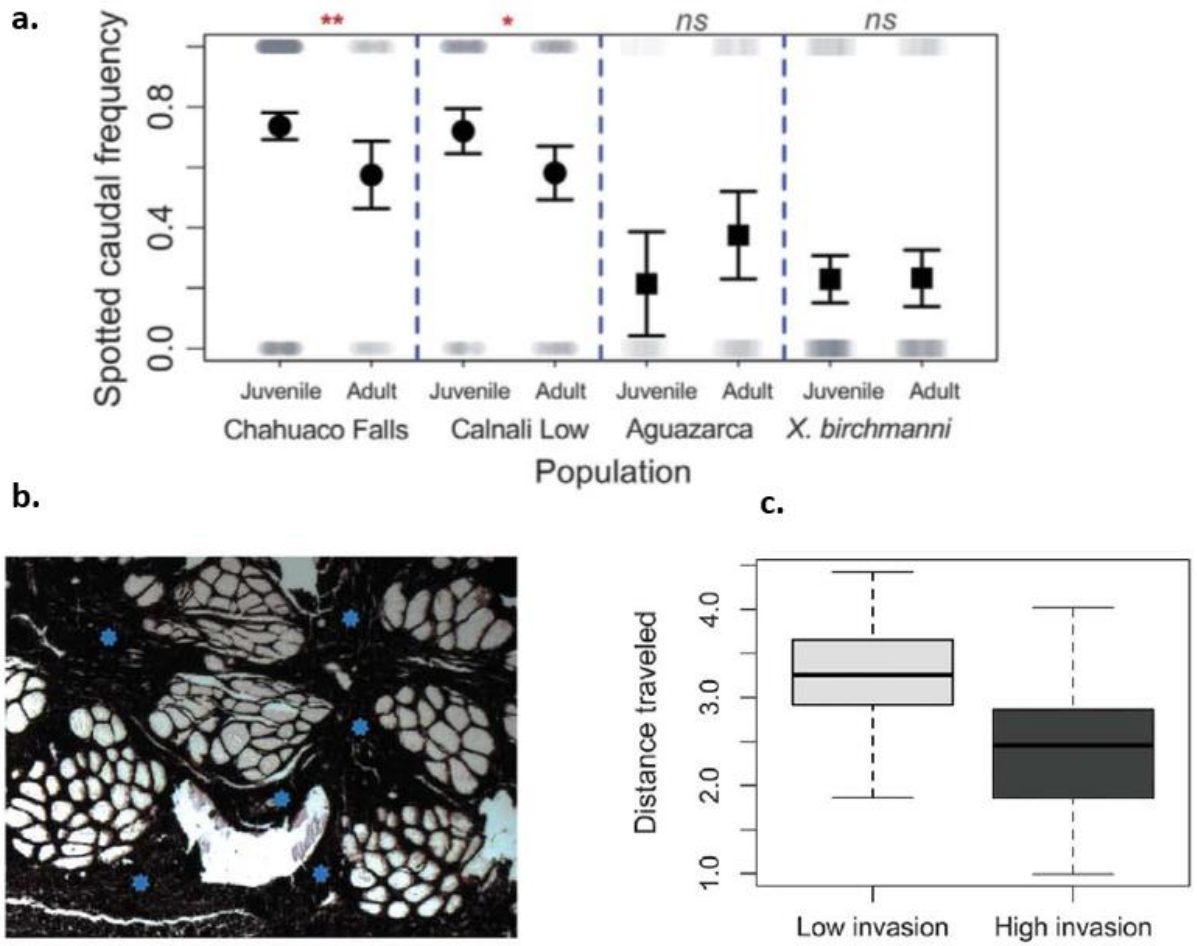


Figure 5. a) Frequency of spotting in juvenile and adult males across hybrid populations with high (circles, Calnali low and Chahuaco falls) or low (squares, Aguazarca and *X. birchmanni*) melanoma incidence. Asterisks indicate significant differences by age class (*P < 0.05, **P < 0.01; ns indicates nonsignificant differences in a two-sample z test). Gray points indicate the raw data, black points indicate the mean, and error bars indicate one standard error of the mean. b) A cross section of the caudal peduncle from a Chahuaco falls hybrid (10× magnification). Melanoma cells invading the body and muscle bundles are visually evident (indicated with blue stars). c) Visualization of the difference in fast-start response between individuals with low and high melanoma invasion (upper and lower 25% quantiles shown here). Published in Powell et al. 2020

Mate choice has been a key concept in sexual selection since Darwin invoked the "taste for the beautiful" (Darwin, 1859). Although mate choice is distinct from sexual selection, in fact, mate choice decisions can turn out to be maladaptive, mate choice is key for sexual selection. After all, "who a chooser mates with and who she pairs with will affect how long she lives and how many healthy children and grandchildren she has. Mate choice determines which sperm fuse with which eggs, and therefore ultimately shapes how lineages split apart or merge together. It can drive the evolution of elaborate traits that position to natural selection" (Rosenthal, 2017). Even weak female mating preferences can produce strong selection on male traits (Kirkpatrick and Ryan, 1991). Mate choice dynamics can ultimately determine whether or not gene flow occurs and persists during hybridization (Brelsford and Irwin, 2009; Fisher et al., 2006). Previous mate choice studies in *Xiphophorus cortezi* have shown that in dichotomous visual choice tests, females spent significantly more time associating with the large Sc males than the small Sc males (Fernandez and Morris, 2008). Moreover, females given a choice between spotted and unspotted males, preferred Sc males in those populations in which female Sc frequency was lower but females from a high incidence of female Sc avoided spotted males (Fernandez and Morris, 2008). Authors of this study interpret this result as a way to avoid an increased probability of offspring having two copies of the oncogene *xmrk*, which would reduce their reproductive lifespan (Sedarti et al., 2002). Furthermore, another study showed that *Xiphophorus* helleri females spent more time with groups of spotted females than with non-spotted females under turbid, but not under clear water conditions and that they preferred spotted males in turbid, but not in clear water (Franck et al., 2001). Given these previous results that suggest Sc might be favored by sexual selection in closely related species and the fact that there is evidence of selection against the hybrids, I tested whether females of *X. birchmanni* or the hybrids favor spotted individuals during

mate choice. Furthermore, I explored if any other traits associated with Sc might indirectly provide spotted individuals with an advantage.

Besides looking at phenotypic traits associated with Sc, I also looked into personality traits. Personality is a consistent behavioral variation among individuals across time and contexts which contributes to individual fitness (Ariyomo and Watt, 2012, 2013). Traits related to personality are under strong ecological selection (Wolf and Weissing, 2012; Reale et al., 2010) and can significantly affect speciation through its effects on mate choice (Ingley and Johnson, 2014). Assortative mating according to personality has been widely reported resulting in benefits to choosers (reviewed in (Schuett et al., 2010)). On the other hand, choosers can avoid specific traits based on personality (Coleman et al., 2004; Sommer-Trembo et al., 2016). Moreover, personality can affect the process of mate sampling and therefore affect choosiness (the amount of resources a chooser invest in a mating decision or "how concentrated chooser responses are in particular areas of trait space" (Rosenthal, 2017)). Both male and female mating preferences have been found to be affected by personality in other fish systems (Bierbach et al., 2015; Sommer-Trembo et al., 2016). Swordtail females have previously been shown to prefer more active male courtship behaviors (Wong et al., 2011) and *X. birchmanni* and *X. malinche* show marked personality differences between them: *X. birchmanni* are bolder whereas *X. malinche* are slower to emerge from shelter and spend less time in open areas (Johnson et al., 2015). We know that boldness covaries with mate choice: bolder individuals of birchmanni were less choosy during mate choice assays (Personal communication: Gil Rosenthal). Therefore, I tested whether Sc could be associated with the shy/bold axis of personality and if there was a relation between mate choice for Sc and personality.

Methods

Mate choice trials

To test whether females preferred Sc males over unspotted males I tested females of *X. birchmanni* and hybrid females from an admixed population using animations (created using the anyFish software platform (Veen et al., 2013)) in dyadic female mate choice trials. This allows me to control all other sources of variation (phenotypic and behavioral) and isolate the effects of Sc. *X. birchmanni* females were offered *X. malinche* vs *X. birchmanni* choice, as a control of mating motivation, in addition to the Sc vs no Sc choice. CHAF hybrids were offered X. variatus vs *X. birchmanni* as a control of mating motivation in addition to the Sc vs no Sc choice. Females are tested for their mating motivation to ensure that if they do not show a preference for Sc, this lack of preference is not due to lack of motivation to choose. *X. birchmanni* females are offered a choice between *X. malinche* and *X. birchmanni* males because we know that *X. birchmanni* females prefer conspecifics over *X. malinche* (Wong and Rosenthal, 2006). CHAF hybrids females are admixed between *X. birchmanni* and *X. malinche* therefore testing them against its parental species would not yield a clear result. However, we know that CHAF hybrids co-exist with *X. variatus* and that there is no evidence of hybridization between them (personal observation). Therefore, it is reasoned that they will prefer *X. birchmanni* over *X. variatus*. Side and order effects were controlled by offering the same stimulus to each female twice, once on each side (the association times from both trials were summed for data analysis), and by offering randomly the control or the Sc choice first across females (Culumber and Rosenthal, 2013).

Trials were conducted by placing females in a 75x19x20 cm trial lane and allowing her to acclimate for five minutes. Lanes were virtually divided into three sections of equal size (one association zone on either side where stimuli were presented and a central neutral zone containing a small acrylic shelter). Each stimulus period lasted five minutes and were divided by five minutes intervals between them. Tanks were drained and refilled between females. Body shapes of animations were based on population means calculated from 42 morphometric landmarks according to the anyfish 2.0 user manual. The body texture was based on a representative male from each species. For the Sc vs unspotted trials, a black spot (Sc index (Sc area/standard length) of 0.05, "big" Sc) was manually added to the image. The animations follow the same simple conserved courtship display typical for both species in which a male swim on screen raises his dorsal fin and tilts toward the direction of the test lane and shimmies briefly then swims off screen. Animations were provided by Dr. Daniel Powell. Association time in each zone was used as a proxy for female preference, as has been shown to be predictive of female mate choice (Walling et al., 2010). If a female failed to visit both preference zones within 300 s, she was defined as unresponsive and excluded from analysis, as done in previous studies (Fisher and Rosenthal, 2006a; Verzijden et al., 2012). If a female failed to recover the control preference she was also excluded from the analyses. 12 females from CHAF, and 34 females from *X. birchmanni* data were included in the analyzes. Data was analyzed using Wilcoxon signed-rank tests.

Boldness trials

To investigate whether there is a correlation between the presence of Sc and personality traits, such as boldness/shyness, spotted and unspotted individuals were submitted to boldness

tests. These tests consisted of placing an individual in a big circular tank (90cm diameter) containing a small acrylic shelter. Latency to emerge from the shelter, and total time spent away from the shelter were measured as a proxy to boldness/shyness. A total of 24 *X. birchmanni* males (15 Non Sc, 9 Sc) and 20 *X. birchmanni* females (previously tested for preference for Sc) were included in the analysis. Non-parametric Kruskal-Wallis tests were performed to analyze the association between preference for Sc and personality in females.

Boldness trials were also performed to 41 hybrid males from an admixed population with high melanoma incidence, Chahuaco Falls. Kruskal-Wallis non parametric tests were performed to test whether spotted and unspotted males differ in their personality.

Phenotypic correlations

I used the individuals photographed for the GWAS (see Chapter 2) to investigate whether there are any phenotypic correlations associated with Sc in *X. birchmanni* that could favor Sc during mate choice indirectly. For phenotypic analysis only mature males were considered (N= 121). Following the phenotyping methods described previously in the GWAS (Chapter 2), standard length, body and dorsal fin standard length and depth were measured from photographs using ImageJ. Variables were corrected by standard length before analyzed.

Results

Net association time of *X. birchmanni* females with Sc males did not differ from the association time spent with unspotted males during the dyadic visual mate choice trials (Figure 6 a, Wilcoxon p-value= 0.602). Net association time of Chahuaco falls (admixed hybrid population) females with Sc males did differ from the association time spent with unspotted males during the visual trials (Figure 6 b, Wilcoxon p-value= 0.57).

X. birchmanni females that participated from the visual mate choice trails were also tested for boldness (N=20). Females that during the visual trials preferred Sc males spent significantly less time away from shelter than females that preferred unspotted males (Kruskal-Wallis p value = 0.049, Figure 7 a). In other words, they spent significantly more time under the protection from the shelter that the females that preferred unspotted males during the visual trials. However, their latency to emerge from the shelter was not significantly different (Kruskal-Wallis p value= 0.18, Figure 7 b). *X. birchmanni* males did not show a difference in their time spent away from shelter (Kruskal-Wallis p value = 0. 08, Figure 7 c) nor their latency to emerge (Kruskal-Wallis p value = 0.70, Figure 7 d) between spotted and unspotted individuals.

Chahuaco Falls (admixed hybrid population) spotted males spent significantly less time away from shelter than unspotted males (Kruskal-Wallis p value= 0.045, Figure 3.4 a) and took longer times to emerge from shelter (Kruskal-Wallis p value= 0.042, Figure 3.4 b)

X. birchmanni spotted males were significantly smaller than unspotted males (ANOVA p value= 0.016, Figure 8 a). Height index was calculated by measuring the body depth (the distance between the beginning of the dorsal fin and the beginning of the annal fin) and diving it by the standard length. Spotted males had a marginally non-significant higher body height indexes

than unspotted individuals (ANOVA p value = 0.056, Figure 8 b). Spotted and unspotted males did not differ in their dorsal fin width index (ANOVA p value= 0.67, Figure 8 c) but they showed a marginally non-significant higher dorsal fin (ANOVA p value = 0.059, Figure 8 d). Overall, spotted individuals were smaller but had bigger dorsal fins than unspotted individuals.

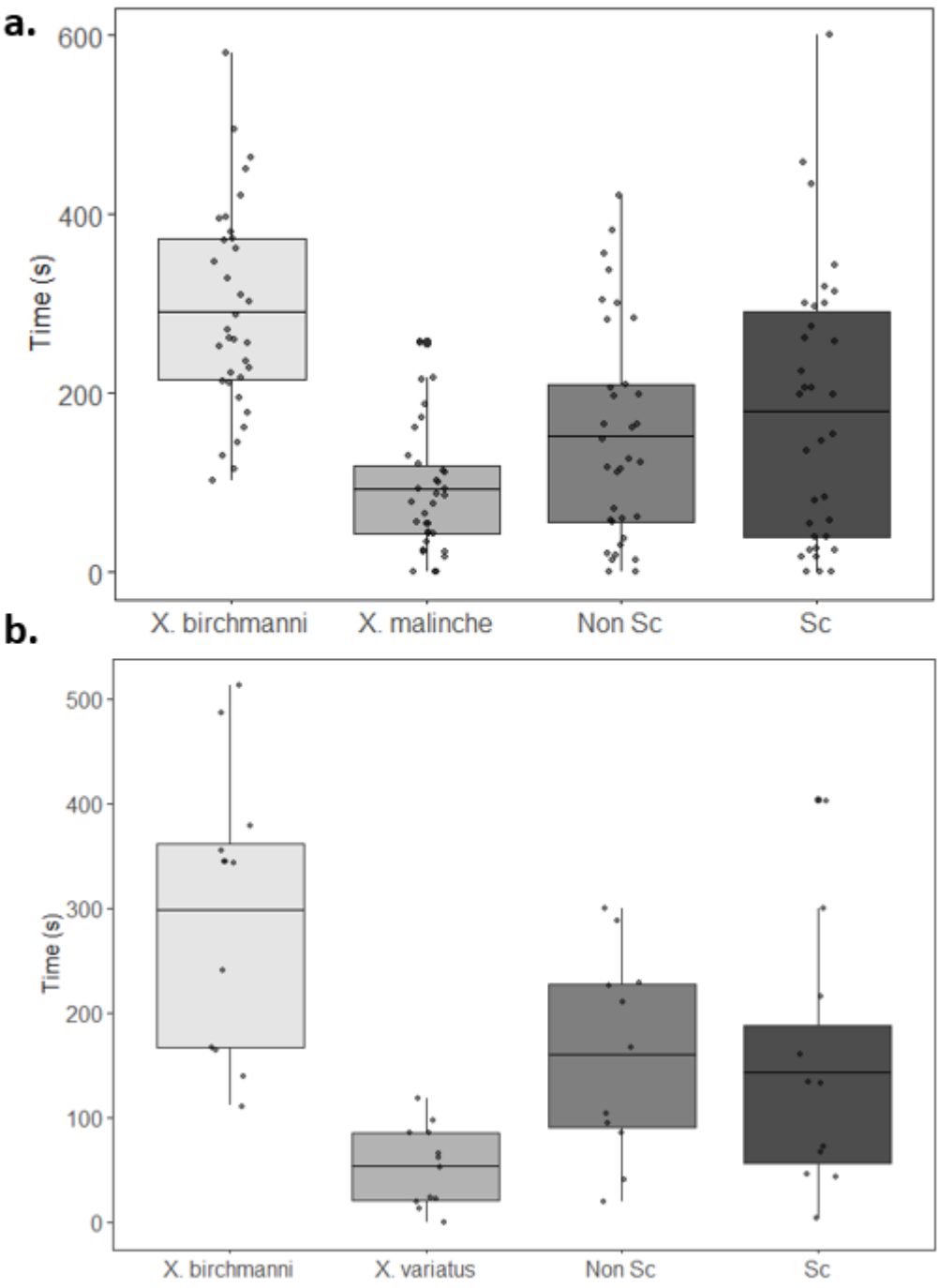


Figure 6. a) *X. birchmanni* association times during dyadic female visual mate choice trials. b) Chahuaco falls (admixed hybrid population) association times during dyadic female visual mate choice trials. The center line denotes the median value (50th percentile), while the colored box contains the 25th to 75th percentiles of dataset. The black whiskers mark the 5th and 95th percentiles, and values beyond these upper and lower bounds are considered outliers.

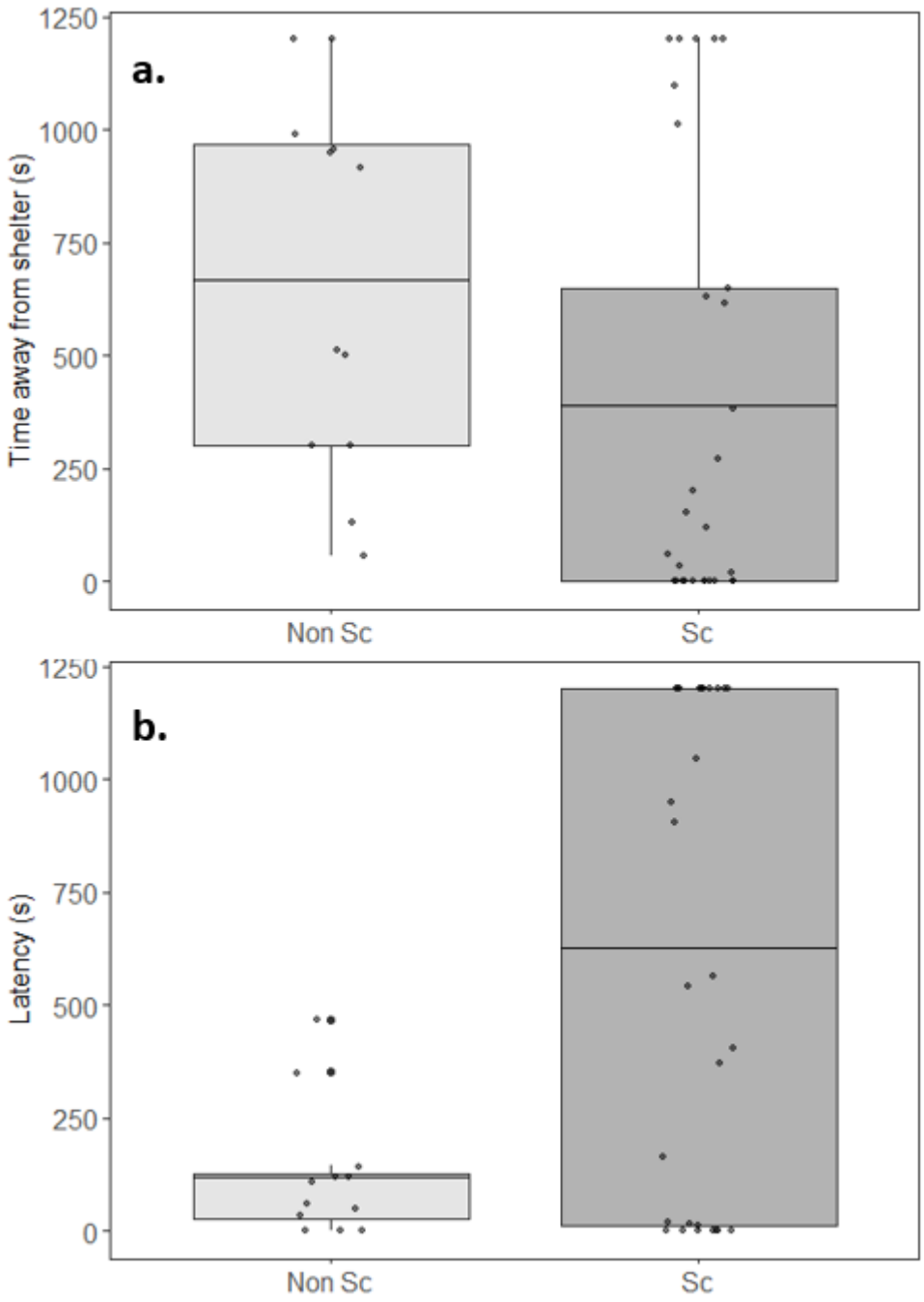


Figure 7. a) Time spent away from shelter by unspotted and spotted hybrid males from an admixed population, Chahuaco Falls. b) Latency to emerge from shelter by unspotted and spotted hybrid males from an admixed population, Chahuaco Falls. The center line denotes the median value (50th percentile), while the colored box contains the 25th to 75th percentiles of dataset. The black whiskers mark the 5th and 95th percentiles, and values beyond these upper and lower bounds are considered outliers.

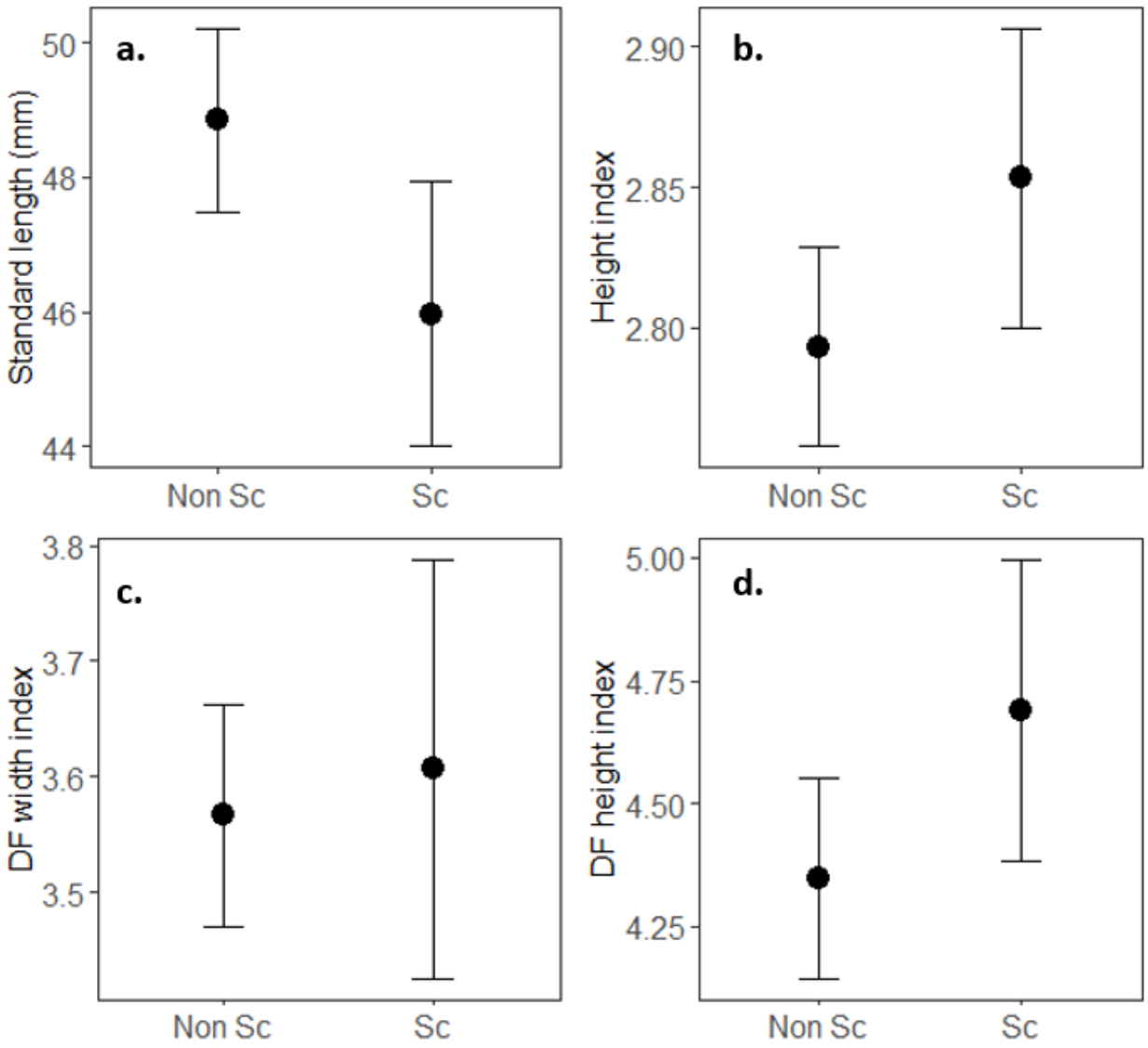


Figure 8. a) *X. birchmanni* unspotted (Non Sc) and spotted (Sc) males standard length. b) Body depth corrected by the standard length for unspotted and spotted *X. birchmanni* males. c) Dorsal fin width corrected by standard length for unspotted and spotted *X. birchmanni* males. d) Dorsal fin height corrected by standard length for unspotted and spotted *X. birchmanni* males. The plot shows the mean, and whiskers indicate two standard errors of the mean.

Discussion

Female mating preferences for traits that make males more visible could have initially evolved if there is genetic covariance between the trait and preference (reviewed in (Rosenthal, 2017), Chapter 15). Fisher (Fisher, 1930) suggested that "males and females inherit genes for both courter traits and chooser preferences but each is expressed only in one sex" (Rosenthal, 2017). Thus, ornaments evolve as a consequence of the genetic correlation between ornament and mate choice (Jones and Ratterman, 2009; Lande, 1981). Alternatively, preferences and traits can be decoupled and be the result of a preexisting sensory bias (Ryan, 1990a, 1990b; Lande, 1981) in which a courter addresses hidden preferences (reviewed in (Ryan and Cummings, 2013) and (Rosenthal, 2017), Chapter 13). Regardless of how the preference arouse, once present, these preferences can be maintained through direct selection because of reduced mate search costs (Westcott, 1994; Reynolds and Gross, 1990).

Fernandez and Morris, 2008, found that females of X. *cortezi* do prefer spotted males (when female Sc frequencies are low. The female Sc frequency in those populations that preferred spotted males was similar to female Sc frequency of my *X. birchmanni* population. Data not shown.) In my visual mate choice trials, females from *X. birchmanni* did not prefer spotted males over unspotted males (Figure 6 a). Although these results may seem contradictory at first, there are several differences between them that might explain the conflicting results. In my trials, I used videos to stimulate females and isolate specifically the effect of Sc on mate choice from other courting aspects such as behavior. Fernandez and Morris, 2008 (Fernandez and Morris, 2008) used live males to which Sc was artificially painted with a dye. Although they repeated trials with alternated treatments within the same males, to control for any unforeseen behavioral or

phenotypical differences between the pairs of stimulus males, it is not possible to discard this might be a confound effect in their tests. A recent review, highlights the control over stimuli and the consistency of video playback (Powell and Rosenthal, 2017).

Females from CHAF, hybrid admixed population, with a high incidence of melanoma did not show preferences for spotted males (Figure 3. 2 b). Given the strength of selection against spotted juveniles observed (Figure 5 a) Powell et al., 2020 estimated that either high levels of gene flow from *X. birchmanni* or mating advantages in individuals with large spots would be required to maintain spotting at observed frequencies in hybrid populations. Future studies should attempt to characterize this further.

The difference in preference observed between studies (at least in populations where Sc occurs in low frequency) might be related to differences in melanoma development between both species. In more than three years of sampling *X. birchmanni* I have never seen an individual with melanoma in the wild. Fernandez and Morris found five males and one female of the 99 individuals in a single day of collecting. Mate choice decisions are subject of selection themselves (reviewed in (Rosenthal, 2017), Chapter 14). Although Sc individuals appear to be more conspicuous to predators, if we consider the lack of differences between juvenile and adult Sc frequencies (Figure 5 a) and the lack of melanoma found while sampling natural populations of *X. birchmanni* as an indicator of presumably weak (or no) selection against Sc, we therefore would not expect mate choice to favor the trait. However, further studies should explore several aspects that could be relevant for mate choice. *Xiphophorus* typically inhabit clear rivers. Yet, during the rainy season many streams become turbid. Since it was shown that turbidity changes preference for Sc (Franck et al., 2001), further studies addressing this should be performed.

A classic review showed that when females prefer traits that deviate from the population mean, they usually prefer traits of greater quantity because of the increased sensory stimulation (Ryan and Keddy-Hector, 1992). It would be interesting to determine whether there are differences in courting intensity or behavior between spotted and unspotted males. In relation to sexual behavior, we know that melanocortins have a positive effect on the production of sexual hormones (Eberle, 1988) and that darker males were sexually more active and had higher levels of testosterone (Ducrest et al., 2008). On this line, it would be also relevant to test whether spotted and unspotted males differ in their sexual pheromones. A recent dissertation from our lab identified specific peaks in *X. birchmanni*'s male pheromones that distinguishes it from *X. malinche.* Moreover, male *X. birchmanni* show two distinct phenotypic clusters: large and ornamented and small and drab, which as well differ in their pheromone profile. Spotted caudal is predominantly a male trait (Sedarti et al., 1995). Given this, it is highly likely Sc could have a downstream effect on male cue production. Future female mate choice olfactory trials using spotted and unspotted male cues, would be a starting point to test this hypothesis.

Even if Sc is not under direct selection during mate choice, it might be indirectly selected if an associated trait is. We know female swordtails typically prefer larger males (Cummings et al., 2008; Ryan and Cummings, 2013; Fernandez and Bowser, 2010) and although in *X. cortezi* Sc males are bigger (Fernandez and Bowser, 2010), I found that *X. birchmanni* spotted males are smaller than unspotted ones (Figure 8 a). We also know that *X. birchmanni* females prefer males with smaller dorsal fins (Fisher and Rosenthal, 2006b) and spotted individuals showed a tendency to have bigger dorsal fins (Figure 8). Despite the disadvantage this may pose during female mate choice, the intended receivers of this signal are males, not females and therefore, intrasexual selection can counterbalance the forces of intersexual selection (Fisher and Rosenthal,

2006b). Indeed, even though the results from this chapter do not support the claim that sexual selection, via mate choice, is contributing to the maintenance of the polymorphism, there are other mechanisms by which sexual selection can operate such as intrasexual selection in the form of male competition. Associations studies between spotted caudal and aggression have shed contradicting results in *Xiphophorus*. Whereas male aggression has been demonstrated to be positively correlated in X. *cortezi* (Fernandez, 2010), no effect was found in X. hellerii (Franck et al., 2001). Future studies should aim to characterize male interactions between spotted and unspotted individuals in this system.

Finally, we know that personality, specifically boldness, can affect mate choice choosiness. For example, high-exploratory zebra finch females showed lower selectivity and lower preference scores (David and Cézilly, 2011). Indeed, boldness trials showed that bolder females showed a preference for spotted individuals (Figure 7 a). This preference instead of reflecting an actual preference for spotted males, it might be result of females being more relaxed in their mating decisions. Males however, did not show differences in their boldness between spotted and unspotted individuals in *X. birchmanni* (Figure 7 c-d) but did show a correlation in CHAF hybrids (Figure 3.4 a). This is an interesting result if we consider the selection pressure difference between both populations (Figure 5 a). Spotted individuals being shyer is in line with what is expected if spotted individuals are being differentially predated.

CHAPTER IV

ADGRE5 ACTS AS A MELANOMA TUMOR SUPPRESSOR IN NATURALLY HYBRIDIZING *XIPHOPHORUS*.

Introduction

Xiphophorus has been a dominant system in the study of melanoma since Gordon (Gordon, 1931) discovered that certain hybrids of the platyfish (*X. maculatus*) and swordtails (*X. hellerii*) develop a highly malignant melanoma (Meierjohann and Schartl, 2006). From the classical crossing experiments, we know that theoretically, there are two genetic explanations to how hybrid melanoma is generated. One it that the observed hybrid melanoma develops because certain individuals of the platyfish carry an oncogene which is under control of a tumor regulator. In this case, the tumor regulator acts as a tumor suppressor and therefore the oncogene effect on the platyfish is only a local dysplasia of melanocytes visible as a black pigment spot of macromelanophores. Since the oncogene and the tumor regulator are located in different linkage groups, when a platyfish hybridizes with a swordtail that lacks both genes some individuals will express the oncogene without the control of the tumor suppressor and therefore develop malignant melanoma. An alternative explanation could be that instead of the platyfish having an oncogene repressed by a tumor suppressor, the platyfish has a pigmentation gene that interacts with a tumor enhancer from the swordtail. In that case, the pigmentation gene makes a minor contribution to pigment cell proliferation, only to become malignant once it interacts with the tumor enhancer (Schartl, 2008).

However, biochemical data and transgenic experiments have shown and identified *xmrk* as a potent tumor gene, whose expression is necessary and sufficient for tumor development (Schartl et al., 1999; Schartl et al., 2010). Even though the main oncogene (*xmrk*) has been identified and isolated almost 30 years ago (Wittbrodt et al., 1989; Wittbrodt et al., 1992), we still have not yet been able to identify the tumor regulator gene. A recent study (Powell et al., 2020) used the naturally hybridizing *X. birchmanni* and *X. malinche* system to carry out an admixture mapping study. Using a high melanoma incidence population, 209 individuals were low coverage whole genome sequenced to infer local ancestry and an admixture mapping for spot presence and melanocyte invasion was performed. The study revealed a strongly associated region on chromosome 21 where spotting correlated with *X. birchmanni* ancestry (which coincides with the same region identified by the GWAS, Chapter 3), and an additional region on chromosome 5 associated with *X. malinche* ancestry. The region in chromosome 21, as expected, corresponds to *xmrk.* The region in chromosome 5 is a gene called *cd97* or *adgre5* (I will call it *adgre5* from now on because it is the HUGO Gene Nomenclature Committee approved name). *adgre5* belongs to a family of G- protein couple transmembrane receptors identified to have varied roles in cell adhesion and signaling (Yona et al., 2008). In humans, *adgre5* is widely expressed in lymphocytes, monocytes, macrophages, dendritic cells, granulocytes and smooth muscle (Eichler et al., 1997). Although *adgre5* plays a vital role in the adhesion and migration of healthy immune cells, it is also a mediator of invasion in a variety of human cancers (reviewed in (Safaee et al., 2013)). Accordingly, *adgre5* is upregulated in RNA-seq data of invasive melanoma, compared to normal spots in the hybrids and five aminoacid changes, one that occurs in a conserved epidermal growth factor–like calcium binding, were found between *X. birchmanni* and *X. malinche*. This gene was also found to be expressed at low levels at the caudal fin tissue in *X. birchmanni*,

regardless of the spotting phenotype, but at higher levels in *X. malinche* and hybrids (natural and artificial) (Powell et al., 2020).

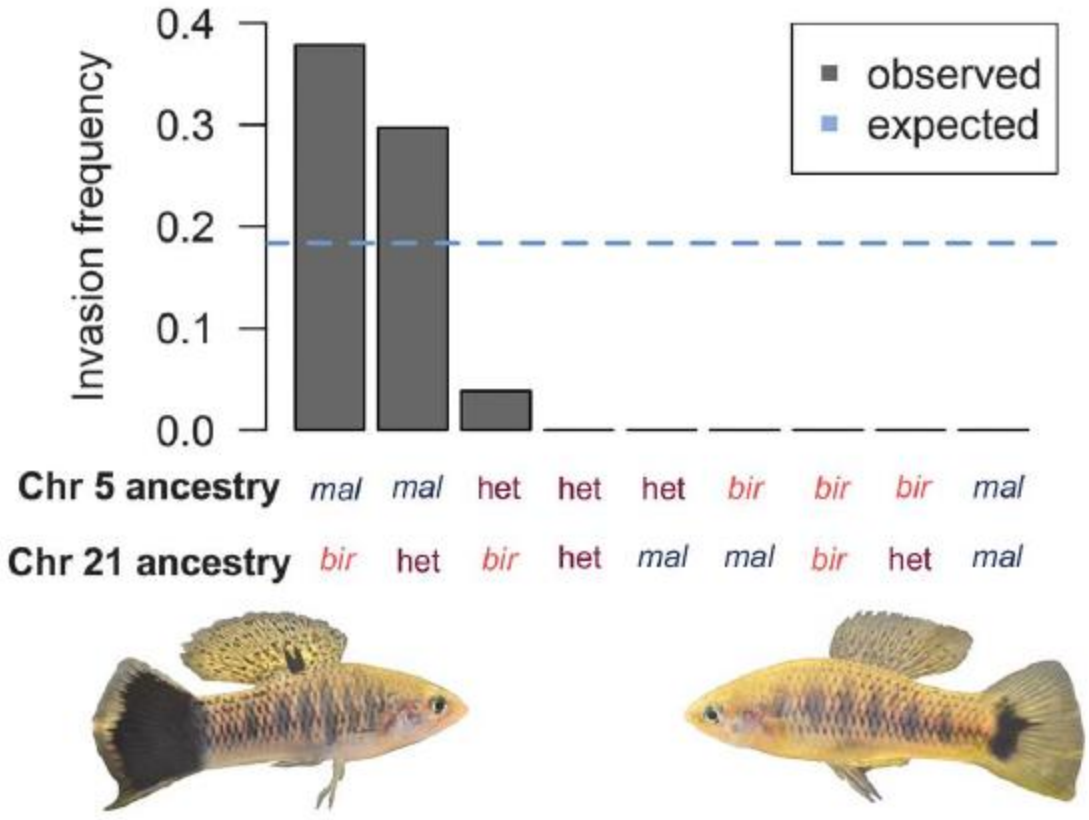


Figure 9. Proportion of individuals with melanoma as a function of ancestry at the associated regions on chromosome 5 and chromosome 21. The expected proportion of cases if melanoma risk were equally distributed among individuals with at least one birchmanni allele at chromosome 21 is indicated by the blue dashed line. Published in Powell et al. 2020

The association with ancestry at these two sites and melanoma is not random: in order for increased melanoma risk, there has to be at least one copy of *X. birchmanni xmrk* and *adgre5* has to come from *X. malinche* (Figure 9). In fact, heterozygous individuals for *adgre5* had smaller spots than *X. malinche* homozygous individuals, which repercusses in lower probability of becoming invasive. This interesting result can be interpreted in two ways: (i) the *X. malinche* allele of *adgre5* is a tumor enhancer that interacts with *xmrk* and triggers a change from benign to malignant pigmentation, or (ii) malignant melanoma only arises in the absence of the tumor suppressor *X. birchmanni* allele of *adgre5.* These two hypotheses, although they would produce same pattern of ancestry associations, imply very different subjacent mechanisms explaining those patterns. Whereas the first hypothesis suggests the action of a tumor enhancer, the second one advocates for the presence of a tumor suppressor. These differences are crucial in the repercussions they may have for future research and therefore this chapter of my dissertation focuses on functionally testing the role of these alleles by analyzing their effect independently on cell growth and in vivo.

I performed cell culture analysis of cell growth and migration using melan-a cells. Melan-a cells are the first known line of non-tumorigenic mouse melanocytes (Rossato et al., 2014). They offer an excellent equivalent non-tumorigenic line for studies of the cellular and molecular basis of melanoma malignancy (Bennett et al., 1987), specially for studies aiming to understand the mechanisms triggering the change from benign pigmentation to malignant melanoma. If we consider that one of cancer's main characteristics is unlimited cell growth and potent metastatic properties (Otto, 1956), I would expect a tumor enhancer gene to cause cells to grow and migrate more whereas tumor suppressor genes to cause the opposite effect. Even though melanoma modifying genes are easiest to investigate in cell culture models, animal models are imperative if we

aim to evaluate their importance in the context of the whole organism. Moreover, the advantage of establishing transgenic organisms is that we can study the isolated effects of the induced gene without other unknown possible genetic interactions. Therefore, in order to study the interaction between tumor modifier candidate genes and the oncogene in vivo, I developed stable transgenic lines using the already established *tg(mitf:xmrk)* transgenic medaka line. I used medaka rather than *Xiphophorus* because *Xiphophorus* is a live bearing fish and therefore transgenic approaches are not easily available. In *Xiphophorus*, *xmrk* is expressed only in the melanocytes but given its potent oncogene characteristics, it can transform other cell types. Thus, in the medaka transgenic model, *xmrk* expression needed to be driven by a pigment cell specific promoter: microphtalmia related transcription factor (*mitf*). *mitf* is a basic helixloop-helix-leucine zipper transcription factor whose expression is specific to pigment cells and sufficient to differentiate blastula derived stem cells to melanocytes (Béjar et al., 2003). With these experiments I aim to isolate the effects of each species *adgre5* allele. I expect that either *X. birchmanni adgre5* transgenic fish develop less melanoma, or *X. malinche adgre5* transgenic fish develop more. These results will allow us to determine whether *adgre5* is a tumor suppressor or a tumor enhancer.

Methods

Cell culture

All cells were thawed under sterile conditions and Melan-a cells were maintained at 37°C and 5% CO2 in an incubator. Melan-a cells were grown in DMEM with pyruvate (Gibco), with 10% FCS and 1% penicillin/streptomycin. Cells were regularly monitored for *Mycoplasma* contamination. Depending upon the confluency, cells were regularly sub-cultured (1-2 times a

week). For this purpose, cells were detached with 1x trypsin in EDTA/PBS and then passaged. For long-term storage, cells were frozen in freezing medium, containing DMEM, 20% FCS and 10% DMSO and were stored at -80°C.

Cell lines

To achieve doxycycline-inducible expression cell lines, I used the vector pSB-ET-iE (M. Gessler, Dept. of Developmental Biochemistry, University of Wurzburg), which allows integration of my genes by *sleeping beauty*-mediated transposition. Dox inducible promoters allow elegant experimental designs in which the expression of the targeted gene can be precisely regulated. Here, the responsive T6 promoter drives expression of my gene and EGFP, with an IRES site between them (Figure 10 a). After transposition, these cells were selected with 1 μg/ml puromycin for 2 weeks.

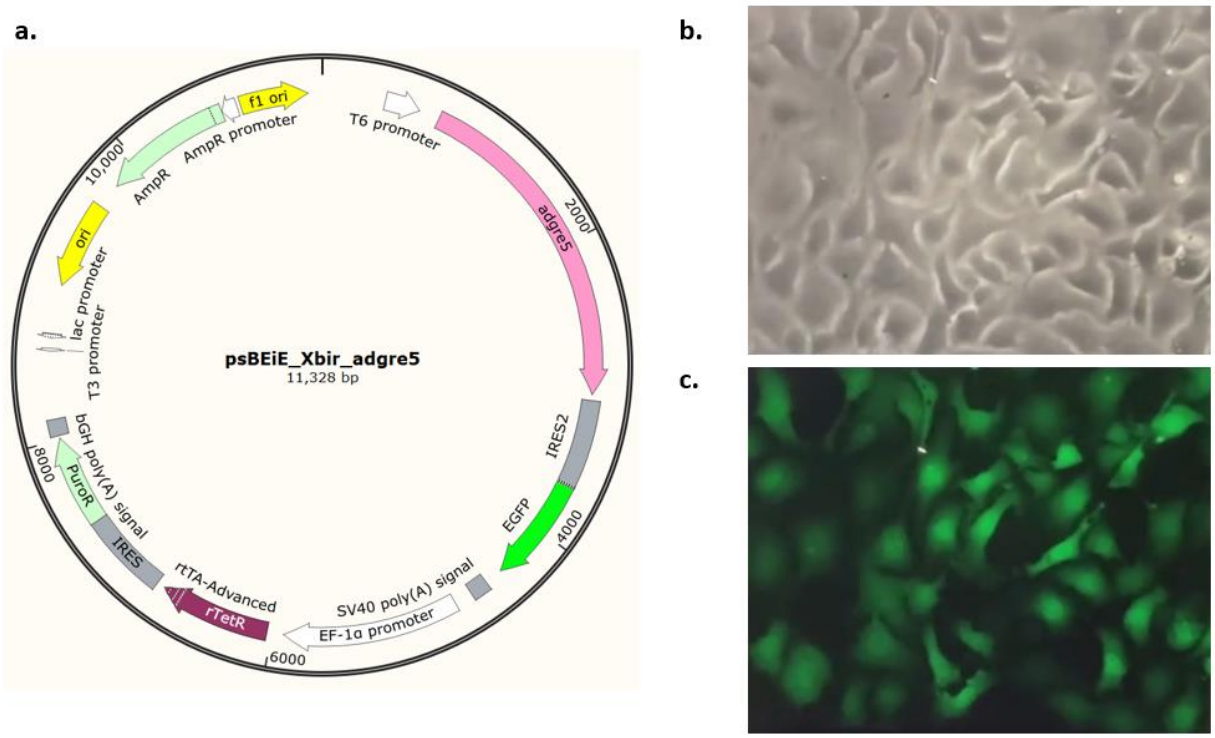


Figure 10. a) Map of the used for cell transfections. *X. birchmanni* and *X. malinche* plasmids are exactly the same except ORF of the *adgre5* which corresponds to each species. b) Bright field view of melan-a cells under an inverted microscope with a 63x objective c) Fluorescence image of b).

First, I amplified *X. malinche* and *X. birchmanni* alleles of *adgre5* using primers with overhang XbaI and ClaI restriction enzyme sites. PCR amplification was done using cDNA from tissue samples of organisms collected in Coacuilco (*X. birchmanni*) and Chicayotla (*X. malinche*). PCR amplification was carried out with the High Fidelity Q5 (New England Bio Labs #M0491S). Respective restriction enzymes were used to clone the PCR product into the vector and wild type melan-a cells were transfected using Fugene transfection reaction (according to manufacturer's protocol). As a result, pSB-ET-iE_*X. malinche_adgre5*, and pSB-ET-iE_*X. birchmanni_adgre5* stable cell lines were generated.

RNA extraction, cDNA synthesis and qPCR

From three independent replicates of cells, per treatment, per cell line, total RNA was extracted from freshly harvested cells or cell pellets frozen at -80 °C. RNA isolation was performed using TRIzol reagent according to the manufacturer's protocol. DNA digestion was performed with DNase I for 1 h at 37 °C. RNA concentration was then determined by a NanoDrop spectrophotometer. Subsequently, 100-4000 ng of RNA was reversely transcribed using a RevertAid First Strand cDNA Synthesis Kit and random hexamer primers according to the manufacturer`s instructions. Fluorescence-based RT-qPCR was performed and analyzed with a Mastercycler ep Realplex using SYBR Green reagent. Gene expression was normalized to a housekeeping gene (*hprt*), which remained unaltered under the treatment conditions using delta-delta Ct method (Figure 11). Moreover, cells were checked for GFP expression an inverted fluorescence microscope with a 63x objective was used to further verify the success of the transfection (Figure 10 c).

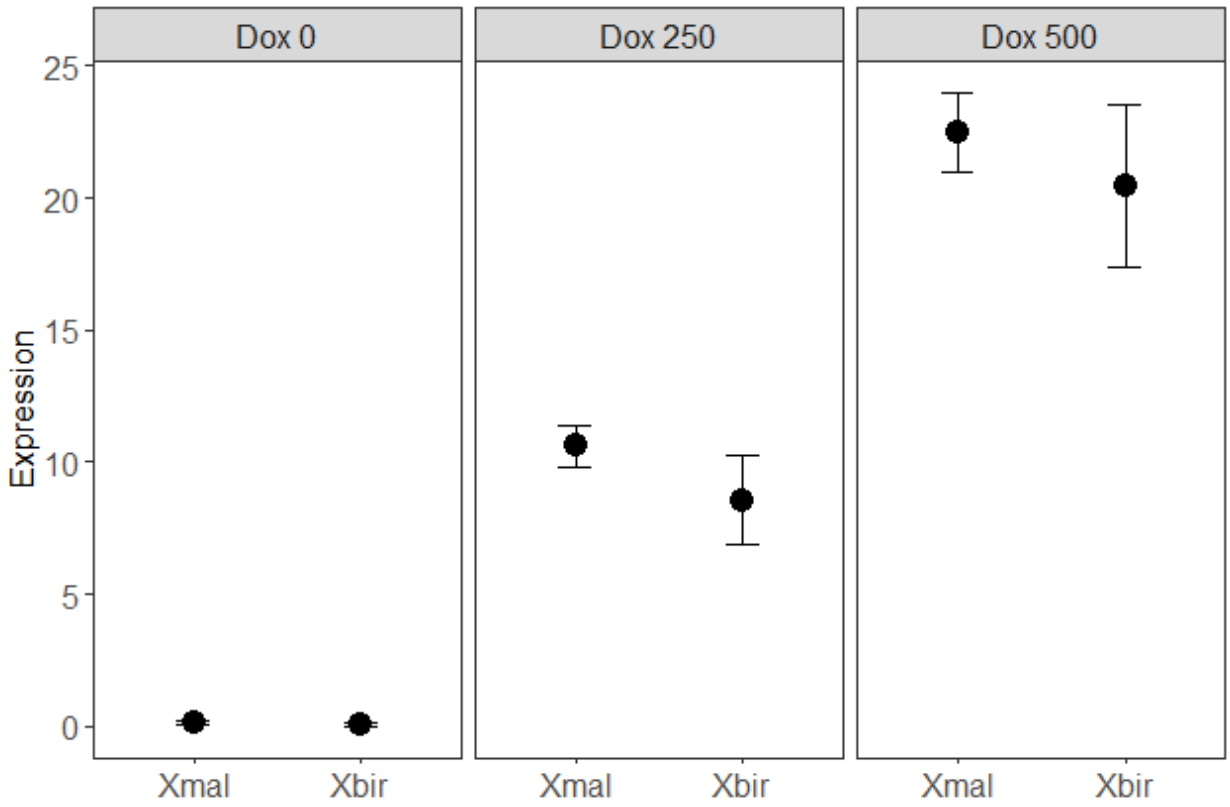


Figure 11. qPCR data of *adgre5* expression relative to the housekeeping gene *hprt* in melan-a cells after one day of dox induction. Xmal = stable cell line transfected with the *X. malinche* allele of *adgre5*. Xbir = stable cell line transfected with the *X. birchmanni* allele of *adgre5*. The number besides Dox indicates how many mg/ml of Dox were added to cell culture media. The plot shows the mean, and whiskers indicate two standard errors of the mean.

Cell growth assay

To investigate the cellular growth MTT assays were performed. Cells were counted and seeded in triplicate at equal density (1-2 x 10^3 cells/well) in 96-well plates. Cells were assigned a dox 0 or dox 500 (500 mg/ml) induction by adding respective dox concentrations to cell media. 2, 3 and 4 days after treatment, 5 mg/ml of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) was added at a ratio of 1:5 (MTT:medium) to each well. After 2 hours of incubation at 37°C, the medium was aspirated and 150 µl of DMSO was added to each well. The plate was then incubated at room temperature for 15 minutes on a shaking device. To measure formazan accumulation, the optical density at 590 nm with a reference filter of 620 nm was measured using a micro-plate reader. Quintuples for each treatment, species and duration of the experiments were performed. Cell growth was calculated as the difference in optical density at 590 nm observed in dox 500 induced cells minus the optical density at 590 nm observed in dox 0 cells. Statistical differences between *adgre5 X. birchmanni* and *X. malinche* alleles in cell growth were determined with an ANOVA and a Tukey post hoc test.

Transwell migration assay

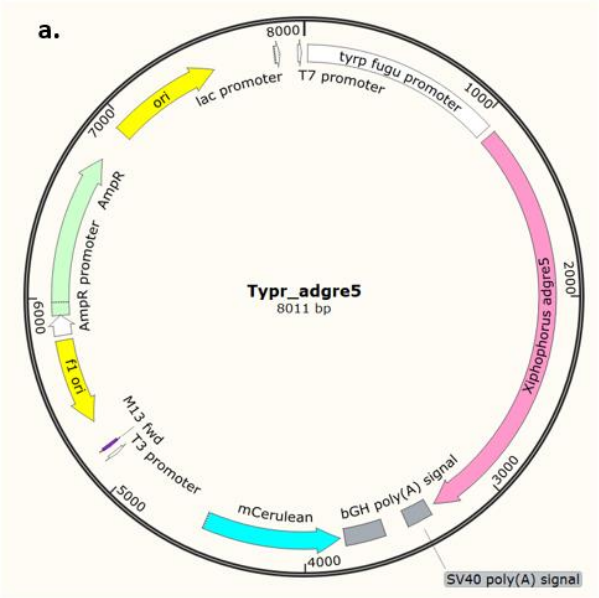
To study the migration ability of cells, transwell migration assays were performed. $2x10^3$ cells, previously starved in 1% dialyzed FCS, and either dox 0 or dox 500 (500mg/ml), for 24 hours were applied in the upper layer of uncoated transwell inlays with 8 µm pore diameter in 24-well plates. To stimulate migration, medium containing 10% FCS, and either dox 0 or dox 500 (500mg/ml), was applied to the lower layer of the transwell and cells were allowed to migrate for 16 hours. Each assay was performed in triplicates per cell line and per dox treatment. Non-migrated cells were removed by cotton swabs and migrated cells were fixed with methanol,

stained by 0.2% crystal-violet dye in 2% ethanol for 15 minutes and washed with PBS. The membrane was cut out of the transwell inlay and embedded with Mowiol (polyvinyl alcohol) on microscope slides. Images of the membranes were made, and cells were counted under the microscope. Migration was calculated as the difference between number of cells that migrated in dox 500 induced cells minus the number of cells that migrated in dox 0 cells

Isolation of the *Xiphophorus adgre5* and construction of expression vectors with *tyrp fugu* promoter

Even though a powerful oncogene's (*xmrk*) expression can thrive when controlled by the *mitf* promoter, that is not necessarily the case for a tumor modifier candidate gene. Moreover, there is evidence (Dr. Frederik Helmprobst, personal communication, paper in review) that the *mitf* promoter is an overall relatively weak promoter. Therefore, *adgre5* was tested driven by a stronger pigment cell specific promoter: *Fugu* rubripes Tyrosinase-related Protein 1 gene (*tyrp*) (Zou et al., 2006). *Tyrp* is an enzyme required for melanin synthesis and is specifically expressed in melanocytes or melanophores (del-Marmol and Beermann, 1996).

To isolate the *adgre5* gene from *Xiphophorus*, a high-fidelity PCR (using Q5 Taq enzyme) was performed from cDNA extracted from either *X. birchmanni* or *X. malinche* using overhang primers containing the XbaI restriction enzyme cutting site.



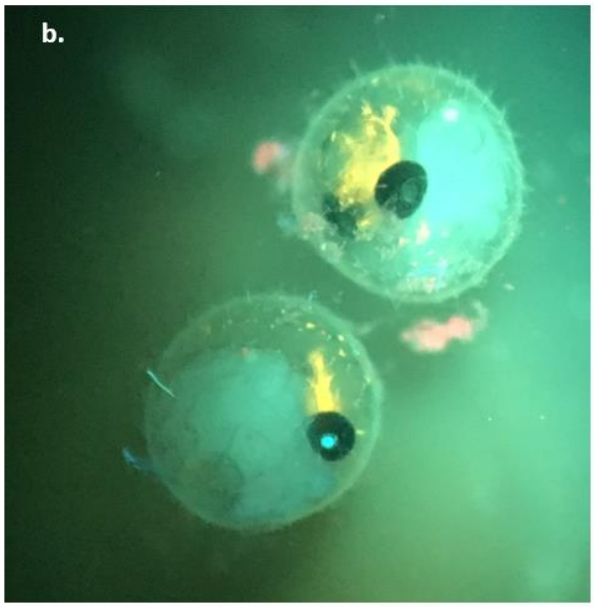


Figure 12. a) Map of the used for embryo injections. *X. birchmanni* and *X. malinche* plasmids are exactly the same except ORF of the *adgre5* which corresponds to each species. b) mCelurean effect: when DNA is successfully integrated blue eyes (bottom) are visible. Wild type embryo (top).

I extracted mCerulean with its promoter and a bGH poly(A) site using KpnI restriction enzyme from the brainbow vector (kindly provided by Brigitta Wilde). Then I linearized the pIsceI-typr-pa 3 plasmid (kindly provided by Dr. Mateus Adolfi) containing the *tyrp fugu* promoter, with KpnI restriction enzyme and ligated them together. The vector created was subsequently digested with XbaI and ligated to the previously digested with XbaI *adgre5* PCR fragments. Therefore, specific plasmids were created containing either the *X. malinche* or *X. birchmanni* alleles of *adgre5* under the expression of the *tyrp fugu* promoter (Figure 12 a)

In order not to extend myself I will summarize each cloning step briefly. Each step was performed according to manufacturer's protocols. Before cloning, each vector was transformed in an agar plate with competent bacteria overnight. Then following day, single colonies were collected and grown overnight in LB medium + antibiotic. Using a Quiagen Mini Prep Kit I extracted the plasmid from the bacteria and then enzyme restrictions were set up. After restrictions, a gel was run and the appropriate band was cut out from the gel and purified using the Quiagen Gel Extraction Kit. Subsequently, another gel is run to verify a single band has been extracted and then the ligation reaction is set up. This newly formed vector need to be transformed again overnight and the following day, a colony PCR was performed. Positive bacteria colonies are grown on LB medium + antibiotic overnight and the following day the plasmid was extracted using the Quiagen Mini Prep Kit. Plasmids were verified by sequencing to check for mutations during cloning.

Production of stable transgenic medaka

Transgenic lines were established using the genetically mixed and not inbred Carbio medaka strain and they developed pigments with a 100% penetrance (Schartl et al., 2010). An interesting characteristic of the *mitf*:*xmrk* transgenic medaka is that the melanoma they develop

shows striking ultrastructural and histopathological similarities to human melanoma. As in humans, the fish melanoma cells contain strongly deformed nuclei and immature melanosomes, which are characteristic for poorly differentiated melanoma and it often metastasizes to the liver (Gimenez-Conti et al., 2001). Further information regarding the isolation of the medaka *mitf* promoter and the production of the stable *tg*(*mitf*:*xmrk*) transgenic lines can be found in Schartl et al 2013 (Schartl et al., 2010).

For the generation of stable transgenic lines, the injection meganuclease protocol was used since it was demonstrated to be more effective than injecting just the plasmid because it reduces mosaic expression, increases frequency of positive founder fish and increases germline transmission rates (Thermes et al., 2002). One-cell stage of the *tg*(*mitf*:*xmrk*) medaka embryos (strain: Carbio) were injected into the cytoplasm with approximately 15–20 pg of total DNA plasmid in a volume of 500 pl injection solution containing I-SceI meganuclease. Adult F0 fish were mated with each other and the offspring were tested for the presence of the transgene by screening of blue eyes under UV light (mCerulean effect). Siblings from positive F1 fish were raised to adulthood and tested by PCR from dorsal fin clips, as described in (Altschmied et al., 1997).

All animal studies have been approved by the authors' Institutional Review Board (Animal Welfare Officer of the University of Wurzburg). Adult fish were maintained under standard conditions (Kirchen and West, 1976) with an artificial photoperiod (10 hours of darkness, 14 hours of light) to induce reproductive activity. Clusters of fertilized eggs were collected 0.5–1 hour after the onset of light and kept in a rearing medium containing 0.1% NaCl, 0.003% KCl, 0.004% CaCl2 × 2H2O, 0.016% MgSO4 × 7H2O, and 0.0001% methylene blue.

F1 Fish (8 *tg*(typr:Xmal_*adgre5*/*mitf*:*xmrk*), and 7 *tg*(typr:Xbir_*adgre5*/*mitf*:*xmrk*)) were anesthetized in MS-222 and photographed with a Nikon D300 digital camera with a Tamron SP 90mm F/2.8 1:1 Macro lenses. Individuals were quantified for hyperpigmented melanin area from the images using ImageJ (Schneider et al., 2012). Data fit normality principles and therefore it was analyzed with an ANOVA and a Tuckey post hoc test.

Results

Melan-a cells integrated and expressed the vectors containing the *X. birchmanni* or *X. malinche* allele *adgre5* and responded accordingly to different levels of dox inductions (Figure 11, ANOVA p value < 0.001 between dox treatments) but no differences were observed in the levels of expression of *adgre5* between cell lines containing the different alleles (Figure 11, ANOVA p value > 0.05 for differences between *X. birchmanni* and *X. malinche adgre5* alleles). Melan- a cells transfected with the *X. birchmanni* allele of *adgre5* showed less growth than melan-a cells transfected with the *X. malinche* allele of *adgre5* after 3 and 4 days of cell growth (Figure 13 a, ANOVA p value < 0.0001, Tukey post hoc pvalues, X.bir-X.mal day 2 = 0.49, X.bir-X. mal day 3 = 0.0006, X.bir-X.mal day 4 = 0.0019). Melan-a cells transfected with the *X. birchmanni* allele of *adgre5* migrated less than melan-a cells transfected with the *X. malinche* allele of *adgre5* (Figure 13 b, t test p value = 0.00582)

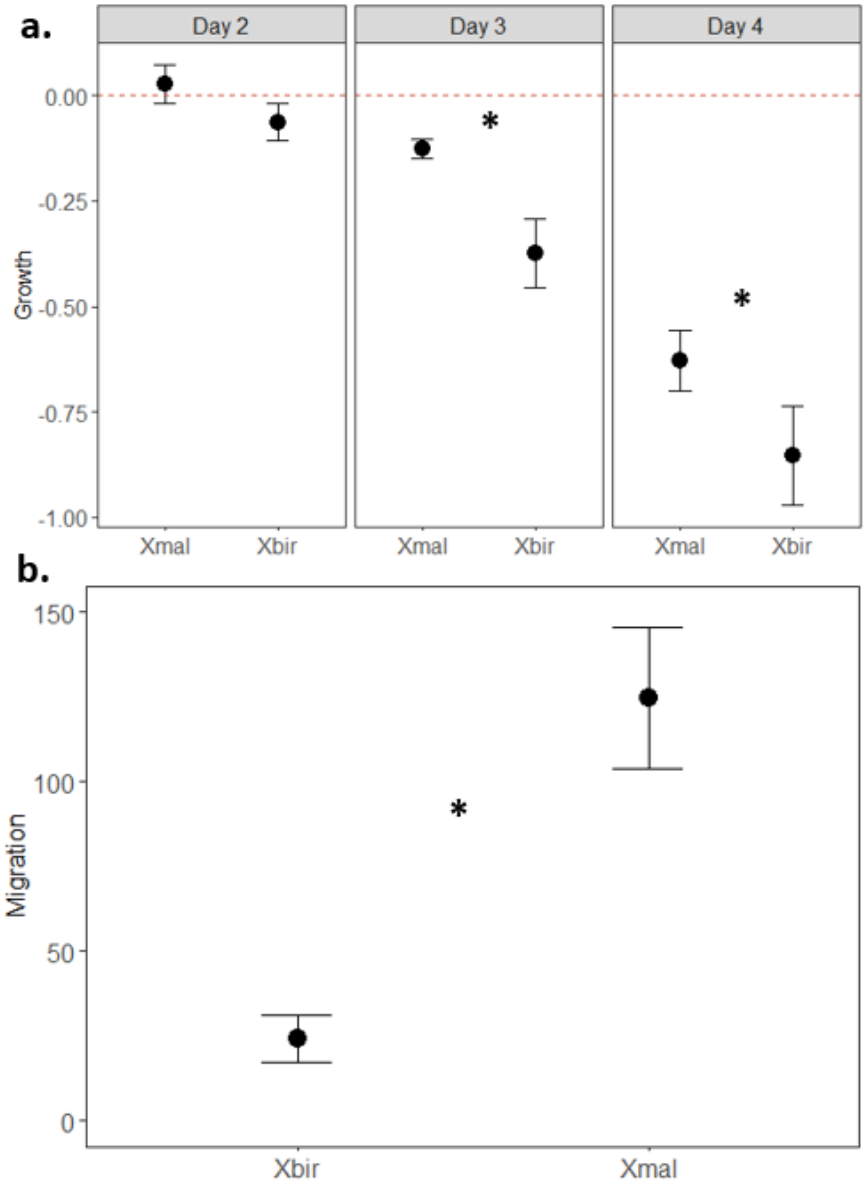


Figure 13. a) Cell growth (calculated as the difference in optical density at 590 nm observed in dox 500 induced cells minus the optical density at 590 nm observed in dox 0 cells.) for each cell line and for each duration of the growth experiment. The dashed red line indicates no difference in growth between dox treated and untreated cells. b) Migration assay (calculated as the difference between number of cells that migrated in dox 500 induced cells minus the number of cells that migrated in dox 0 cells). Xmal = stable cell line transfected with the *X. malinche* allele of *adgre5*. Xbir = stable cell line transfected with the *X. birchmanni* allele of *adgre5*. Asterisk means the difference is statistically significant. The plot shows the mean, and whiskers indicate two standard errors of the mean.

Transgenic medaka lines containing the *X. birchmanni* allele of *adgre5*, *tg(tyrp:X.bir_adgre5/mift:xmrk)*, showed significant reductions of pigmentation compared to *X. malinche adgre5* medaka lines, *tg(tyrp:X.mal_adgre5/mift:xmrk),* or *tg(mitf:xmrk)* medaka line (Figure 14, Figure 15. ANOVA p value < 0.0001, Tukey post hoc p values, *tg(tyrp:X.bir_adgre5/mift:xmrk)- tg(mitf:xmrk)* < 0.00001, *tg(tyrp:X.bir_adgre5/mift:xmrk)- tg(tyrp:X.mal_adgre5/mift:xmrk)* < 0.00001). *tg(tyrp:X.mal_adgre5/mift:xmrk)* and *tg(mitf:xmrk)* medaka lines did not show significant differences between them (Tukey post hoc p values, *tg(tyrp:X.mal_adgre5/mift:xmrk)- tg(mitf:xmrk)* = 0.85).

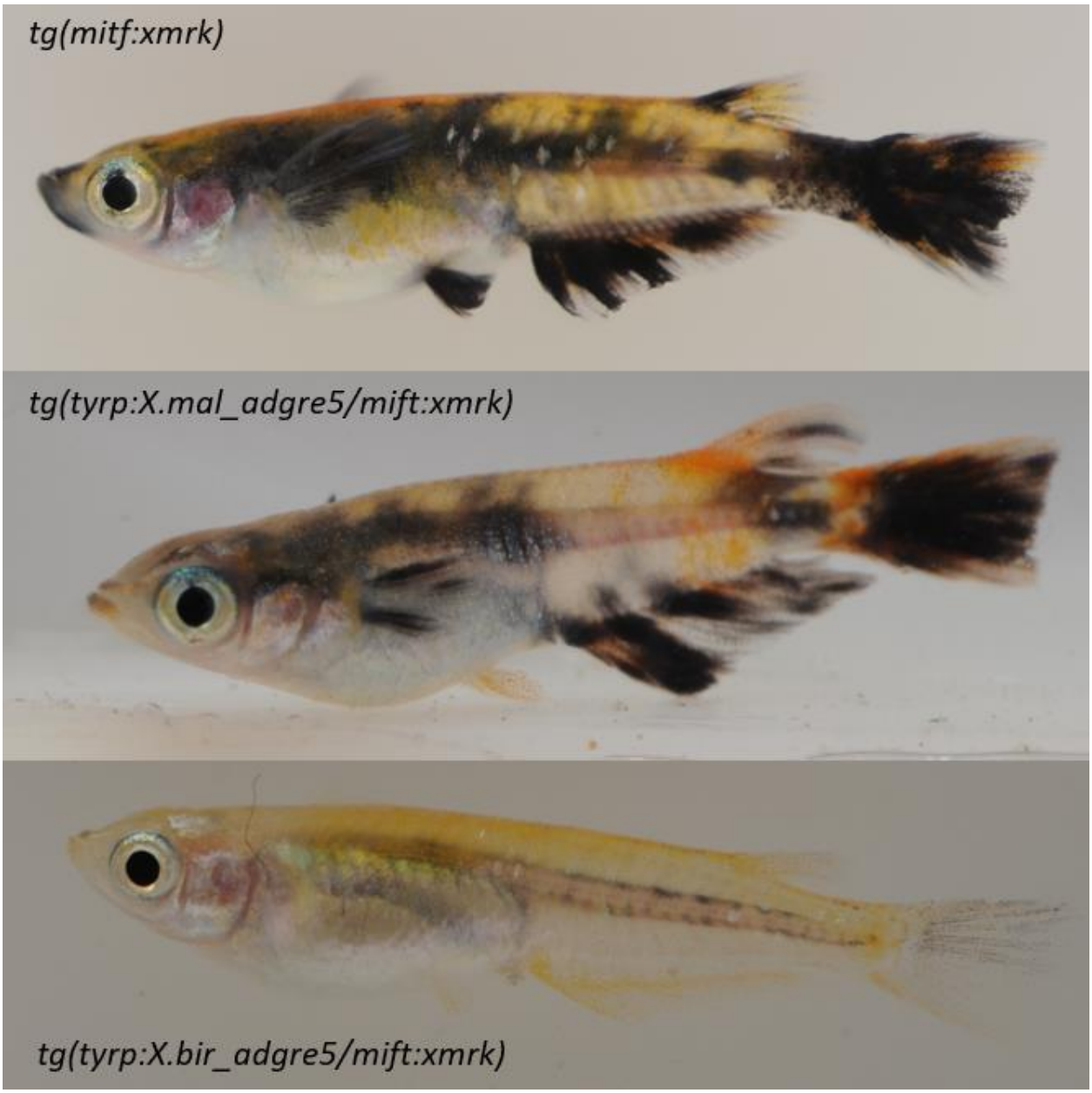


Figure 14. Close up photographs of the different transgenic lines. All individuals photographed of *tg(typr:Xmal_adgre5/mitf:xmrk)*, *tg(typr:Xbir_adgre5/mitf:xmrk)* are F1s, *tg(mitf:xmrk)* are more than 30 generations old.

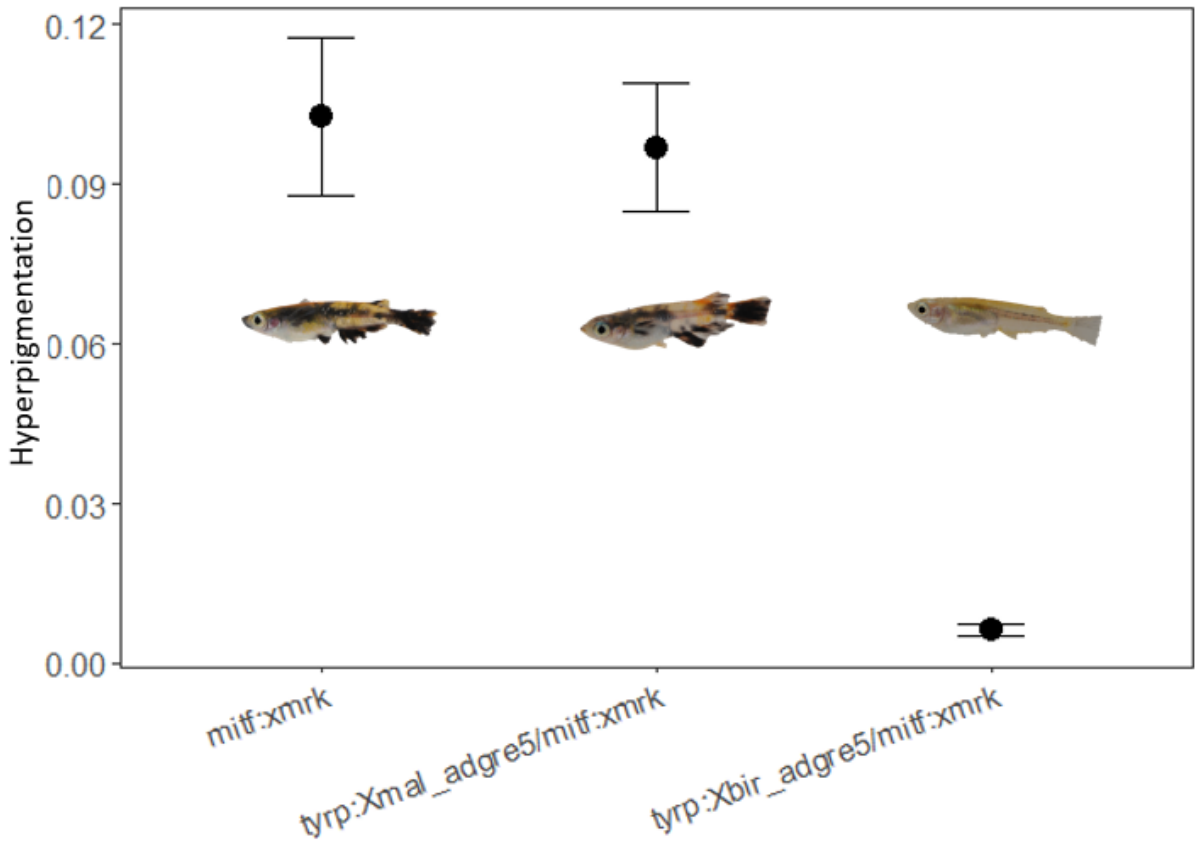


Figure 15. Hyperpigmentation area (corrected by standard length) for each transgenic line generated. Asterisk means the difference is statistically significant. The plot shows the mean, and whiskers indicate two standard errors of the mean.

Discussion

The lack of differences in pigmentation between the *tg(typr:Xmal_adgre5/mitf:xmrk)* and *tg(mitf:xmrk)*, as well as the striking reduction of pigmentation of *tg(typr:Xbir_adgre5/mitf:xmrk)*, are clear indicators that *adgre5* acts as a tumor suppressor. This agrees with the fact that pigment cells transfected with the *X. birchmanni* allele of *adgre5* caused cells to grow slower and migrate less than those transfected with the *X. malinche* allele of *adgre5*. *ADGRE5* is a prototypic member of the Adhesion class of G protein-coupled receptors (Adhesion GPCRs), which plays vital roles in numerous developmental processes as well as in tumorigenesis. Although it has been demonstrated that *adgre5*, under apoptotic conditions, can increase tumor cell viability by inhibiting of caspase activation and modulation of anti- and pro-apoptotic members of the BCL-2 superfamily (Hsiao et al., 2015), it has also been demonstrated that G protein-coupled receptors inhibit melanoma tumor growth and metastasis (Xu et al., 2006). In fact, a recent review (Langenhan, 2020) proposes Adhesion G protein–coupled receptors as candidate metabotropic mechanosensors and novel drug targets for numerous cancer types because they are surface molecules that act as mechanosensors. Moreover, another recent study (Regneri et al., 2019) also demonstrated that *xmrk* can be suppressed when co-expressed with a known potent tumor suppressor in medaka transgenic lines. Future studies to further characterize the role of *adgre5* as a tumor suppressor are much needed. For example, I would expect that a tissue specific knockout of medaka *adgre5* should cause even more malignant melanoma phenotypes (Regneri et al., 2019). Moreover, comparing these cell cultures results to studies performed in the PSM cell line should be particularly informative. The PSM cell line was established from melanoma tissue of

an adult F1 maculatus – hellerii (albino) hybrids and they highly overexpress *xmrk* (Wakamatsu, 1981).

Powell et al. 2020 had a big impact because it was the first study to propose a candidate single gene hybrid incompatibility in naturally occurring hybrids. Almost simultaneously, another important example of hybrid incompatibility single gene identification came from the closely related hybrids between X. maculatus and X. helleri. A recent study by Lu et al. 2020 (Lu et al., 2020) finally tackled the mystery that has been haunting researchers since *xmrk* was discovered: which is the tumor modifier gene that interacts with *xmrk* to cause hybrid melanoma. By a successive backcross of hybrids that lack tumors but show enhanced pigmentation of their dorsal fin (*xmrk*X. mac/− and heterozygous for R(diff)X. mac/X. hel) they successfully introgressed the R(diff) locus from X. maculatus into an X. helleri genetic background. Combining genome sequencing with association analyses, they identified a ∼100-kb interval containing three genes. RNA sequencing confirmed that only *rab3d* is expressed in the dorsal fin and they found an aminoacid change in the *RAB3D* protein between X. hellerii and X. maculatus *rab3d*. Future studies should focus on experimentally test the role of rabd3. Transgenic studies with the medaka *tg*(*mitf*:*xmrk*) line as well as cell culture experiments, should be particularly informative.

These mapping results are interesting because they propose a different gene (7 Mb apart from *adgre5*) interacting with *xmrk* as responsible hybrid incompatibilities in closely related pairs of hybrids. This suggests that a melanoma incompatibility involving *xmrk* originated independently in two distinct lineages. Although worded differently, this was hypothesized considerably before these studies were published. Schartl 2008 discussed that R (as it was called the then unknown tumor regulator gene) “could have preexisted before *xmrk* arose and would have sup-

pressed the melanoma from the moment when the oncogene arose". The most parsimonious hypothesis is that that *xmrk* originated in a common ancestor to all *Xiphophorus* and that it has been repeatedly lost in several branches of the tree. On the contrary, it appears as if all species have R, but different alleles, because different hybrid crosses have different levels of *xmrk* suppression (Anders, 1967; Atz, 1962). Now that R has been identified independently in two different *Xiphophorus* hybrids, we can say that R "different alleles" are actually different genes. This is not surprising if we consider that in both cases, this hybrid incompatibility appeared in similar situations: both X. helleri and *X. malinche* lack *xmrk* and therefore the interacting gene can evolve constraint free and independently in both lineages. As more incompatibilities are mapped to the single gene level, our understanding of whether there are unifying molecular or evolutionary forces driving them will increase.

The search for Bateson-Dobzhansky-Muller hybrid incompatibilities has seen a renewed interest since we realized hybridization is more common that we thought (Jhonson, 2002; Schumer et al., 2014). Yet, not many studies have been able to precisely pinpoint the interacting genes responsible for the incompatibility (Presgraves, 2010). Of those that have, most identify candidate genes by using hybrids between crosses of model species that no longer hybridize naturally. For example, using *Saccharomyces bayanus* - *Saccharomyces cerevisiae* hybrids, *ATPase expression 2* and *oligomycin resistance 1* were the genes identified to be responsible for a sporulation defect causing hybrid sterility in yeast (Lee et al., 2008). *Arabidopsis thaliana* laboratory hybrid crosses have also been used to determine *histidinol-phosphate amino-transferase* 1 and *histidinol-phosphate amino-transferase 2* as the genes responsible for embryo lethality (Bi-

kard et al., 2009) and *dangerous mix 1* as the gene activating hybrid necrosis triggered by an autoimmune response (Bomblies et al., 2007). *Drosophila* (Masly et al., 2006) and other less conventional species (Zuellig and Sweigart, 2018; Yu et al., 2018) also include this short list.

The list grows meaningfully shorter if considered which, from the studies able to pinpoint single gene effects, actually are able to experimentally test them and effectively assign a causal relationship between the gene interaction and the hybrid incompatibility. Among those few, there is two studies in *Drosophila* that generated transgenic lines to show that *hybrid male rescue* (*Hmr*) causes lethality and female sterility in hybrids among *D. melanogaster* and its sibling species (Barbash et al., 2003) and that *Lethal hybrid rescue* (*Lhr*) has functionally diverged in *Drosophila simulans* and interacts with *Hmr*, to cause lethality in F1 hybrid males (Brideauheather et al., 2006). To date, there is only one other study that experimentally tested a hybrid incompatibility in vertebrates. Mihola et al. 2009 generated six transgenic mouse lines to test, individually, gene by gene, each of the six genes associated with the *hybrid sterility 1* (*Hst1*) region (Trachtulec et al., 2008). With this Herculean effort, they were able to identify *Pr domaincontaining 9 (Prmd9)* as the gene responsible for the incompatibility, caused by spermatogenic failure via activation of genes essential for meiosis by methylation of histone H3 at lysine 4.

As far as I know, this is the only study that has experimentally validated a candidate gene involved in a hybrid incompatibility in species that currently hybridize naturally. Similar studies to this will shed important light to how incompatibilities affect speciation, more specifically, whether mapped incompatibilities were responsible for triggering divergence between species or they simply appeared after these lineages had stopped exchanging genes.

The results of this chapter demonstrate pretty clearly that *adgre5* acts as a tumor modifier and highlights the importance of following mapping studies with functional ones. Future research characterizing the specific mechanism by which *adgre5* suppresses *xmrk* promises to be of great biomedical significance. Only in 2021, 106110 new cases of melanoma and 7180 deaths occurred in the U.S. due to melanoma. Invasive melanoma accounts for about 1% of all skin cancer cases, but the vast majority of skin cancer deaths (American Cancer Society, 2021). Any insights on how a tumor suppressor gene might act to counteract an oncogene will prove to be of great value in developing cures for melanoma.

CHAPTER V

CONCLUSIONS

The study of color polymorphism is and has been important to understand how evolution operates in several contexts (Kronforst et al. 2012). Research traditionally focused on natural history, ecology, and behavior related to color and color polymorphisms but our understanding of the actual genetic material responsible for evolutionary change stills needs to catch up (Gray and McKinnon, 2007). Most of our knowledge comes from well-established animal models, but there are plenty of emerging systems. The genus *Xiphophorus* of livebearing fishes presents a striking variation in melanin-based pigmentation patterns, which are polymorphic within and between species (Basolo, 2006; Culumber, 2014). Moreover, *Xiphophorus* benefits from recently developed genomic resources (Schartl et al., 2013; Schumer et al., 2012, 2014, 2016) and has an intricated history of hybridization (Cui et al., 2013). Hybridization provides a great resource to pinpoint genes of interest, disentangle multivariate effects on fitness, and measure evolutionary change in real time.

The Bateson- Dobzhansky-Muller model of hybrid incompatibility proposes that new mutations arising in diverging species can interact negatively in hybrids, casing hybrids to have a reduced fitness in comparison to parental species (Orr, 1996). After realizing hybridization is more common than we thought (Jhonson, 2002) the search for hybrid incompatibilities has seen a renewed interest. And yet, only a few studies have been able to identify the hybrid incompatibilities

to the single-gene level (Presgraves, 2010). This study uses naturally hybridizing *Xiphophorus* species (*X. birchmanni* and *X. malinche*) to identify and test novel genes associated with melanoma and explores several hypotheses that attempt to explain how can melanoma persist in the wild.

Xiphophorus has been a dominant system in the study of melanoma since Gordon (Gordon, 1931) discovered that certain hybrids of the platyfish (*X. maculatus*) and swordtails (*X. hellerii*) develop a highly malignant melanoma (Meierjohann and Schartl, 2006). We know that in this hybrids melanoma is likely caused by a genetic imbalance between an oncogene and a tumor regulator gene. The hypothesis is that in *X. maculatus* there is an oncogene controlled by a tumor suppressor and that in *X. hellerii* these genes are both lacking. When both the oncogene and the tumor suppressor are present the phenotype results in benign pigmentation. However, in certain hybrids, because their genome is a combination of both parental species' genomes, it sometimes happens that the oncogene is expressed without the control of the tumor suppressor and therefore they develop malignant melanoma. *X. birchmanni* is polymorphic for a pigmentation spot on its caudal fin called spotted caudal. *X. malinche* lacks this pigmentation pattern. Hybrids between them are also polymorphic for the spot and can develop malignant melanoma from them. I looked into whether this melanoma between naturally occurring hybrids was also caused by a hybrid incompatibility.

Interestingly, the frequency of Sc and melanoma varies a lot between different hybrid populations. In Chapter II of my dissertation, I looked into the relationship between population ancestry structure and Sc frequency and malignancy. As expected in a hybrid incompatibility scenario, I found that ancestry admixed populations showed higher frequencies and malignancy

of Sc compared to ancestry structured population. Therefore, in order to identify the genes involved in this hybrid incompatibility I first performed a GWAS in *X. birchmanni*. The GWAS identified *xmrk*, a potent known oncogene as the gene driving the expression of Sc.

In Chapter III of my dissertation, I investigated how is it possible that a phenotype that lowers survivorship (Powell et al., 2020) is still present in the wild. Females did not show a preference for spotted individuals in visual choice trails in neither hybrids nor *X. birchmanni*. Future work should focus on characterizing if environmental factors, such as turbidity (Franck et al., 2001), or intrasexual differences associated with Sc, such as personality or morphological traits that covary with Sc, can help maintain the trait in the population.

Finally, an admixture mapping study that followed up the results of Chapter II of my dissertation proposed *adgre5* and *xmrk* as the genes responsible for the hybrid incompatibility causing melanoma (Powell et al., 2020). Therefore, in Chapter IV of my dissertation, I performed functional cell culture and transgenic experiments to demonstrate that *adgre5* acts as a tumor suppressor. As far as I know, this is the only study that combines behavior, genomics and molecular biology techniques in an integrative approach to identify and functionally test a hybrid incompatibility (melanoma) to the single gene level in naturally occurring hybridizing species.

REFERENCES

Abbott R., Albach D., Ansell S., Arntzen J, Baird S, Bierne N., Boughman J, Brelsford A., Buerkle C., Buggs R., Butlin R., Dieckmann U., Eroukhmanoff F., Grill A., Cahan S., Hermansen J., Hewitt G., Hudson A., Jiggins C., Jones J., Keller B., Marczewski T., Mallet J., Martinez-Rodriguez P., Most M., Mullen S., Nichols R., Nolte A., Parisod C., Pfennig K., Rice A., Ritchie M., Seifert B., Smadja C., Stelkens R., Szymura J., Vainola R., Wolf J and Zinner D. 2013. Hybridization and speciation. Journal of Evolutionary Biology, 26, 229-246.

Altschmied J., Hornung U., Schlupp I., Gadau J., Kolb R. and Schartl M. 1997. Isolation of DNA suitable for PCR for field and laboratory work. Biotechniques, 23, 228-229.

American Cancer Society 2021. Cancer Facts & Figures 2021.

Anders F. 1967. Tumor formation in Platyfish-Swordtail hybrids as a problem of gene regulation. Experientia, 23, 1-10.

Anders F. 1991. Contributions of the Gordon-Kosswig melanoma system to the present concept of neoplasia. Pigment Cell Res., 4, 7-29.

Andersson M. 1994. Sexual Selection. Princeton University Press, Princeton, NJ.

Ariyomo T. and Watt P. 2012. The effect of variation in boldness and aggressiveness on the reproductive success of zebrafish. Animal Behavior, 83, 41-46.

Ariyomo T. and Watt P. 2013. Disassortative mating for boldness decreases reproductive success in the guppy. Behavioral Ecology, 24, 1320-1326.

Atz J. 1962. Effects of hybridization on pigmentation in fishes of the genus *Xiphophorus*. Zoologica, 47, 153-181.

Avise J. 1983. Polymorphism of mitochondrial DNA in populations of higher animals. In Evolution of Genes and Proteins (ed. M. Nei & R. K. Koehn), 147-164.
Bagnara J and Hadley M. 1973. Chromatophores and Color Change: The Comparative Physiology of Animal Pigmentation. (Englewood Cliffs, NJ: Prentice-Hall, Inc.).

Barbash D., Siino D., Tarone A. and Roote J. 2003. A rapidly evolving MYB-related protein causes species isolation in *Drosophila*. Proceedings of the National Academy of Sciences of the USA, 100, 5302-5307.

Barton G. and Hewitt N. 1985. Analysis of hybrid zones. Annual Review of Ecology, Evolution, and Systematics, 16, 113-148.

Basolo A. 2006a. Genetic linkage and color polymorphism in the southern platyfish (*Xiphophorus maculatus*): a model system for studies of color pattern evolution. Zebrafish, 3, 65-83.

Bateson W. 1909. Heredity and variation in modern lights. Darwin and Modern Science, ed. AC Seward, pp. 85–101. Cambridge, UK: Cambridge Univ. Press.

Bateson W. and Gregory R. 1905. On the inheritance of heterostylism in Primula. Proceedings of the Royal Society B: Biological Sciences, 76, 581-586.

Béjar J., Hong Y. and Schartl M. 2003. *Mitf* expression is sufficient to direct differentiation of medaka blastula derived stem cells to melanocytes. Development 130, 6545-6553.

Bennett D., Cooper P. and Hart I. 1987. A line of non-tumorigenic mouse melanocytes, syngeneic with the B16 melanoma and requiring a tumor promoter for growth. International Journal of Cancer, 39, 414-418.

Bernstein F. 1925. Zusammenfassende Betrachtungen aus der Theorie der Blu*tg*ruppen. Zeitschrift fur induktiveAbstammungs- und Vererbungslehre, 37, 237-269.

Bierbach D., Sommer-Trembo C., Hanisch J., Wolf M. and Plath M. 2015. Personality affects mate choice: bolder males show stronger audience effects under high competition. Behavioral Ecology, 26, 1314-1325.

Bikard D., Patel D., Le-Metté C., Giorgi V., Camilleri C., Bennett M. and Loudet O. 2009. Divergent evolution of duplicate genes leads to genetic incompatibilities within *A. thaliana*. Science, 323, 623-626.

Black W., Baer C., Antolin M. and DuTeau N. 2001. Population genomics: genome-wide sampling of insect populations. Annual Review of Entomology, 41, 441-469.

Blanpain C. 2013. Tracing the cellular origin of cancer Nature Cell Biology, 15, 126-134.

Bomblies K., Lempe J., Epple P., Warthmann N., Lanz C., Dangl J. and Weigel D. 2007. Autoimmune response as a mechanism for a Dobzhansky–Muller-type incompatibility syndrome in plants. Plos Biology, 5, e236.

Breese M and Liu Y. 2013. NGSUtils: a software suite for analyzing and manipulating next-generation sequencing datasets. Bioinformatics, 29, 494-496.

Brelsford D. and Irwin A. 2009. Incipient speciation despite little assortative mating: the yellow-rumped warbler hybrid zone. Evolution: International Journal of Organic Evolution, 63, 3050-3060.

Brideauheather N., Flores H., Wang J., Maheshwari S., Wang X. and Barbash D. 2006. Two Dobzhansky-muller genes interact to cause hybrid lethality in *Drosophila*. Science, 314, 1292-1295.

Brisson J., Wilder J. and Hollocher H. 2006. Phylogenetic analysis of the Cardini group of *Drosophila* with respect to changes in pigmentation. Evolution, 60, 1228-1241.

Chan R., Stuart-Fox D. and Jessop T. 2009. Why are females ornamented? A test of the courtship stimulation and courtship rejection hypotheses Behavioral Ecology, 20, 1334–1342.

Chen G., Marjoram P. and Wall J. 2009. Fast and flexible simulation of DNA sequence data. Genome Research, 19, 136-142.

Clusella-Trullas S., Van-Wyk J. and Spotila J. 2007. Thermal melanism in ectotherms. Journal of Thermal Biology, 32, 235–245.

Coleman S., Patricelli G. and Borgia G. 2004. Variable female preferences drive complex male displays. Nature, 428, 742-745.

Corbett-Detig R. and Nielsen R.2017. A Hidden Markov Model approach for simultaneously estimating local ancestry and admixture time using next generation sequence data in samples of arbitrary ploidy. Plos Genetics, 13, e1006529.

Crespi B and Summers K. 2006. Positive selection in the evolution of cancer. Biological Reviews of the Cambridge Philosophical Society 81.

Cui R., Schumer M., Kruesi K., Walter R., Andolfatto P. and Rosenthal G. 2013. Phylogenomics reveals extensive reticulate evolution in *Xiphophorus* fishes. Evolution, 67, 2166-2179.

Culumber Z. 2014. Pigmentation in *Xiphophorus*: An Emerging System in Ecological and Evolutionary Genetics. Zebrafish, 11, 57-70.

Culumber Z. and Rosenthal G. 2013. Population-level mating patterns and fluctuating asymmetry in swordtail hybrids. Naturwissenschaften 100.

Culumber, Z. and Rosenthal, G. 2013. Mating preferences do not maintain the tailspot polymorphism in the platyfish, *Xiphophorus variatus.* Behavioral Ecology, 24.

Cummings M., Larkins-Ford J., Reilly C., Wong R., Ramsey M. and Hofmann H. 2008. Sexual and social stimuli elicit rapid and contrasting genomic responses. Proceedings of the Royal Society B: Biological Sciences, 275, 393-402.

Darwin C. 1859. On the origin of species by means of natural selection. London: J. Murray, 247.

Darwin C. 1868. The variation of animals and plants under domestication. 2 Vols. London: John Murray.

David F. and Cézilly M. 2011. Personality may confound common measures of mate-choice. Plos One, 6, e24778.

del-Marmol F. and Beermann V. 1996. Tyrosinase and related proteins in mammalian pigmentation. FEBS Letters, 381, 165-168.

Dobzhansky T. 1934. Studies on hybrid sterility. Z.Zellforsch 21, 169-223.

Dombeck J. and Jaenike I. 2004. Ecological genetics of abdominal pigmentation in *Drosophila falleni*: a pleiotropic link to nematode parasitism. Evolution, 58, 587-596.

Ducrest A., Keller L. and Roulin A. 2008. Pleitropy in the melanocortin system, coloration and behavioral syndromes. Trends in Ecology and Evolution, 23.

Eberle A. 1988. The Melanotropins: Chemistry, Physiology and Mechanism of Action. Karger.

Eichler W., Hamann J. and Aust G. 1997. Expression characteristics of the human *CD97* antigen. Tissue Antigens, 50, 429-438.

Fernandez A and Morris M. 2008a. Mate choice for more melanin as a mechanism to maintain a functional oncogene. Proceedings of the National Academy of Sciences of the USA, 105, 13503-13507.

Fernandez A. 2010. A cancer-causing gene is positively correlated with male aggression in *Xiphophorus cortezi*. Journal of Evolutionary Biology, 23, 386-396.

Fernandez A. and Bowser P. 2010. Selection for a dominant oncogene and large male size as a risk factor for melanoma in the *Xiphophorus* animal model. Molecular Ecology 19, 3114-3123.

Fisher H and Rosenthal G. 2006b. Male swordtails court with an audience in mind. Biology Letters, 3, 5-7.

Fisher H. and Rosenthal G. 2006a. Hungry females show stronger mating preferences. Behavioral Ecology, 17, 979-981.

Fisher H., Wong B. and Rosenthal G. 2006. Alteration of the chemical environment disrupts communication in a freshwater fish. Proceeding Biological Sciences, 273, 1187-1193.

Fisher H., Wong B. and Rosenthal G. 2006. Alteration of the chemical environment disrupts communication in a freshwater fish. Proceedings of the Royal Society B-Biological Sciences, 273, 1187-1193.

Fisher R. 1930. The Genetical Theory of Natural Selection. Oxford: Clarendon Press.

Franck D., Dikomey M. and Schartl M. 2001. Selection and the maintenance of a colour pattern polymorphism in the green swordtail (*Xiphophorus helleri*). Behavior 138, 467-486.

Frank S and Nowak M. 2004. Problems of somatic mutation and cancer. BioEssays, 26, 291-299.

Froschauer A., Körting C., Katagiri T., Aoki T., Asakawa S., Shimizu N., Schartl M. and Volff J. 2002. Construction and initial analysis of bacterial artificial chromosome (BAC) contigs from the sex-determining region of the platyfish *Xiphophorus* maculatus. Gene, 295, 247-254.

Gibert P., Moreteau B., Munjal A. and David J. 1999. Phenotypic plasticity of abdominal pigmentation in *Drosophila kikkawai*: multiple interactions between a major gene, sex, abdomen segment and growth temperature. Genetica, 105, 165-176.

Gimenez-Conti I., Woodhead A., Harshbarger J., Kazianis S., Setlow R., Nairn R. and Walter R. 2001. A Proposed Classification Scheme for *Xiphophorus* Melanomas Based on Histopathologic Analyses. Marine Biotechnology, 3, S100-S106.

Gomez A., Wellbrock C., Gutbrod H., Dimitrijevic N. and Schartl M. 2001. Ligand-independent dimerization and activation of the oncogenic *Xmrk* receptor by two mutations in the extracellular domain. Journal of Biological Chemistry, 276, 3333-3340.

Gordon M. 1931. Hereditary Basis of Melanosis in Hybrid Fishes. The American Journal of Cancer, 15, 1495-1523.

Graham J. 1992a. Cancer selection: the new theory of evolution. Lexington VA: Aculeus Press.

Gray S and McKinnon J. 2007. Linking color polymorphism maintenance and speciation. Trends in Ecology and Evolution, 22, 71-79.

Greaves M. 2000. Cancer: The Evolutionary Legacy. Oxford University Press, Oxford.

Haig, D. 1993. Genetic conflicts in human pregnancy. Quaterly review of Biology, 68, 495-532.

Haldane J. 1922. Sex ratio and unisexual sterility in animal hybrids. Journal of Genetics, 12, 101-109.

Harris H. 1966. Enzyme polymorphisms in man. Proceedings of the Royal Society B: Biological Sciences, 164, 298-310.

Hewitt G. 1998. Hybrid zones - natural laboratories for evolutionary studies. Trends in Ecology & Evolution, 3.

Hsiao C., Keysselt K., Chen H., Sittig D., Hamann J., Lin H. and Aust G. 2015. The Adhesion GPCR *CD97/ADGRE5* inhibits apoptosis. The International Journal of Biochemistry & Cell Biology, 65, 197-208.

Hubby J. and Lewontin R. 1966. A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. Genetics, 54, 577-594.

Ingley S. and Johnson S. 2014. Animal personality as a driver of reproductive isolation. Trends in Ecology and Evolution, 29, 369-371.

Jacobs M. 1985. Role of beta-alanine in cuticular tanning, sclerotization, and temperature regulation in *Drosophila melanogaster*. Journal of Insect Physiology, 31, 509-515.

Jeong S., Rebeiz M., Andolfatto P., Werner T., True J. and Carroll S. 2008. The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. Cell, 132, 783-793.

Jhonson N. 2002. Sixty years after "Isolating Mechanisms, Evolution and Temperature": Muller's legacy. Genetics, 161, 939-944.

Johnson J., Culumber Z., Easterling R. and Rosenthal G. 2015. Boldness and predator evasion in naturally hybridizing swordtails (Teleostei: *Xiphophorus*). Current Zoology, 61, 596-603.

Johnston S., Gratten J., Berenos C., Pilkington J., Clutton-Brock T., Pemberton J. and Slate J. 2013. Life history trade-offs at a single locus maintain sexually selected genetic variation. Nature, 502, 93-95.

Jones A. and Ratterman N. 2009. Mate choice and sexual selection: What have we learned since Darwin? Proceedings of the National Academy of Sciences of the USA, 106, 10001-1008.

Kan A. and Dozy Y. 1978. Polymorphism of DNA sequence adjacent to human beta-globin structural gene: relation to sickle mutation. Proceedings of the National Academy of Sciences of the USA, 75, 5631-5635.

Kimura M. 1983a. The Neutral Theory of Molecular Evolution. Cambridge: Cambridge University Press.

Kirchen R. and West W. 1976. The Japanese Medaka: its Care and Development. Carolina Biological Supply Company, 36.

Kirkpatrick M. and Ryan M. 1991. The evolution of mating preferences and the paradox of the lek. Nature, 350, 33-38.

Kleene K. 2005. Sexual selection, genetic conflict, selfish genes and the atypical patterns of gene expression in spermatogenic cells. Development Biology, 277, 16-26.

Klein R., Zeiss C., Chew E., Tsai J., Sackler R., Haynes C., Henning A., SanGiovanni J., Mane S., Mayne S., Bracken M., Ferris F., Ott J., Barnstable C. and Hoh J. 2005. Complement Factor H Polymorphism in Age-Related Macular Degeneration. Science, 308.
Koopp A. 2009. Metamodels and phylogenetic replication: a systematic approach to the evolution of developmental pathways. Evolution, 63, 2771–2789.

Korte A. and Farlow A. 2013. The advantages and limitations of trait analysis with GWAS: a review. Plant Methods, 9.

Kouprina N., Mullokandov M., Rogozin I., Collins N., Solomon G., Otstot J., Risinger J.,

Koonin E., Barrett J. and Larionov V. 2004. The SPANX gene family of cancer/testis-specific antigens: rapid evolution and amplification in African great apes and hominids. Proceedings of the National Academy of Sciences of the USA, 101, 3077-3082.

Kreitman M. 1983b. Nucleotide polymorphism at the Alcohol dehydrogenase locus of *Drosophila melanogaster*. Nature 304, 412-417.

Kronforst M., Barsh G., Kopp A., Mallet J., Monteiro A., Mullen S., Protas M., Rosenblum E., Schneider C. and Hoekstra H. 2012. Unraveling the thread of nature's tapestry: the genetics of diversity and convergence in animal pigmentation. Pigment Cell & Melanoma, 25, 411-433.

Lande R. 1981. Models of speciation by sexual selection on polygenic traits. Proceedings of the National Academy of Sciences of the USA, 78, 3721-3725.

Langenhan T. 2020. Adhesion G protein–coupled receptors—Candidate metabotropic mechanosensors and novel drug targets. Basic & Clinical Pharmacology & Toxicology, 126, 5-16.

Langley C. and Aquadro C. 1987. Restriction map variation in natural populations of *Drosophila melanogaster*: white locus region Molecular Biology and Evolution, 4, 651-663.

Lee H., Chou J., Cheong L., Chang N., Yang S. and Leu J. 2008. Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species. Cell, 135, 1065-1073.

Lewontin R. 1974. The Genetic Basis of Evolutionary Change. New York, NY: Columbia University Press.

Lewontin R. 1985. Population genetics. Annual Review of Genetics, 19, 81-102.

Lexer C., Joseph J., van-Loo M., Barbara T., Heinze B., Bartha D., Castiglione S., Fay M. and Buerkle C. 2010. Genomic Admixture Analysis in European Populus spp. Reveals Unexpected Patterns of Reproductive Isolation and Mating. Genetics, 186, 699-U391.

Li H and Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 25.

Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics, 27, 2987-2993.

Liang B., Ding H., Huang L., Luo H. and Zhu X. 2020. GWAS in cancer: progress and challenges. Molecular Genetics and Genomics, 295, 537-561.

Lindsley D., Grell E. and Bridges C. 1968. Genetic Variations of *Drosophila melanogaster*. (Washington, DC: CarnegieInst.).

Lu Y., Sandoval A., Voss S., Lai Z., Kneitz S., Boswell W., Boswell M., Savage M., Walter C., Warren W., Schartl M. and Walter R. 2020. Oncogenic allelic interaction in *Xiphophorus* highlights hybrid incompatibility. Proceedings of the National Academy of Sciences of the USA, 47, 29786-29794.

Luikart G., England P., Tallmon D., Jordan S. and Taberlet P. 2003. The power and promise of population genomics: from genotyping to genome typing. Nature Reviews Genetics, 4, 981-994.

Lynch A. and Force M. 2000. The origin of interspecific genomic incompatibility via gene duplication. The American Naturalist, 156, 590-605.

MacArthur J., Bowler E., Cerezo M., Gil L., Hall P., Hastings E., Junkins H., McMahon A., Milano A., Morales J., Pendlington Z., Welter D., Burdett T., Hindorff L., Flicek P., Cunningham F. and Parkinson H. 2017. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleid Acids Research, 45, 896-901.

Masly J., Jones C., Noor M., Locke J. and Orr H. 2006. Gene transposition as a cause of hybrid sterility in *Drosophila*. Science, 313, 1448-1450.

Meierjohann S and Schartl M. 2006. From Mendelian to molecular genetics: the *Xiphophorus* melanoma model. Trends in Genetics, 22, 654-661.

Meyer A and Van-de-Peer Y. 2005. From 2R to 3R: Evidence for a fish specific genome duplication (FSGD). BioEssays, 27, 937-945.

Mihola O., Trachtulec Z., Vlcek C., Schimenti J. and Forejt J. 2009. A Mouse Speciation Gene Encodes a Meiotic Histone H3 Methyltransferase. Science, 323, 373-375.

Muller H. 1928. The measurement of mutation rate in *Drosophila*, its high variability and its dependence on temperature. Genetics, 13, 279-357.

Nei M. and Li W. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Sciences of the USA, 76, 5269-5276.

Nosil C. and Sandoval P. 2008. Ecological niche dimensionality and the evolutionary diversification of stick insects. Plos One, 3.

Nosil P., Harmon L. and Seehausen O. 2009. Ecological explanations for (incomplete) speciation. Trends in Ecology and Evolution, 24, 145-156.

Orr H. 1997. Haldane's Rules. Annual Review of Ecology, Evolution, and Systematics, 28, 195-218.

Orr H. 2010. The population genetics of beneficial mutations. Philosophical Transactions of the Royal Society B: Biological Sciences, 365, 1195–1201.

Orr H.1996. Dobzhansky, Bateson, and the genetics of speciation. Genetics, 144, 1331-1335.

Otto W.1956. On the origin of cancer cells. Science, 123, 309-314.

Palopoli C. and Wu M. 1994. Genetics of hybrid male-sterility between *Drosophila* sibling species: a complex web of epistasis is revealed in interspecific studies. Genetics, 138, 329-341.

Powell D. and Rosenthal G. 2017a. What artifice can and cannot tell us about animal behavior. Current Zoology, 63, 21-26.

Powell D., García-Olazábal M., Keegan M., Reilly P., Du K., Díaz-Loyo A., Banerjee S., Blakkan D., Reich D., Andolfatto P., Rosenthal G., Schartl M. and Schumer M. 2020. Natural hybridization reveals incompatible alleles that cause melanoma in swordtail fish. Science, 368, 731-736.

Powell D., Payne C., Banerjee S., Keegan M., Bashkirova E., Cui R., Andolfatto P., Rosenthal G. and Schumer M. 2021. The genetic architecture of variation in the sexually selected sword ornament and its evolution in hybrid populations. Current Biology, 31, 923-935.

Presgraves D. 2010. The molecular evolutionary basis of species formation. Nature Review Genetics, 11, 175-180.

Price A., Patterson N., Plenge R., Weinblatt M., Shadick N. and Reich D. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics, 38, 904-909.

Punnet R. 1915. Mimicry in Butterflies. Cambridge: Cambridge University Press.
Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M., Bender M., Maller J., Sklar P., de-Bakker P., Daly M. and Sham P. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics, 81, 559-575.

Ramalho J., Lopes V., Tarafder A., Seabra M. and Hume A. 2009. Myrip uses distinct domains in the cellular activation of myosin VA and myosin VIIA in melanosome transport. Pigment Cell & Melanoma 22, 461-476.

Rambaut A. and Grassly N. 1997. Seq-Gen: An application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic frees. Computer Applications in the Biosciences 13, 235-238.

Reale D., Dingemanse N., Kazem A. and Wright J. 2010. Evolutionary and ecological approaches to the study of personality. Philosophical Transactions of the Royal Society B: Biological Sciences, 365, 3937-3946.

Regneri J., Klotz B., Wilde B., Kottler V., Hausmann M., Kneitz S., Regensburger M., Maurus K., Götz R., Lu Y., Walter R., Herpin A. and Schartl M. 2019. Analysis of the putative tumor suppressor gene cdkn2ab in pigment cells and melanoma of *Xiphophorus* and medaka. Pigment Cell & Melanoma, 32, 248-258.

Reynolds J and Gross M. 1990a. Costs and benefits of female mate choice: Is there a lek paradox? The American Naturalist, 136, 230-243.

Rosemblum E and Harmon L. 2011. ‘‘Same but different’’: replicated ecological speciation at White Sands Evolution, 65, 946-960.

Rosemblum E. 2006. Convergent evolution and divergent selection: lizards at the White Sands ecotone. The American Naturalist, 167, 1-15.

Rosenblum E., Hoekstra H. and Nachman M. 2004. Adaptive reptile color variation and the evolution of the Mc1r gene. Evolution, 1794–1808.

Rosenblum E., Rompler H., Schoneberg T. and Hoekstra H. 2010. Molecular and functional basis of phenotypic convergence in white lizards at White Sands. PNAS, 107, 2113–2117.

Rosenthal G. 2017b. Mate Choice: The Evolution of Sexual Decision Making from Microbes to Humans. Princeton University Press, 647.

Rosenthal G., de-la-Rosa X., Kazianis S., Stephens M., Morizot D., Ryan M. and García-de-León F. 2003. Dissolution of sexual signal complexes in a hybrid zone between the swordtails *Xiphophorus birchmanni* and *Xiphophorus malinche* (Poeciliidae). Copeia, 2, 299-307.

Rossato F., Zecchin K., La-Guardia P., Ortega R., Alberici L., Costa R., Catharino R., Graner E., Castilho R. and Vercesi A. 2014. Fatty acid synthase inhibitors induce apoptosis in non-tumorigenic melan-a cells associated with inhibition of mitochondrial respiration. Plos One, 9, e101060.

Ryan M and Cummings M. 2013. Perceptual Biases and Mate Choice. Annual Review of Ecology, Evolution, and Systematics, 44, 437-459.

Ryan M and Keddy-Hector A. 1992. Directional patterns of female mate choice and the role of sensory biases. American Society of Naturalists, 139, 4-35.

Ryan M. 1990b. Sensory systems, sexual selection, and sensory exploitation Oxford Surveys in Evolutionary Biology, 17, 157-195.

Ryan M. 1990c. Signals, species, and sexual selection. Oxford Surveys in Evolutionary Biology, 7, 157-195.

Safaee M., Clark A., Ivan M., Oh M., Bloch O., Sun M., Oh T. and Parsa A. 2013. *CD97* is a multifunctional leukocyte receptor with distinct roles in human cancers. International Journal of Oncology, 43, 1343-1350.

Safran R., Scordato E., Symes L., Rodríguez R. and Mendelson T. 2013. Contributions of natural and sexual selection to the evolution of premating reproductive isolation: a research agenda. Trends in Ecology and Evolution, 28, 643-650.

Savage J. and Slowinski J. 1992b. The coloration of the venomous coral snakes (family Elapidae) and their mimics (families Aniliidae and Colubridae). Biological Journal of the Linneal Society, 45, 235–254.

Schartl M. 2008b. Evolution of *Xmrk*: an oncogene, but also a speciation gene? BioEssays, 30, 822-832.

Schartl M., Hornung U., Gutbrod H., Volff J. and Wittbrodt J. 1999. Melanoma loss-of-function mutants in *Xiphophorus* caused by *Xmrk*- oncogene deletion and gene disruption by a transposable element. Genetics, 153, 1385-1394.

Schartl M., Kneitz S Ormanns J., Schmidt C., Anderson J., Amores A., Catchen J., Wilson C., Geiger D., Du K., Garcia-Olazábal M., Sudaram S., Winkler C., Hedrich R., Warren W., Walter R., Meyer A. and Postlethwait J. 2021. The Developmental and Genetic Architecture of the Sexually Selected Male Ornament of Swordtails. Current Biology, 31, 911-922.

Schartl M., Ronald W., Shen Y., Garcia T., Catchen J., Amores A., Braasch I., Chalopin D., Volff J., Lesch K., Bisazza A., Minx P., Hillier L., Wilson R., Fuerstenberg S., Boore J., Searle S., Postlethwait J. and Warren W. 2013. The genome of the platyfish, *Xiphophorus maculatus*, provides insights into evolutionary adaptation and several complex traits. Nature Genetics, 45, 567-572.

Schartl M., Wilde B., Laisney J., Taniguchi Y., Takeda S. and Meierjohann S. 2010. A mutated egfr is sufficient to induce malignant melanoma with genetic background-dependent histopathologies. Journal of Investigative Dermatology, 130, 249-258.

Schneider C., Rasband W. and Eliceiri K. 2012. NIH Image to ImageJ: 25 years of Image Analysis. Nature Methods, 9, 671-675.

Schuett W., Tregenza T. and Dall S. 2010. Sexual selection and animal personality. Biological Reviews, 85, 217-246.

Schumer M., Cui R., Boussau B., Walter R., Rosenthal G. and Andolfatto P. 2012. An evaluation of the hybrid speciation hypothesis for *Xiphophorus clemenciae* based on whole genome sequences. Evolution, 67, 1155–1168.

Schumer M., Cui R., Powell D., Dresner R., Rosenthal G. and Andolfatto P. 2014a. High-resolution mapping reveals hundreds of genetic incompatibilities in hybridizing fish species. Elife, e02535.

Schumer M., Cui R., Rosenthal G. and Andolfatto P. 2016. simMSG: an experimental design tool for high-throughput genotyping of hybrids. Molecular Ecology Resources, 16, 183-192.

Schumer M., Powell D., Delclós P., Squire M., Cui R., Andolfatto P. and Rosenthal G. 2017. Assortative mating and persistent reproductive isolation in hybrids. Proceedings of the National Academy of Sciences of the USA, 114, 10936–10941.

Schumer M., Rosenthal G. and Andolfatto P. 2014b. How common is homoploid hybrid speciation? Evolution, 68, 1553-1560.

Schumer M., Xu C., Powell D., Durvasula A., Skov L., Holland C., Blazier J., Sankararaman S., Andolfatto P., Rosenthal G. and Przeworski M. 2018. Natural selection interacts with recombination to shape the evolution of hybrid genomes Science, 360, 656.

Sedarti A., Malitschek B., Kazianis S., Borowsky R. and Schartl M. 1995. Spontaneous Melanoma Formation in Nonhybrid *Xiphophorus*. Cancer Research, 55, 159-165.

Sommer-Trembo C., Bierbach D., Arias-Rodriguez L., Verel Y., Jourdan J., Zimmer C., Riesch R., Streit B. and Plath M. 2016. Does personality affect premating isolation between locally-adapted populations? BMC Evolutionary Biology, 16, 138.

Stamatakis A. 2006b. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics, 22, 2688-2690.

Steiner C., Rompler H., Boet*tg*er L., Schoneberg T. and Hoekstra H. 2009. The genetic basis of phenotypic convergence in beach mice: similar pigment patterns but different genes. Molecular Biology and Evolution, 25, 35-45.

Summers B., and Crespi K. 2008. The androgen receptor and prostate cancer: a role for sexual selection and sexual conflict? Medical Hypotheses, 70, 435-443.
Tam V., Patel N., Turcotte M., Bossé Y., Paré G. and Meyre D. 2019. Benefits and limitations of genome wide association studies. Nature Review Genetics, 20, 467-484.

Thermes V., Grabher C., Ristoratore F., Bourrat F., Choulika A., Wittbrodt J. and Joly J. 2002. I-SceI meganuclease mediates highly efficient transgenesis in fish. Mechanisms of Development, 118, 91-98.

Trachtulec Z., Vlcek C., Mihola O., Gregorova S., Fotopulosova V. and Forejt J. 2008. Fine haplotype structure of a chromosome 17 region in the laboratory and wild mouse. Genetics, 178, 1777-1784.

True J. 2003. Insect melanism: the molecules matter. Trends in Ecology and Evolution, 18, 640–647.

True J., Yeh S., Hovemann B., Kemme T., Meinertzhagen I., Edwards T., Liou S., Han Q. and Li J. 2005. *Drosophila* tan encodes a novel hydrolase required in pigmentation and vision. Plos Genetics, 1.

Uffelmann E., Huang Q., Munung N., de-Vries J., Okada Y., Martin A., Martin H., Lappalainen T. and Posthuma D. 2021. Genome-wide association studies. Nature Review Methods Primers, 1.
Veen T., Ingley S., Cui R., Simpson J., Asl M., Zhang J., Butkowski T., Li W., Hash C., Johnson J., Yan W. and Rosenthal G. 2013. anyFish: an open-source software to generate animated fish models for behavioral studies. Evolutionary Ecology Research, 15, 361-375.

Verzijden M., Culumber Z. and Rosenthal G. 2012. Opposite effects of learning cause asymmetric mate preferences in hybridizing species. Behavioral Ecology, 23, 1133-1139.

Volff J. and Schartl M. 2003. Evolution of signal transduction by gene and genome duplication in fish. Journal of Structural and Functional Genomics, 3, 139-150.

Walling C., Royle N., Lindström J. and Metcalfe N. 2010. Do female association preferences predict the likelihood of reproduction? Behavioral Ecology and Sociobiology, 64, 541-548.

Wakamatsu Y. 1981. Establishment of a cell line from the platyfish-swordtail hybrid melanoma. Cancer Research, 41, 679-680.

Weis S. and Schartl M. 1998. The macromelanophore locus and the melanoma oncogene *Xmrk* are separate genetic entities in the genome of *Xiphophorus*. Genetics, 149, 1909-1920.

Westcott D. 1994. Leks of leks: A role for hotspots in lek evolution? Proceedings of the Royal Society B: Biological Sciences, 258, 281-286.

Winberry G. and Shefner L. 2009. Intraspecific polymorphism to interspecific divergence: genetics of pigmentation in *Drosophila*. Science, 326, 540-544.

Winkler C., Wittbrodt J., Lammers R., Ullrich A. and Schartl M. 1994. Ligand dependent tumor induction in medaka fish embryos by a *Xmrk* receptor tyrosine kinase transgene. Oncogene, 9, 1517-1525.

Wittbrodt J., Adam D., Malitschek B., Mäueler W., Raulf F., Telling A., Robertson S. and Schartl M. 1989. Novel putative receptor tyrosine kinase encoded by the melanoma-inducing Tu locus in *Xiphophorus*. Nature, 341, 415-421.

Wittbrodt J., Lammers R., Malitschek B., Ullrich A. and Schartl M. 1992. The *Xmrk* receptor tyrosine kinase is activated in *Xiphophorus* malignant melanoma. EMBO J, 11, 4239-4246.

Wittbrodt J., Meyer A. and Schartl M. 1998. More genes in fish? BioEssays, 20, 511-515.

Wittbrodt J., Shima A. and Schartl M. 2002. Medaka—a model organism from the far East. Nature Review Genetics, 3, 53-64.

Wittkopp P., Smith-Winberry G., Arnold L., Thompson E., Cooley A., Yuan D., Song Q. and Mcallister B. 2011. Local adaptation for body color in *Drosophila* americana. Heredity, 106, 592-602.

Wittkopp P., Stewart E., Arnold L., Neidert A., Haerum B., Thompson E., Akhras S., Smith-

Wolf F. and Weissing M.2012. Animal personalities: consequences for ecology and evolution. Trends in Ecology and Evolution, 27, 452-461.

Wong B and Rosenthal G. 2006c. Female disdain for swords in a swordtail fish. The American Naturalist, 167, 136-140.

Wong R., So P. and Cummings M. 2011. How female size and male displays influence mate preference in a swordtail. Animal Behavior, 82, 691-697.

Wright S. 1922. The Effects of Inbreeding and Crossbreeding on Guinea pigs: I. Decline in Vigour. Bullletin 1090. Washington, DC: U. S. Department of Agriculture.

Wright T. 1987. The genetics of biogenic amine metabolism, sclerotization, and melanization in *Drosophila melanogaster*. Advanced Genetics, 24, 127-222.

Wu, C. 2001. The genic view of the process of speciation. Journal of Ecology and Evolution, 14, 851-865.

Xu L., Begum S., Hearn J. and Hynes R. 2006. GPR56, an atypical G protein-coupled receptor, binds tissue transglutaminase, *TG*2, and inhibits melanoma tumor growth and metastasis. Proceedings of the National Academy of Sciences of the USA, 103, 9023-9028.

Yona S., Lin H., Siu W., Gordon S. and Stacey M. 2008. Adhesion-GPCRs: emerging roles for novel receptors. Trends in Biochemical Sciences, 33, 491-500.

Yu X., Zhao Z., Zheng X., Zhou J., Kong W., Wang P., Bai W., Zheng H., Zhang H., Li J., Liu J., Wang Q., Zhang L., Liu K., Yu Y., Guo X., Wang J., Lin Q., Wu F., Ren Y., Zhu S., Zhang X., Cheng Z., Lei C., Liu S., Liu X., Tian Y., Jiang L., Ge S., Wu C., Tao D., Wang H. and Wan J. 2018. A selfish genetic element confers non-Mendelian inheritance in rice. 360, 1130-1132.

Zou J., Beermann F., Wang J., Kawakami K. and Wei X. 2006. The *Fugu tyrp*1 promoter directs specific GFP expression in zebrafish: tools to study the RPE and the neural crest-derived melanophores. Pigment Cell Research, 19, 615-627.

Zuellig A. and Sweigart M. 2018. Gene duplicates cause hybrid lethality between sympatric species of Mimulus. Plos Genetics, 14, e1007130.