MIRCOEVOLUTION IN NATURAL AND EXPERIMENTAL HYBRID POPULATIONS OF SWORDTAIL FISH

A Dissertation

by

RICHARD STEPHEN BOVIO

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 Rosenthal
Co-Chair of Committee,	William Murphy
Committee Members,	Charles Criscione
	Kira Delmore
Interdisciplinary Faculty Chair,	Mickey Eubanks

August 2022

Major Subject: Ecology & Evolutionary Biology

ABSTRACT

Hybridization is a common phenomenon that serves as an important evolutionary mechanism by which diversity can arise. When two genetically divergent species hybridize, the resulting admixture generates novel genotypic and phenotypic combinations. Depending on a host of demographic, genetic, and environmental conditions, hybrid zone evolution can have various evolutionary outcomes. Natural hybrid zones between two species can offer a unique opportunity to study how hybridization impacts adaptive evolution in admixed populations due to their natural history and ecological circumstances. In my dissertation, I use two freshwater species of fish, *Xiphophorus birchmanni* and *X. malinche*, that form replicated natural hybrid zones to address how hybrid populations evolve and to identify barriers that contribute to maintaining isolation between species.

To uncover how morphology evolves in hybrid populations, I measured sexually selected male traits in three independent natural hybrid populations as well as three experimental, early generation hybrid populations located at different elevational treatments for approximately 5 years. Principal component analysis revealed the two natural hybrid populations located at lower elevations to share similar morphology with the lowland parental species *X. birchmanni*, and that the third natural hybrid population located at higher elevations shared more similarity with the highland parental species, *X. malinche*. Overall, I found in the lowland natural hybrid populations, environment drives morphology to resemble the lowland parental species. In the highland natural hybrid population, morphology aligns with genome-wide hybrid index. These findings reveal independent hybrid populations can exhibit similar morphometric combinations of traits despite genome-wide ancestry composition, but that it is context dependent. In other cases, morphology aligns with the genome-wide ancestry. Most notably, within the experimental hybrid populations, I found more variation was shared between sites than within sites. Across each experimental hybrid site, the amount of variation in morphology among replicates was relatively the same. Additionally, I make use of whole genome-wide sequence data for a subset of the individuals in the study and found the average hybrid index at each site trended in the direction that would be expected if

environmental selection was driving genome-wide changes in ancestry.

When species hybridize along climate gradients, selection should favor the introgression of alleles that mitigate the fitness cost of environmental change. Quantifying the introgression of loci associated with fitness in hybrid zones is critical to predicting how hybridization aligns with responses to climate change. To this end, we performed a QTL analysis on intercrossed experimental, early generation hybrids with the aim to identify genomic regions associated with interspecific differences in thermal tolerance and found an underdominant QTL located on chromosome 22. However, ancestry at this QTL did not significantly change over time in natural hybrid populations nor was it considered a candidate for selection after performing geographic cline analysis among two independent hybrid zone clines. Interestingly, however, I found several *X. malinche* heat-shock protein alleles to be introgressing downstream in one of the river drainages.

Finally, despite abundant work on postmating-prezygotic sexual selection within the same genus of fishes, little is known about how sperm competition and cryptic female choice may mediate gene flow between species. This is particularly important as premating isolating mechanisms may be weak and are frequently susceptible to disruption of communication channels. I investigate the postmating-prezygotic barriers to reproduction and found females inseminated with equal proportions of conspecific and heterospecific sperm to bias fertilization towards the conspecific male sperm. These findings suggest postmating-prezygotic process has the potential to mediate the cost of hybridization.

DEDICATION

To the love of my life, Sarah Bovio.

ACKNOWLEDGMENTS

I want to acknowledge my graduate advisor, mentor, and friend Dr. Gil Rosenthal for his support and guidance throughout this journey. Gil's undying enthusiasm for science and dedication to his students made my time in the lab deeply rewarding. Thank you to my committee members Dr. William Murphy, Dr. Kira Delmore, and Dr. Charles Criscione for your technical guidance and challenging me each step of the way - I am a better scientist today because of it.

Many individuals of the Rosenthal Lab, past and present, deserve recognition for their assistance and camaraderie throughout this pursuit. These include Angie Achorn, Irene Andermarcher, Keisuke Atsumi, Amanda Beckman, Radim Blazek, Max Chin, Pablo Delclos, Owen Dorsey, Megan Exnicios, Paola Fascinetto Zago, Mateo Garcia, Faith Hardin, Christopher Holland, Gaston Jofre, Will Ledbetter, Emma Lehmberg, Elizabeth Lewis, Rebecca Mangold, Annais Muschett-Bonilla, Annabel Perry, Dan Powell, Keerthana Rameshbabu, Giorgia Rando, Breann Richey, Aaron Rose, and Jeffery White. I want to give recognition to Heath Blackman, Adam Jones, Jessica Light, Alan Pepper, Manfred Schartl, and Molly Schumer. Each of these faculty members played a direct part towards my growth as a scientific thinker. Thank you for being incredible educators. There are many other students, collaborators, and friends that contributed to my success. An incomplete list includes Jamie Alfieri, Shreya Banerjee, Luke Bower, Alyson Brokaw, Nicole Foley, Janelle Goeke, Zach Hancock, Samantha Levell, Cheyenne Payne, Whitney Preisser, David Saenz, and Baruc Zago Mazzocco. Thanks to each and every one of you for everything you offered.

The EEB program coordinators Jason Martina, Nick Jacobsen, and Jennifer Bradford assisted in many ways behind the scenes. Thank you for handling many of those duties. An entire chapter of my dissertation would have not been possible without the services provided by Dickson Varner, Rosanna Serafini, Charles Love, and Sheila Teague. These individuals trained and allowed me to use highly specialized equipment that was unavailable in my own lab. I would like to thank the Calnali community for always welcoming me and being supportive in my research. Most

notably, the Hernandez Perez family provided essential help by maintaining CICHAZ and assisting in fieldwork. Thank you to Gaby Vazquez who organized and coordinated our stay at the field station.

Thank you to my family. Specifically my parents, Sylvia Gutierrez and Ernie Bovio, who not only encouraged me throughout this arduous process, but have loved and supported me my entire life. The love provided by Rebecca and Anil Shetty, Holly, John and Lilly Bovio, and Gary Carrington was received with gratitude and appreciation. Finally, I am most thankful to the love of my life, Sarah Bovio. Sarah is a constant inspiration in my life. I am so fortunate that we have the opportunity to share our lives together. "You and me together could do anything, baby." -Dave

I would like to thank the Texas A&M University Graduate and Professional School to allow me to construct this LaTeX thesis template.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Dr. Gil Rosenthal (advisor) of the Department of Biology, Dr. William Murphy (co-advisor) of the Department of Veterinary Integrative Biosciences, and Dr. Kira Delmore and Dr. Charles Criscione of the Department of Biology.

The field work conducted in Chapter II and III was made possible by a team of graduate and undergraduate students affiliated with the Department of Biology. The morphometric data was measured by Max Chin, Keerthana Rameshbabu, William Ledbetter, Rebecca Mangold, and Irene Andermarcher of the Department of Biology.

In Chapter IV, the geographic cline data was collected by Dr. Gaston Jofre of the Department of Biology at The University of North Carolina at Chapel Hill, and the quantitative trait locus analyses were conducted in part by Cheyenne Payne of the Department of Biology at Stanford University and were published in 2022 in an article listed in Molecular Ecology.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a Graduate Research Fellowship from the National Science Foundation (NSF), a Long Term Research in Environmental Biology Grant from the NSF, and a Student Research Grant from the Animal Behavior Society.

NOMENCLATURE

ACUA	Acuapa natural hybrid zone
TLMC	Tlatemaco natural hybrid zone
AGZC	Aguazarca natural hybrid zone
HIGH	Highland experimental, early generation hybrid site
MID	Intermediate experimental, early generation hybrid site
LOW	Lowland experimental, early generation hybrid site
MAL	*Xiphophorus malinche*
BIR	*X. birchmanni*
MC	AGZC *X. malinche*-skewed cluster
BC	AGZC *X. birchmanni*-skewed cluster
QTL	Quantitative trait locus
PCA	Principal compoenent analysis
CTmax	Critical thermal maxima
AIC	Akaike information criterion
NCBI	National Center for Biotechnology Information
FDR	False discover rate
CSP	Conspecific sperm precedence

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION

Historically, hybridization was underappreciated as an important evolutionary process across the tree of life. We now recognize hybridization is a ubiquitous process in nature. In hybridizing species, gene flow can lead to novel genetic and phenotypic variation for selection to act upon [1]. This mechanism enables genomic regions to introgress from one species into another, either as a neutral or adaptive process [1, 2]. We often see remnants of introgressed haplotypes from ancestral admixture in the genomes of extant species as a consequence [3]. In some instances, if interbreeding populations have not accumulated enough differences to maintain reproductive isolation and hybrids remain fit, then homogenization can lead to the reformation and merging of populations [4]. When the hybridization event itself causes reproductive isolation between the hybrid and parental species, then new hybrid species can form [5, 6]. On the other hand, when genetic incompatibilities have accumulated to a sufficient degree, reinforcement of reproductive barriers can be accentuated due to low hybrid fitness [7].

When two genetically divergent species hybridize, the resulting admixture generates novel genotypic and phenotypic combinations that selection can act upon. Transgressive phenotypes that arise by this process can outcompete parental phenotypes at intermediate habitats along the environmental gradient if hybrids have a greater relative fitness with respect to the parental species [2, 8, 9, 10, 11]. These hybrid zones are referred to as bounded hybrid superiority [12]. By contrast, if genetic incompatibilities have accumulated in the genomic background, hybrids may have reduced fitness. Hybrid zones following this motif, coined tension zones, are maintained by a balance between dispersal of parental forms into the hybrid zone and selection against hybrids [13, 14, 15]. Numerous examples of the two hybrid formations described above have been well documented [16]. However, not all hybrid zones can be neatly categorized as one or the other.

After hybridization and the establishment of a hybrid population, the genomic contributions from both parental species will segregate. During this time, recombination will decay linkage blocks that are not identical-by-descent. This in turn effects the distribution of phenotypes for se-

lection to act upon. However, hybrid zones are dynamic, complex systems driven by an interaction between ecology, demography, and evolutionary history [17]. Population sizes, migration rates, the strength of selection, and changes in local environments can all influence how hybrid populations evolve. While the scientific community has uncovered many explanations to what governs hybrid evolution, many questions remain. What is the interplay between selection and drift in driving hybrid phenotypic and genotypic evolution? How do early generation hybrids either constrain or promote phenotypic and genotypic variability? In other words, is there a large amount of variation in hybrid populations or are hybrid populations constrained to certain genotypic and phenotypic combinations. How are genotypes and phenotypes among hybrid populations changing over time and across different environmental regimes? Is there more variation within or between hybrid populations? Beginning to address these questions will provide novel insight into understanding the evolutionary outcomes of hybridization and its importance in speciation.

To explore these questions, I make use of two closely related swordtail fish, *Xiphophorus birchmanni* and *X. malinche* and the hybrid populations they form. The *birchmanni-malinche* system offers a unique opportunity to study long-standing evolutionary questions regarding the consequences of ecological selection on hybrid populations due to their unique natural history and ecological circumstances. These fish are endemic to the Sierra Madre Oriental in Mexico [18], and their distributions are in part determined by the thermal environment [19]. Populations of *X. malinche* are found at high elevations in cooler water temperatures (7-25 °C), while *X. birchmanni* are found at lower elevations in warmer waters (15-35 °C) – at intermediate elevations, hybrids form, likely as a consequence of anthropogenic disturbances in the chemical environment [20]. Females in this system rely heavily on the olfactory modality for conspecific mate recognition [21]. When pheromonal communication becomes disrupted, females become incapable of discerning between conspecific and heterospecific males. Hybrid formation subsequently ensues giving rise to novel combinations of genotypes and phenotypes. Despite hundreds of genetic incompatibilities having been identified in the system [22] and evidence that pure *X. birchmanni* females have larger brood sizes when crossed with pure *X. birchmanni* males versus hybrid males [23], hybrids are still viable

and can backcross and intercross successfully.

Multiple naturally occurring replicated hybrid zones that have experienced several generations of selection have been explored over the past decade [18, 22, 24]. Consequently, we have gained a tremendous amount of insight into the evolutionary dynamics of hybrid populations. For instance, we find some populations are structured, while others are not [24], and interestingly, the genomic cline between hybrid populations distributed along independent river drainages are centered at different elevations [25]. Access to reference genomes of both parental species and a robust genotyping pipeline enables us to conduct whole genome analyses. In addition, we can maintain them in a laboratory setting which allows us to breed and design controlled laboratory crosses for experiments. This includes F1 and later generation backcrossed or intercrossed hybrids which we can perform behavioral studies on that would difficult to conduct in the field. Together, by pairing whole genome and phenotypic data from natural and experimental hybrid populations I am able to address the inquiries proposed above.

In Chapter II of my dissertation, I investigate how phenotypic distributions of sexually selected male traits change over time among independent natural hybrid populations. This chapter is closely paired with Chapter III where I explore the evolutionary change in experimental, early generation hybrid populations. Early generation hybrids between *X. birchmanni* and *X. malinche* are infrequent in nature [24]. Nearly all populations are characterized by later generation hybrids that have experienced several bouts of selection over the past several decades. To get a clear understanding of how these processes are shaping hybrid zones and the distribution of species-specific alleles at functional loci during early generations of hybridization, it's important to monitor these populations using controlled experimental crosses.

The distribution of *X. birchmanni* and *X. malinche* are in part determined by their thermal environment [19]. As a consequence of local adaption to their respective thermal habitats, thermal tolerance varies between *X. birchmanni* and *X. malinche* in natural populations [19]. In Chapter IV, I aim to investigate putative loci under thermal selection and detect whether allele frequencies at these regions significantly deviate from the genome-wide expectation over time in natural hy-

brid populations. I use a combination QTL analysis and genomic cline analysis to investigate the genomic architecture underlying ecological traits and how these genomic regions evolve in hybrid populations. The genetic architecture of a trait will give insight into the evolutionary potential for these alleles to introgress across hybrid populations.

In my penultimate chapter, I investigate the role of postmating-prezygotic sexual selection in the *birchmanni-malinche* system. Specifically, I test whether females bias fertilization towards conspecific males when inseminated with equal proportions of conspecific and heterospecific sperm and whether sperm physiological traits correlate with fertilization success.

2. MORPHOLOGICAL EVOLUTION OF NATURAL HYBRID POPULATIONS OF SWORDTAIL FISH

2.1 Abstract

Hybridization is a common phenomenon that serves as an important evolutionary mechanism by which diversity can arise. Specifically, admixture generates novel combinations of phenotypes that selection can subsequently act upon. In this study, my aim is to characterize changes and differences in morphology among three independent natural hybrid zones between hybridizing sister species *Xiphophorus birchmanni* and *X. malinche*. Using a combination of univariate and multivariate approaches, I compare variation in male traits within and between populations. I found significant differences among populations for most of the phenotypic traits measured. In addition, each population exhibited some sign of stabilizing selection as trait means remain relatively unchanged while trait variance decreases. Principal component analysis revealed the two hybrid populations located at lower elevations to share similar morphotypes with the lowland parental species *X. birchmanni*, and that the third hybrid population located at higher elevations shared more similarity with the highland parental species, *X. malinche*. Overall, I found in the lowland populations, environment drives morphology to resemble the lowland parental species. In the highland hybrid population, morphology aligns with genome-wide hybrid index. These findings reveal independent hybrid populations can exhibit similar morphometric combinations of traits despite genome-wide ancestry composition, but that it is context dependent. In other cases, morphology aligns with genome-wide ancestry.

2.2 Introduction

Hybridization is a common phenomenon, but there are still questions about how hybrid phenotypes continue to evolve after hybridization [26]. For instance, a major question is how the stabilization of the hybrid genomes influences morphological trait evolution. Hybrid populations may express morphological traits that skew towards the parent species it derives a majority of its

genome from over time [27]. Alternatively, other selective regimes (e.g., sexual selection, thermal selection) and evolutionary forces like drift may take precedence in governing the distribution of phenotypes in hybrids [28]. Hybrid populations often resolve incompatibilities that arise as a consequence of admixture by selecting for allelic combinations that align with the majority parent ancestry genome-wide [22]. In turn, regions of the genome responsible for morphological variation should also have skewed ancestry, thereby effecting morphological trait variation. However, sexual selection is a strong evolutionary force driving the differentiation of morphological traits in this system that could act in opposition to selection against incompatibilities [21, 29]. Similarly, if the parental species of a hybrid system are adapted to distinct thermal environments, thermal selection could yield comparable signals [30]. For example, if genes regulating morphology are linked to loci governing thermal tolerance, then selection at the thermal tolerance loci can cause changes in allele frequencies at morphological genes that oppose the direction of intrinsic selection.

In this chapter, my primary aim is to explore patterns of morphological variation among three independent populations where hybrid swordtails between *X. birchmanni* and *X. malinche* naturally occur. These include Acuapa (ACUA, 476 m), Tlatemaco (TLMC, 480 m), and Aguazarca (AGZC, 980 m). ACUA is characterized by individuals that have X. birchmanni-skewed ancestry $\sim$25% *X. malinche* ancestry; [31]). TLMC is at a similar elevation as ACUA, but instead has *X. malinche*-skewed ancestry ($\sim$72% *X. malinche* ancestry; [24]). AGZC, on the other hand, is a structured population. There are two clusters that live in sympatry, but do not interbreed [24]. One cluster has *X. malinche*-skewed ancestry ($\sim$95% *X. malinche* ancestry) and the other has *X. birchmanni*-skewed ancestry ($\sim$25% *X. malinche* ancestry). If intrinsic postzygotic selection takes precedence, then I expect populations where the genome-wide ancestry is skewed towards *X. malinche* to resemble *X. malinche*-like morphology. In contrast, I expect populations where the genome-wide ancestry is skewed towards the lowland species, *X. birchmanni*, to exhibit more *X. birchmanni*-like morphology. Alternatively, morphology may be a function of environment, whereby hybrid populations located in environments similar to either parental species will express skewed morphology in the respective direction. Because population sizes can be relatively low in this system, genetic

drift may govern morphology trait distributions. In this case, I expect a reduction in variance over time and trait distributions that don't necessarily align with ecological expectations. Together, these independent hybrid zones offer a unique opportunity to study how wild hybrid populations are evolving on a seasonal and annual timescale.

I focus on male morphological traits. The two species are sexually dimorphic and are differentiated by several characteristics. *X. malinche* males are characterized by pigmented extensions of their caudal fins ("swords"), medium sized dorsal fins, and leaner body depths. Male *X. birchmanni* lack the sword, have larger dorsal fins, and wider body depths [32]. Hybrids display a wide range of intermediate and transgressive phenotypes as a result of admixture [32].

I conducted principal component analysis and analysis of variance tests on morphometric data collected semiannually for four years to determine how hybrid populations compare to one another and to the pure parental populations. In addition, I decomposed the time series data set and fit a seasonally adjusted linear regression to changes in morphology for each hybrid population over time to detect how morphology evolves over time.

I found significant differences among populations for most of the phenotypic traits measured. In the ACUA population, sword length residuals decreased, and in the AGZC population, body depth residuals increased over time. No traits significantly changed in the TLMC population. Principal component analysis revealed ACUA and TLMC share similar morphology with the lowland species, *X. birchmanni*, and that AGZC shares more similarity with the highland species, *X. malinche*.

2.3 Methods

2.3.1 Sample collection

We visited each natural hybrid site semiannually (January and June) to collect morphometric data. These sites included AGZC, TLMC, and AGZC (Table 2.1). At each site, fish were collected from the same pool using baited minnow traps. We sought to capture 50 mature males during each collection. When we were unable to capture up to 50 males we recorded juvenile males to

supplement. AGZC is a structured population that consists of a *X. birchmanni*-skewed cluster and an *X. malinche*-skewed cluster that are reproductively isolated. The two clusters can be visually distinguished by the presence or absence of a sword. Individuals that have a protrusion extending past the tangent of the caudal fin are considered sworded test (Figure 2.1). We photographed each fish next to a ruler on both sides so morphological traits could be measured. For all males, the dorsal fin, caudal fin, sword, and gonopodium were fully displayed. In addition, we collected morphometric data of pure *X. birchmanni* from the Rio Coacuilco (COAC) and pure *X. malinche* from the Rio Xontla (CHIC, Table 2.1). These two populations are characterized by nearly a million fixed, ancestry informative markers genome-wide. Before handling, all fish were anesthetized with tricaine methanesulfonate (MS-222). After handling, the fish were placed in fresh river water until they regained equilibrium and were released back into the river.

Site	Elevation (m)	Latitude	Longitude
ACUA	476	20.953	-98.568
TLMC	480	21.217	-98.790
AGZC	980	20.953	-98.568
COAC	183	20.552	-98.343
CHIC	1400	21.550	-98.352

Table 2.1: Elevation and coordinates of each natural hybrid population site and the parental species populations.

2.3.2 Morphometric analysis

For all mature males across the natural hybrid populations, we took measurements on continuous traits including standard length, body depth, sword length, dorsal width, dorsal height, gonopodium length, and peduncle height (Figure 2.1). Each trait was standardized by the standard length by fitting a linear model and the residuals were used for subsequent analysis. All morphometric data was scored in ImageJ [33]. I conducted a one-way ANOVA followed by a post-hoc Tukey honest significant difference test to determine significant differences in trait values

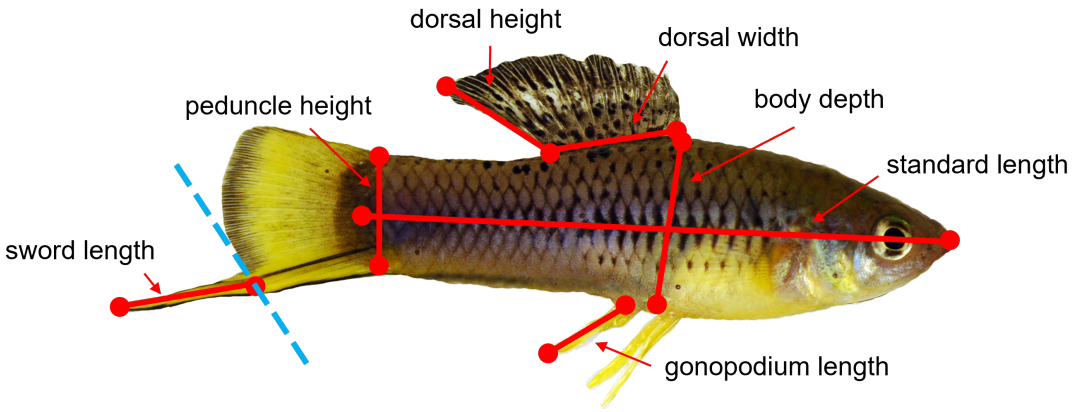

Figure 2.1: Measurement guide of continuous traits: standard length, body depth, sword length, dorsal fin width, dorsal fin height, peduncle height, and gonopodium length. Blue dash line indicates cut-off for sworded (i.e. *X. malinche*-skewed) and non-sworded (*X. birchmanni*-skewed) individuals. Individuals that have an extension past the line are considered sworded individuals.

between sites and fit decomposed seasonally adjusted linear models to determine how trait values and variance were changing over time. A seasonal time series consists of a trend component, a seasonal component and an irregular component. Decomposing the time series means separating and estimating these three components. All statistics were conducted in R.

2.3.3 Principal component analysis and Bayesian ellipse overlap analysis

Principal components analysis (PCA) was conducted on all continuous traits using the `prcomp` function in the R package stats. I calculated the 95% CI ellipse of each site and used the `bayesianOverlap` function in the R package SIBER to calculate the area of overlap between two ellipses. This is done pairwise between sites. This function returns the proportion of overlap between two ellipses, each representing the values for a particular draw from the posterior estimates so that a posterior distribution of the estimated overlap is obtained. The number of draws was set to 100. I use Kalman smoothing to impute missing data for any traits that were not recorded in some collection dates [34]. This fits the ellipses using an Inverse Wishart prior on the covariance matrix Sigma, and a vague normal prior on the means. This will range from 0 when the ellipses are completely dis-

tinct, to 1 when the ellipses are completely coincidental. The final set of posterior distributions are concatenated and used to determine which sites share more variation with one another. In addition to the continuous traits measured, I used the first two principal components (PC1 and PC2) in the analyses comparing means between populations and changes through time since they can explain most of the variance.

2.4 Results

2.4.1 Morphological change over time

I measured 7 morphological traits over time among 3 independent hybrid populations. After decomposition of the time series dataset, I found sword length residuals (Estimate: -0.02614, $F(2,9) = 7.754$, $p = 0.011$) to be significantly decreasing over time in the ACUA population (Figure 2.2). I found body depth residuals (Estimate: 0.105, $F(2,8) = 8.439$, $p = 0.0107$) to be significantly increasing over time in the AGZC population (Figure A.1.D). While non-significant, sword length residuals trended negatively in the AGZC population (Figure 2.2; Estimate: -0.578, $F(2,8) = 3.356$, $p = 0.087$). No traits significantly differed from a slope of zero in the TLMC population (Figure A.1).

Many traits exhibit a decrease in variance over time. In AGZC, variance in PC2 (driven primarily by sword length; $R^2 = 0.4$, $p = 0.036$), standard length ($R^2 = 0.63$, $p = 0.0036$), body depth ($R^2 = 0.49$, $p = 0.017$), sword length residuals ($R^2 = 0.39$, $p = 0.039$) and peduncle height ($R^2 = 0.46$, $p = 0.021$) decreased over time (Figure A.1). While non-significant, both ACUA and TLMC exhibit decreasing trends in variance for standard length, body depth, and peduncle height. Variance in PC2 and sword length residuals remained constant in both ACUA and TLMC (Figure A.1.B,E).

Most of the signal from the AGZC population derives from *X. malinche*-skewed individuals. I repeated the analysis within the AGZC population treating *X. malinche*-skewed and *X. birchmanni*-skewed individuals as separate groups. Variance in standard length ($R^2 = 0.65$, $p = 0.0048$), body depth ($R^2 = 0.71$, $p = 0.0023$), peduncle height ($R^2 = 0.64$, $p = 0.0053$), dorsal width ($R^2 = 0.63$, $p = 0.0062$), and dorsal height ($R^2 = 0.72$, $p = 0.0018$) significantly decreased in *X. malinche*-skewed,

but not *X. birchmanni*-skewed individuals. In all cases, variance remained unchanged over time in the *X. birchmanni*-skewedindividuals. While non-significant, variance in sword length had a decreasing trend in *X. malinche*-skewed, but not *X. birchmanni*-skewed individuals ($R^2 = 0.39$, $p = 0.054$).

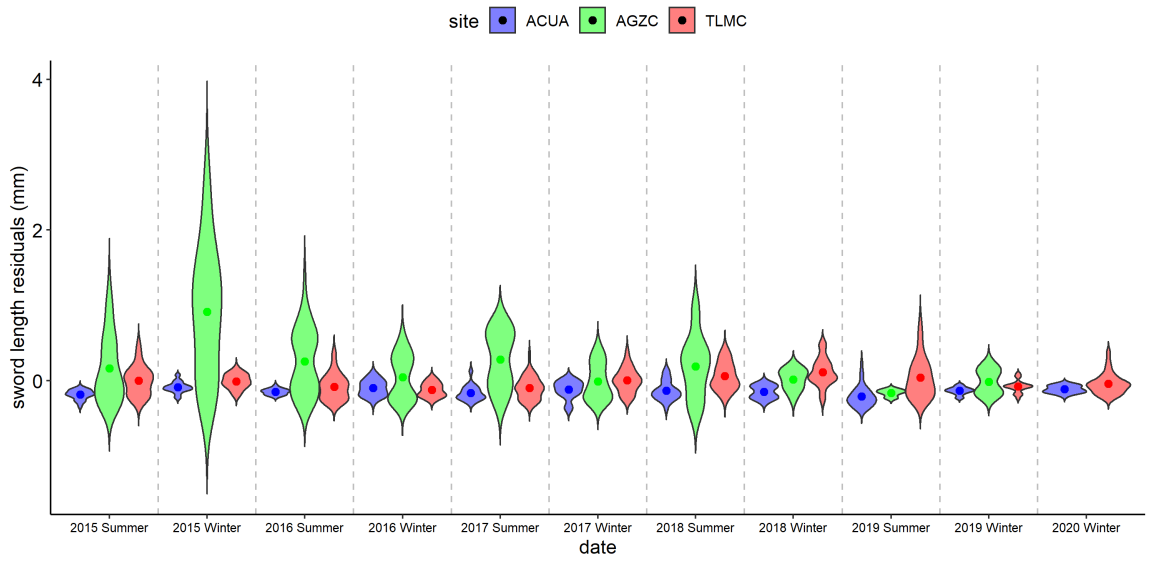

Figure 2.2: Distribution of sword length residuals for the natural hybrid populations over all collection dates.

2.4.2 Morphological trait distributions

I conducted one-way ANOVAs followed by post-hoc Tukey HSD test to determine if residual trait values significantly differed between hybrid and parental populations. For PC1 ($F(4,1132) = 10.06$, $p < 0.001$) and standard length ($F(4,1132) = 8.132$, $p < 0.001$), AGZC was significantly different from the other two hybrid populations, but not from either of the parental populations (Figure 2.3.A,C). ACUA and TLMC did not significantly differ in PC1 values (Figure 2.3.A). All groups significantly differed from each other for PC2 ($F(4,1132) = 120.9$, $p < 0.001$) and sword length residuals ($F(4,1132) = 120.5$, $p < 0.001$), except ACUA and wild *X. birchmanni* (Figure 2.3.B,E). Wild *X. malinche* had the greatest residual sword length values followed by AGZC,

TLMC, ACUA and wild *X. birchmanni*, respectively (Figure 2.3.E). The wild *X. birchmanni* population had significantly greater body depth residual trait values among the different groups (Figure 2.3.D, $F(4,1132) = 11.65$, $p < 0.001$). No other populations differed from one another for body depth residual trait values. AGZC and TLMC shared similar dorsal width residual trait values and both were significantly different from all other populations (Figure 2.3.G, $F(4,1132) = 93.39$, $p < 0.001$). ACUA did not significantly differ from either of the parental populations in dorsal width residual trait values (Figure 2.3.G). TLMC had significantly different dorsal height residual trait values than all other populations (Figure 2.3.F, $F(4,1132) = 9.926$, $p < 0.001$), while AGZC, ACUA, and wild *X. malinche* shared similar trait values. Only ACUA significantly differed in gonopodium length between populations (Figure 2.3.H, $F(4,1132) = 7.599$, $p < 0.001$). TLMC had the lowest peduncle height residual trait values and significantly differed from all other populations (Figure 2.3.J, $F(4,1132) = 20.99$, $p < 0.001$). ACUA and AGZC shared similar peduncle height residual trait values and differed from all other groups (Figure 2.3.J).

2.4.3 Summer vs winter collections

In the ACUA population, summer and winter collections significantly differed in standard length (Figure 2.3.C; $t(302.48) = 2.7853$, $p = 0.005686$), sword length residuals (Figure 2.3.E; $t(290.35) = -5.0964$, $p = 6.247e-07$), and PC2 (Figure 2.3.B; $t(280.02) = -6.1967$, p-value = 2.054e-09). Individuals collected in the winter had increased values. Standard length (Figure 2.3.C; $t(439.28) = 4.8757$, $p = 1.518e-06$) and PC1 (Figure 2.3.A; $t(444.37) = 5.4209$, $p = 9.745e-08$) were significantly different between summer and winter collections in TLMC. Only PC1 was significantly different between seasons in AGZC (Figure 2.3.A; $t(292.62) = 3.4727$, $p = 0.0005931$).

2.4.4 Principal component analysis

I conducted a principal components analysis that included all morphometric data from the three hybrid populations as well as the parental species. The first principal component explained 71.1% of the variance and was driven equally between all traits except for sword length (Figure 2.4). Principal component 2 explained 14.3% of the variation and was strongly driven by sword

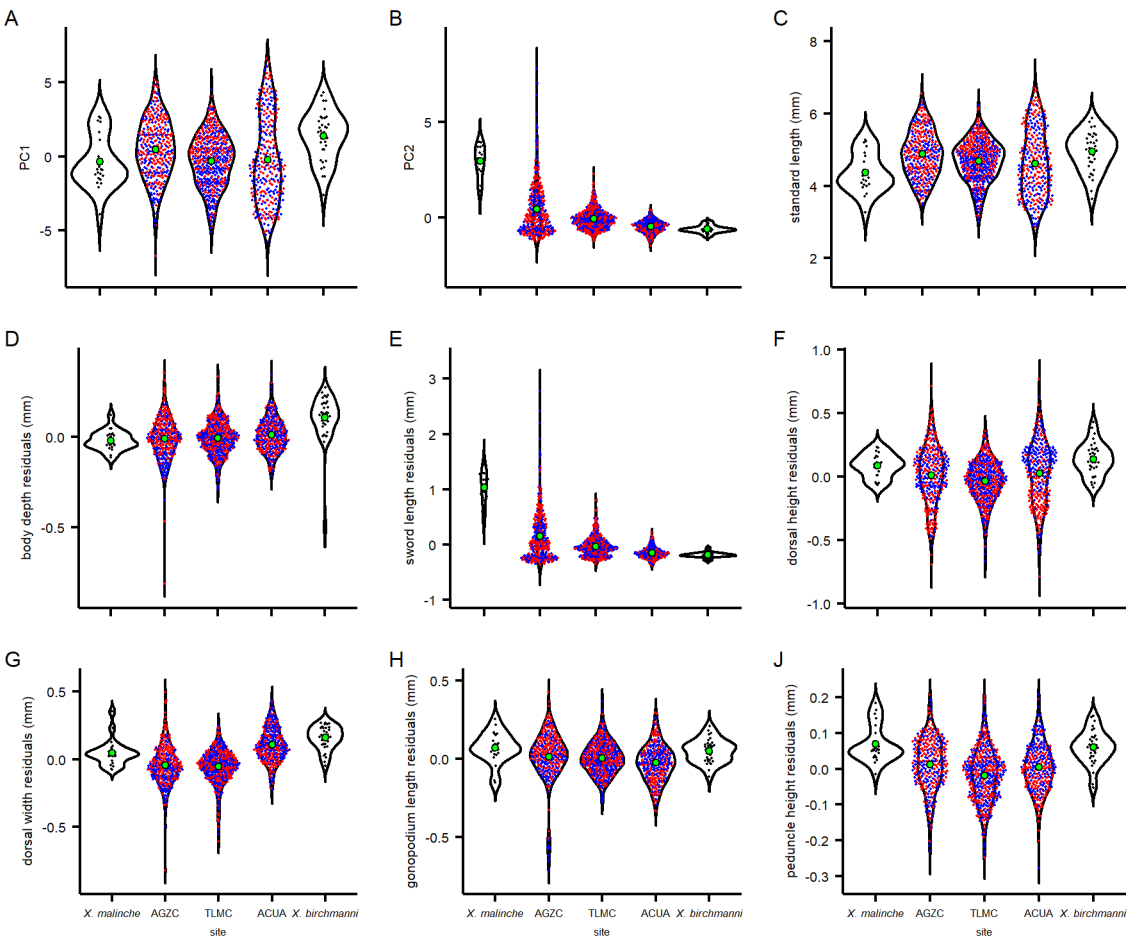

Figure 2.3: Distributions of A) PC1, B) PC2, C) standard length, D) body depth residuals, E) sword length residuals, F) dorsal height residuals, G) dorsal width residuals, H) gonopodium length residuals, and J) peduncle height residuals in the natural hybrid populations and parental populations among all collection dates combined. Red and blue points represent individuals collected in summer and winter, respectively. Green point represents population mean.

length (Figure A.2). TLMC and ACUA cluster closely and overlap with each other as well as with the *X. birchmanni* parent population. The parent *X. malinche* samples and the hybrid AGZC population have the greatest variation (presumably due to high variation in sword length). I reran the principal component analysis removing the lowland hybrid populations and separating AGZC into two groups: *X. malinche*-skewed and *X. birchmanni*-skewed clusters (Figure 2.5, Figure A.3). The *X. malinche*-skewed individuals contribute a majority of the variation present in AGZC. I also analyzed each site independently treating each date as its own cluster. Across each hybrid site,

variation in morphology is decreasing over time, as the area of 95% confidence interval of each site gets smaller over time (Figure A.4).

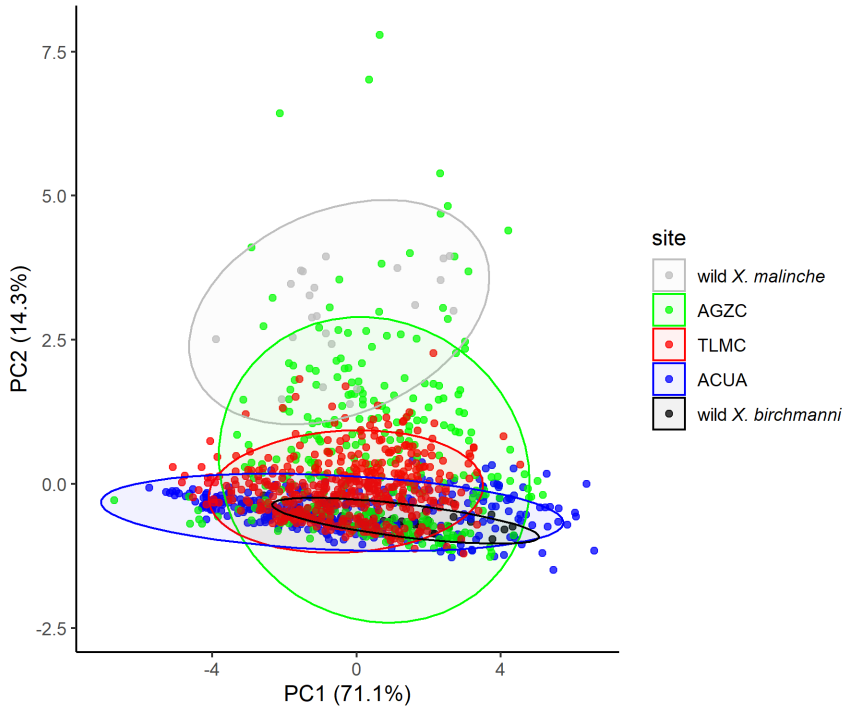

Figure 2.4: Principal components 1 and 2 of all morphological data scored between the natural hybrid sites and the parental species populations.

2.4.5 Pairwise overlap comparison

I conducted pairwise comparisons to calculate the percentage of overlap between principal component ellipses (Table 2.2, Figure A.5). ACUA and TLMC overlapped with the *X. birchmanni* samples (BIR) 26 and 20%, respectively. AGZC overlapped with BIR only 9%. AGZC was the only population to overlap with MAL at 26%. ACUA and TLMC shared the greatest amount of overlap of 46%. AGZC overlapped with ACUA and TLMC 23 and 32%, respectively. Pure parental species did not share any overlap.

In the AGZC population, *X. malinche*-skewed individuals overlapped with BIR 12% and with MAL 35% (Table 2.3, Figure A.6). Male *X. birchmanni*-skewed individuals overlapped with BIR

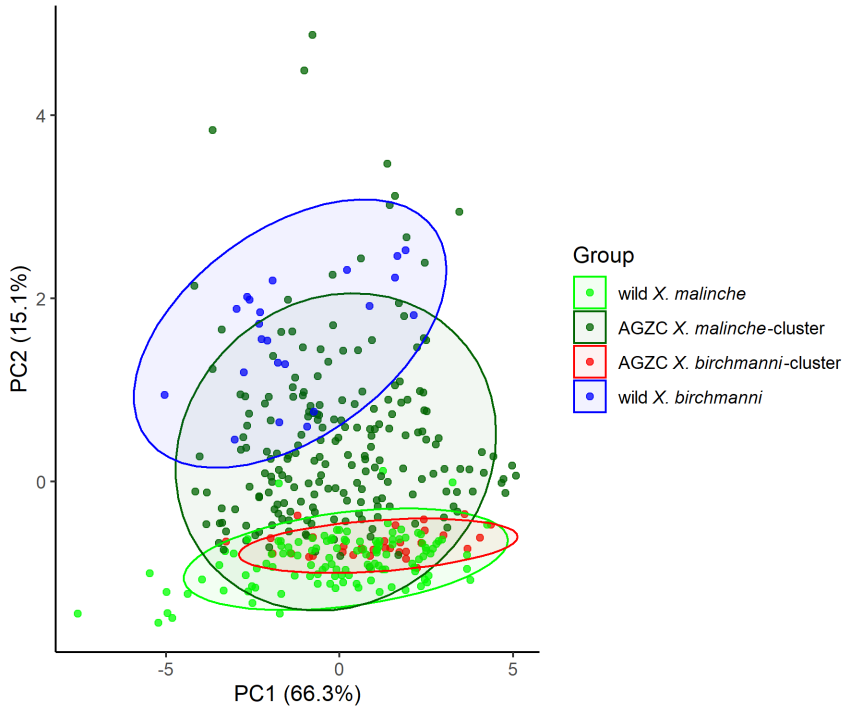

Figure 2.5: Principal components 1 and 2 of all morphological data scored between the AGZC hybrid clusters and the parental species populations.

43% and with MAL 0%. Male *X. malinche*-skewed individuals and *X. birchmanni*-skewed individuals shared 23% overlap.

2.5 Discussion

The primary objective of this study was to monitor phenotypic evolution across independent natural hybrid populations and compare variation within and between sites. The three hybrid populations consist of two lowland sites located at similar elevations yet differ in genome-wide hybrid index, and a structured population located at higher elevations that consist of a *X. malinche*-skewed cluster and an *X. birchmanni*-skewed cluster. Studies that track changes in morphology over time in hybrid populations have found hybrid phenotypic distributions to fall intermediately between parental distributions or express transgressive phenotypes that fall outside the boundaries of either parental species [35]. However, these are continuous distributions that fall along a continuum, and in nature, phenotypes can be distributed and centered anywhere along that continuum; that is, hy-

Population 1	Population 2	Percent overlap
BIR	ACUA	25.95
BIR	AGZC	8.73
BIR	TLMC	19.22
BIR	MAL	0.00
ACUA	AGZC	23.41
ACUA	TLMC	47.71
ACUA	MAL	0.10
AGZC	TLMC	32.28
AGZC	MAL	25.25
TLMC	MAL	1.95

Table 2.2: Percent of ellipse overlap between all pairwise combinations of natural hybrid sites.

Population 1	Population 2	Percent overlap
BIR	MC	11.85
BIR	BC	42.58
BIR	MAL	0.00
MC	BC	23.29
MC	MAL	34.94
BC	MAL	0.00

Table 2.3: Percent of ellipse overlap between all pairwise combinations of AGZC hybrid clusters and parental species. BIR = *X. birchmmani*, MAL = *X. malinche*, MC = AGZC *X. malinche*-cluster, BC = AGZC *X. birchmanni*-cluster.

brid phenotypes can exhibit skew towards one parental species over the other and don't necessarily fall directly in between parental phenotypes. This pattern is demonstrated in *Hyla arborea*. Using similar univariate and multivariate approaches to those conducted in my analysis, morphotypes of both parental species and hybrids differed from one another and hybrids were neither intermediate nor transgressive, but instead were skewed more towards one of the parental species over the other [36].

To determine how much variation was shared within and between populations, I conducted a principal component analysis on seven morphological traits in males and calculated the proportion of principal component space each site shared among each other. ACUA and TLMC shared 26

and 20% of principal component space with the *X. birchmanni* samples, respectively. Overall, I found ACUA and TLMC to share the greatest proportion of principal component space among any two sites. I expect this pattern if ecological selection were driving morphological evolution. Conversely, AGZC shared only 9% with *X. birchmanni*, but 26% with *X. malinche*. Similarity to the highland species in the AGZC population was driven by *X. malinche*-skewed males. When I analyzed the AGZC population separately, I found *X. malinche*-skewed males (i.e. individuals with genome-wide hybrid index skewed towards *X. malinche*) to share more principal component space with *X. malinche* and vice versa for *X. birchmanni*-skewed males, suggesting in this hybrid population, morphology aligns more with genome-wide hybrid index.

For PC1 and standard length, AGZC was significantly different from the other two hybrid populations. ACUA and TLMC did not significantly differ in PC1 values. Since variance in PC1 is explained equally among all continuous measured traits (with the exception of sword length) this provides further evidence that these two independent hybrid populations are more similar to one another and to the parental *X. birchmanni* species compared to the highland hybrid populations and *X. malinche*. This pattern is observed in TLMC despite individuals in this populations having a greater genome-wide hybrid index compared to ACUA samples. A near identical pattern was observed in a 50 year long-term analysis of *Helianthus bolanderi* - *H. annuus* hybrids [37]. In this population, hybrids expressed an initial bias towards *H. bolanderi*, but later shifted to a predominance of H. annuus-like plants. Moreover, morphological similarity was more pronounced than neutral genetic markers. These results are consistent with lowland hybrid populations being driven to express characteristics derived from the lowland parental species and vice versa in the highlands.

The decomposed time series linear regressions models revealed how means and variance components of various morphological traits were evolving over time. In ACUA, sword length significantly decreased. This pattern matches what we would expect if ecological selection were responsible for governing the evolution of the sword in this population. That is, as sword lengths decrease is ACUA, they will resemble more *X. birchmanni*-like morphology. While non-significant, AGZC

exhibited a negative trend for sword lengths that was driven primarily by the *X. malinche*-skewed males in this population.

Interestingly, the *X. malinche*-skewed males in the AGZC population expressed larger body depths over time. Larger body depths are associated with *X. birchmanni* relative to *X. malinche* and other hybrids. The wild *X. birchmanni* population had significantly greater body depth trait values among the different groups. No other populations differed from one another for body depth trait values. In addition, over the course of my sampling period, the frequency of *X. malinche*-skewed individuals in the AGZC population has decreased (Figure A.7). These findings suggest the *X. malinche*-skewed cluster of males at AGZC is evolving phenotypes that resemble more *X. birchmanni*-like characteristics (i.e. smaller swords and larger body depths). We expect highland hybrid populations effected by increasing water temperatures due to climate change to evolve characteristics that more closely resemble *X. birchmanni*.

In the ACUA population, summer and winter collections significantly differed in sword length. Individuals collected in winter expressed larger swords than those collected in summer. While general water temperature is an important evolutionary driving forces for adaption in this system, rapid fluctuations and hot and cold extremes are equally important [19]. Seasonal effects play an important role in determining phenotypic distributions among many taxa (e.g. insects [38], mammals [39], and birds [40]).

No traits significantly changed in TLMC over the course of the study suggesting strong stabilizing selection in this population. Variance in several traits decreased over time in AGZC – which suggest stabilizing selection may be an important evolutionary force in this population as well. This signal was driven by population structure at AGZC. When separated into *X. malinche*-skewed and *X. birchmanni*-skewed groups, only *X. malinche*-skewed individuals exhibited the decrease in variance while the *X. birchmanni*-skewed individuals did not. While non-significant, the other populations exhibited decreasing trends in trait variance – further evidence for stabilizing selection.

All groups significantly differed from each other for PC2 and sword length residuals, except ACUA and wild *X. birchmanni*. Since sword length explains most of the variance in PC2, I expect

ACUA and wild *X. birchmanni* to share similar values for sword lengths and PC2 if ecological selection is driving trait evolution. Among the hybrid sites, AGZC had the greatest sword length values followed by TLMC and ACUA, respectively. Since AGZC has a cluster of *X. malinche*-skewed hybrids and is a highland hybrid site, this finding is consistent with ecological expectations. However, we expect this pattern to attenuate over time since the *X. malinche*-skewed hybrids at this site are exhibiting directional selection towards *X. birchmanni* phenotypes as well as decreasing in frequency in general.

During speciation, trait divergence depends critically on the selective forces acting on the population. Populations under stabilizing selection increases transgression in hybrids and thereby increases the possibility of novel adaptation [41]. Each population exhibited some sign of stabilizing selection as trait means remain relatively unchanged while trait variance decreases. Overall, I found lowland hybrid populations in this system to resemble the lowland parental species over time and vice versa for the highland hybrid population.

Overall, I found in the lowland populations, environment drives morphology to resemble the lowland parental species. In the highland hybrid population, morphology aligns with genome-wide hybrid index. These findings reveal independent hybrid populations can exhibit similar morphometric combinations of traits despite genome-wide ancestry composition, but that it is context dependent. In other cases, morphology aligns with genome-wide ancestry. Together, this study sheds insight into how independent replicated natural hybrid populations evolve over time.

3. MORPHOLOGICAL AND GENOMIC EVOLUTION OF EXPERIMENTAL EARLY GENERATION HYBRID SWORDTAIL FISH

3.1 Abstract

The incipient stages of hybridization and hybrid zone formation significantly influence how hybrid populations evolve. Early generation hybrids can express greater trait variation compared to parental species due to recombination generating novel combinations of phenotypes. How variation and mean trait values change and compare among treatments will yield valuable insight into the way hybrid populations respond immediately after hybrid formation. Therefore, this project aims to characterize changes and differences in morphology among replicated experimental hybrid populations located at low, intermediate, and high elevations. Using a combination of univariate, multivariate, and Bayesian approaches, I compare variation in male traits within and between populations over time. Most notably, I found more variation was shared between sites than within sites. Across each experimental hybrid site, the amount of variation in morphology among replicates was relatively the same. Additionally, I make use of whole genome-wide sequence data for a subset of the individuals in the study and found the average hybrid index at each site trended in the direction that would be expected if environmental selection was driving genome-wide changes in ancestry.

3.2 Introduction

The first few generations of hybridization can have a crucial impact on the distribution of morphological traits, species-specific alleles, and the genome-wide ancestry of subsequent generations [1, 42]. For example, F1 hybrids often have asymmetric fitness in one cross versus the other [43, 44]. Further, in species with sex chromosomes, the heterogametic sex can suffer from reduced fitness due to Haldane's rule, whereby the heterogametic sex expresses recessive deleterious alleles on the X chromosome that cannot be rescued [45, 46, 47]. If hybrids are viable and survive beyond the first few generations, recombination will generate novel combinations of alleles that in

turn can produce extreme phenotypes that fall outside the distribution of either parental species, a phenomenon known as transgressive segregation. Alternatively, hybrids may express phenotypes that are intermediate between the parental species. In any case, these novel phenotypes may be advantageous or deleterious with respect to natural or sexual selection [48], and different selective forces can act on the same phenotype in opposition. For instance, some traits may be favored by sexual selection, but suffer from intrinsic natural selection due to genetic incompatibilities. Of course, a suite of environmental and demographic factors influence the evolutionary trajectory of traits as well, particular migration rates and population sizes.

Early generation hybrids in the *Xiphophorus birchmanni-malinche* system are rarely found in nature. The hybrid zones we monitor and sample from regularly are anywhere from 30-60 generations old. In this chapter, I am interested in exploring a similar set of questions as in chapter 2 but tested using replicated experimental populations of early generation hybrids. Specifically, how do early generation hybrid phenotypes evolve during the incipient stages of hybridization across different thermal environments? Do hybrids skew morphometrically and genotypically towards the parental species it shares a common environment to? Is there more variation within or between hybrid populations?

If ecological selection is playing a major role from generation to generation, then I predict hybrids seeded in the highlands will evolve to become more *X. malinche*-like, both morphometrically and genotypically, while hybrids seeded in the lowlands will evolve to become more *X. birchmanni*-like. Hybrids seeded at the intermediate site will evolve intermediate phenotypes and genotypes (e.g. intermediate sword length relative to parentals and a hybrid index $\sim$0.5). If ecological selection is driving differentiation between hybrids at highland vs lowland, I predict there will be more variation between sites than within sites.

To this end, we generated F1 generation hybrid populations. We seeded replicated mesocosm stock tanks at high, intermediate, and low elevations with the F1 hybrids. Each treatment experiences divergent ecological selection with respect to water temperature. Semiannually, we collect genotype and phenotype data from these experimental hybrid populations for the purpose of mon-

itoring morphometric and genomic evolution through time. Specifically, we measure standard length, body depth, dorsal width, dorsal height, sword length, gonopodium length, and peduncle height in males. Female mating preferences for several of these traits have been characterized by numerous studies [21, 29, 49, 50, 51, 52, 53, 54, 55].

We found the lowland experimental hybrid population to share slightly more principal component space with the parental *X. birchmanni* population than the other two hybrid populations and that more variation was shared between sites than within sites. Across each experimental hybrid site, the amount of variation in morphology among replicates was relatively the same. After decomposing the time series dataset, we found significant changes in several morphological traits in the intermediate experimental hybrid population, but no others. The intermediate hybrid population also exhibited a significantly lower overall distribution of traits values among the hybrid populations. While no sites significantly differed from a slope of zero, the average hybrid index at each site trended in the direction that would be expected if environmental selection was driving genome-wide changes in ancestry. Interestingly, in the lowland hybrid population, sword length was positively associated with hybrid index. Individuals from the intermediate hybrid population on average had the greatest lifespan followed by the highland and the lowland populations, respectively. The intermediate population had larger growth rates for several morphological traits relative to the highland and lowland populations, except for sword length, where the lowland population had the greatest growth rate.

3.3 Methods

3.3.1 Design of experimental hybrid populations

Eight replicate 2000 L artificial mesocosm stock tanks were built at high (1514 m, 20°53’29.8"N 98°38’30.0"W), intermediate (980 m, 20°53’52.2"N 98°36’05.1"W), and low (186 m, 20°59’23.8"N 98°22’15.3"W) elevations near the CICHAZ field station in Calnali, Hidalgo, MX. These tanks were fed by a constant low-rate flow of dechlorinated municipal tap water. Each replicate was seeded with 10 male and 10 female F1 hybrids. To generate F1 hybrids, gravid female *X. ma-*

linche were collected with baited minnow traps from the Chicayotla locality on the Rio Xontla (20°55'27.24"N 98°34'34.50W) and housed in the mesocosms at the CICHAZ field station. Because females store sperm for several months in poecilid fishes [56, 57], we needed to rear virgin females for the cross. Thus, after parturition, fry were removed from the adults and reared to maturity in the absence of males to ensure virginity. Once mature, these females were then crossed with wild *X. birchmanni* collected from the Rio Coacuilco (21°5'50.85 N, 98°35'19.46 W). The progeny of this cross were used as the founding F1 population for each replicate across all sites. These F1s were allowed to freely interbreed for the remainder of the experiment. We raised *X. birchmanni* in a common garden at the same elevation as the intermediate population as a baseline comparison between hybrid and parental phenotypes. We attempted to raise *X. malinche* in a similar way, but all samples died prior to maturity, thus we only compare hybrid phenotypes to the lowland parental species, *X. birchmanni*. Fish at all sites were fed the same diet daily (Ken's Premium Spirulina Pellets). Two HOBO temperature loggers were deployed at each stock tank site that record the water temperature every 6 hours.

3.3.2 Sample collections

We visited each stock tank site semiannually to collect morphometric and genetic data. For each replicate, all mature fish were sorted by sex and into either untagged or previously tagged categories. Untagged fish (i.e. newly captured fish; fish born after the previous collection date) were tagged with a unique color code elastomer ID (Northwest Marine Technologies) so that the individual could be tracked in future collection dates. Next, we photographed the fish on both sides so morphological traits can be measured. For mature males, the dorsal fin, caudal fin, sword, and gonopodium were fully displayed for at least one of the two photographs. Finally, a 2 mm fin clip is taken from the distal top quarter of the caudal fin and stored in 95% ethanol for future genotyping. Previously tagged fish (i.e. fish that were collected in a previous season and have already been ID'd) were only photographed to measure phenotypic growth through time. Before handling, all fish were anesthetized with tricaine methanesulfonate (MS-222). After handling, the fish were placed in an aerated recovery tank until they regained equilibrium and were returned to

their respective stock tank.

3.3.3 Morphometric analysis

For all mature males in the experimental hybrid populations, we took measurements on continuous traits including standard length, body depth, sword length, dorsal width, dorsal height, gonopodium length, and peduncle height (Figure 2.1). Each trait was standardized by the standard length by fitting a linear model and the residuals were used for subsequent analysis. All morphometric data was scored in ImageJ [33]. This dataset includes all males that have been collected up to March 2021. If a male had been recaptured one or more times, I only used the first set of scored morphometrics beginning at its maturity. I conducted a one-way ANOVA followed by a post-hoc Tukey honest significant difference test to determine significant differences in trait values between sites and fit decomposed seasonally adjusted linear models to determine how trait values and variance were changing over time. A seasonal time series consists of a trend component, a seasonal component and an irregular component. Decomposing the time series means separating and estimating these three components. Lastly, I calculated the lifespan and growth rate of individuals that survived at least one collection date. I calculate the minimum number of days each individual survived as the last date collected subtracted by the first date collected. If an individual lived for only two collection dates (e.g. collected at timepoint 0 as untagged and again at timepoint 1 as previously tagged) then the growth rate was calculated as:

$$growth\ rate = phenotype_{t1} - phenotype_{t0}/days\ alive$$

If an individual lived for more than two collection dates then the growth rate was calculated as the coefficient of a linear model:

$$growth\ rate = coeff[lm(phenotype \sim time)]$$

I conducted independent one-way ANOVAs to determine if site effects lifespan or growth rates. All statistics were conducted in R.

3.3.4 Principal component analysis and ellipse overlap

Principal components analysis (PCA) was conducted on the continuous traits using the prcomp function in the R package stats. I calculated the 95% CI ellipse of each site and tank and used the `bayesianOverlap` function in the R package SIBER to calculate the area of overlap between two ellipses. This is done pairwise between and within sites. This function returns the proportion of overlap between two ellipses, each representing the values for a particular draw from the posterior estimates so that a posterior distribution of the estimated overlap is obtained. The number of draws was set to 100. I use Kalman smoothing to impute missing data for any traits that were not recorded in some collection dates [34]. This fits the ellipses using an Inverse Wishart prior on the covariance matrix Sigma, and a vague normal prior on the means. This will range from 0 when the ellipses are completely distinct, to 1 when the ellipses are completely coincidental. The final set of posterior distributions are concatenated and used to determine whether there is more variation within or between sites. In addition to the continuous traits measured, I used the first two principal components (PC1 and PC2) in the analyses comparing means between populations and changes through time since they can explain most of the variance.

3.3.5 DNA extraction, library preparation, and sequencing

I applied the multiplexed shotgun genotyping (MSG) approach for genome-wide genotyping of hybrids [24, 58]. DNA was extracted using an Agencourt DNAdvance bead-based purification kit (Beckman Coulter) following the manufacturer's instructions except we used half reactions and diluted to a 2.5 ng/μL. Each sample was sheared using the Illumina Tagment DNA TDE1 Enzyme and Buffer Kits, and amplified in a dual-index PCR reaction with the conditions outlined in Table 3.1.

PCR products were pooled and purified with 18% SPRI magnetic beads (Beckman Coulter). The library was quantified using a Qubit 3.0 fluorometer, run on a Agilent 4200 TapeStation to assess the library size distribution, and finally sequenced on an Illumina Genome Analyzer (HiSeq 4000 PE 150 bp reads).

Step	Temp	Time
1	68°C	3 min – Extend Tn5 transposon ends
2	95°C	30 sec
3	95°C	10 sec
4	55°C	30 sec
4	68°C	30 sec
5		Cycle to Step 3 12 times
6	68°C	5 minutes (final extension)
7	4°C	Hold

Table 3.1: Polymerase chain reaction conditions for Illumina tagmentation library preparation.

3.3.6 Genome-wide genotyping

Raw reads were parsed by individual and then mapped to both the parental reference genomes using BWA [59]. A hidden Markov model was applied to reads from each sample which assigns ancestry probabilities at hundreds of thousands of fixed ancestry informative markers across the genome. Hard calls are made at ancestry informative markers that pass a 95% threshold. The hybrid index for each individual can be calculated as the number of *X. malinche*-called alleles divided by the total number of alleles called genome-wide. I conducted one-way ANOVAs to determine if there were significant differences in hybrid index between sites. In addition, I fit linear models to determine if the hybrid index was significantly increasing or decreasing over time and whether hybrid index was positively or negatively associated with morphological traits.

3.4 Results

3.4.1 Principal component analysis and pairwise overlap comparison

I conducted a principal components analysis that included all morphometric data from the three experimental hybrid populations as well as the common garden parental *X. birchmanni* species. The first principal component explained 75.2% of the variance and was driven equally between all traits except for sword length (Figure 3.1). Principal component 2 explained 11.7% of the overall variation and was strongly driven by sword length (77.4%, Figure B.1). The highland site had the greatest amount of variation (i.e. the ellipse with the largest area), followed by the other

two hybrid populations and then the parent *X. birchmanni* population. We also analyzed each site independently with each replicate representing its own cluster. Across each hybrid site, the amount of variation in morphology among replicates is relatively the same (i.e. each replicate has a similar ellipse area, Figure 3.2).

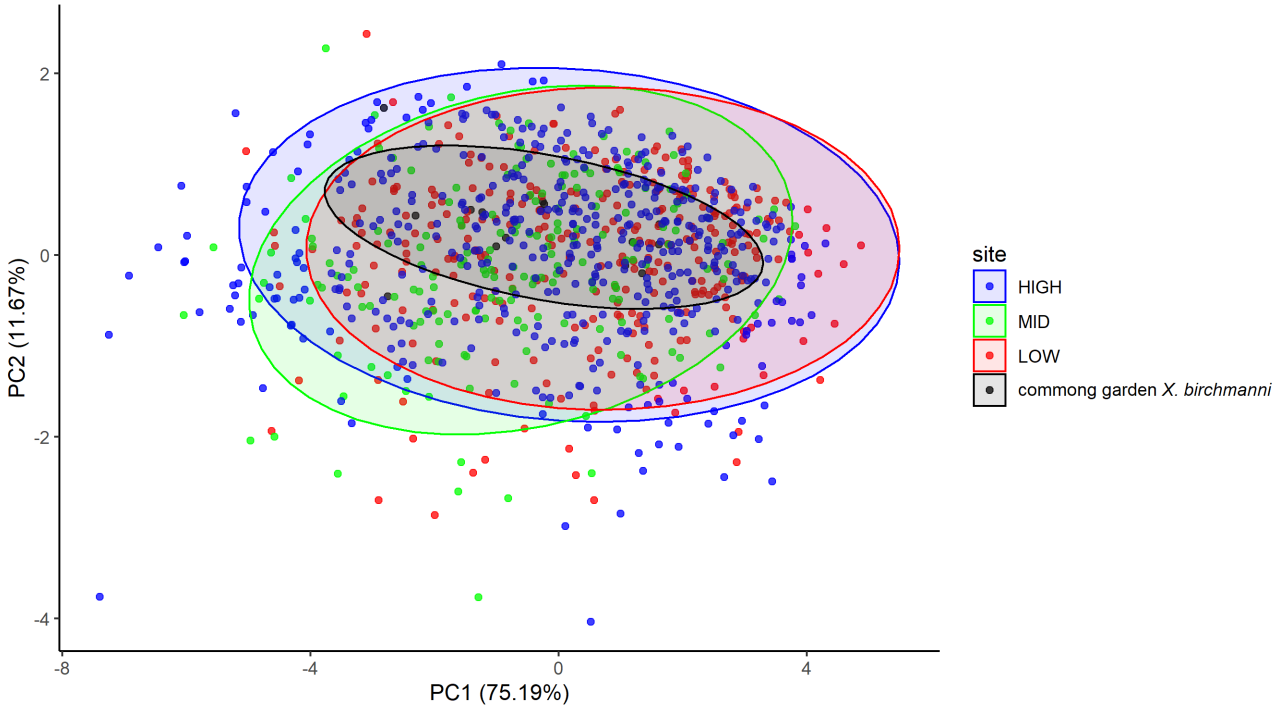

Figure 3.1: Principal components 1 and 2 of all morphological data scored between the experimental, early generation hybrid sites and the parental species, *X. birchmanni*.

The parent *X. birchmanni* samples shared the most variation with the lowland (38%) and the intermediate (38%) populations (Table 3.2, Figure B.2). The parent *X. birchmanni* samples shared only 32% of the principal component space with the highland hybrid samples. Interestingly, among the experimental hybrid populations, the lowland and highland populations shared the most variation (80%). The intermediate population shared 71% and 72% with the lowland and highland experimental hybrid populations, respectively. Overall, we found more variation was shared between sites than within sites when we concatenated pairwise ellipse overlap proportions from within sites

Population 1	Population 2	Percent overlap
BIR	LOW	37.52
BIR	MID	38.03
BIR	HIGH	32.35
LOW	MID	71.49
LOW	HIGH	78.98
MID	HIGH	72.46

Table 3.2: Percent of ellipse overlap between all pairwise combinations of experimental, early generation hybrid sites and the parental species, *X. birchmanni*.

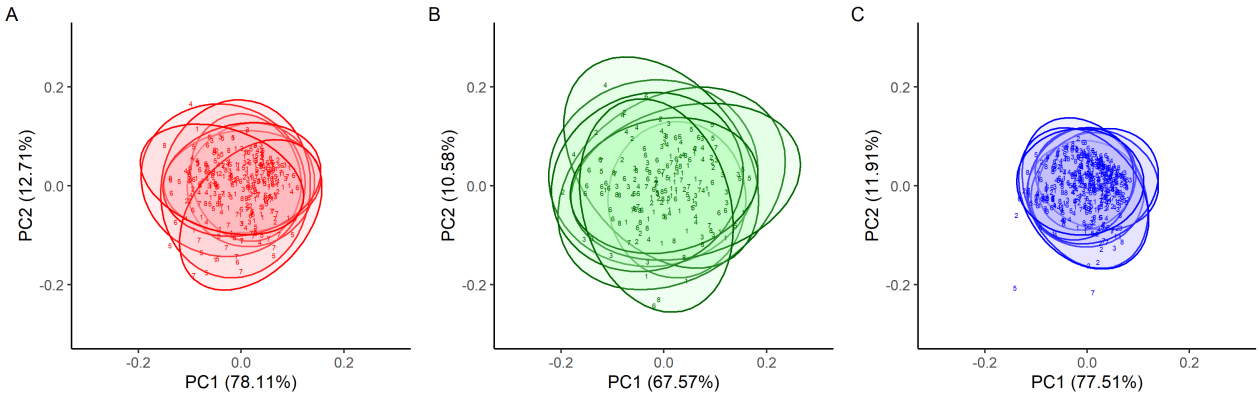

Figure 3.2: Principal components 1 and 2 for all morphological data conducted among replicates within A) LOW, B) MID, and C) HIGH. Numbered points represent an individual's respective tank.

and between sites (Figure B.2). That is, there was more similarity among sites than among replicates within sites.

3.4.2 Morphological change over time

I measured 7 morphological traits for four years among 3 independent experimental hybrid populations located at different elevations. After decomposition of the time series dataset, I found season to have a positive effect on sword length residuals in the intermediate hybrid population over time (Figure 3.4.E, $F(2,4) = 16$, $p = 0.012$). Individuals that were collected in the summer had greater sword length values relative to individuals collected in the winter in the intermediate ($t(57) = 4.4333$, $p < 0.001$) and highland ($t(250) = 5.5181$, $p < 0.001$) hybrid populations (Figure 3.5.E).

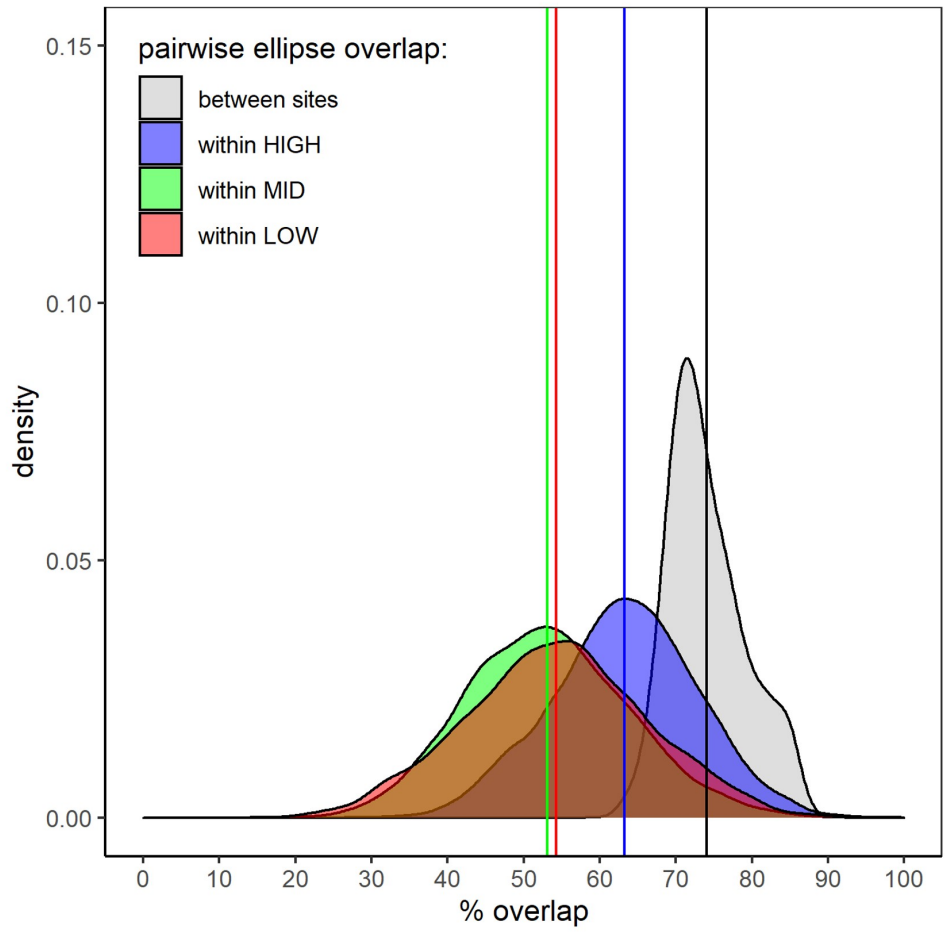

Figure 3.3: Concatenated posterior probability distributions of pairwise ellipse overlap comparisons between versus within sites. Values closer to 100 indicade complete ellipse overlap and values closer to zero indicate no overlap between ellipses.

While non-significant, both the lowland and highland populations exhibited decreasing trends in variance for standard length ($R^2 = 0.46$, $p = 0.063$, $R^2 = 0.49$, $p = 0.052$), while the intermediate site exhibited increasing trends ($R^2 = 0.56$, $p = 0.052$). Variance in dorsal width significantly decreased in the highland hybrid population ($R^2 = 0.51$, $p = 0.047$). The variance for all other traits across all hybrid populations did not significantly differ from a slope of zero. However, the overall mean variance between sites for dorsal height ($F(2,20) = 5.926$, $p = 0.0095$) and width residuals ($F(2,20) = 5.351$, $p = 0.014$) was significantly greater in the intermediate population compared to the others.

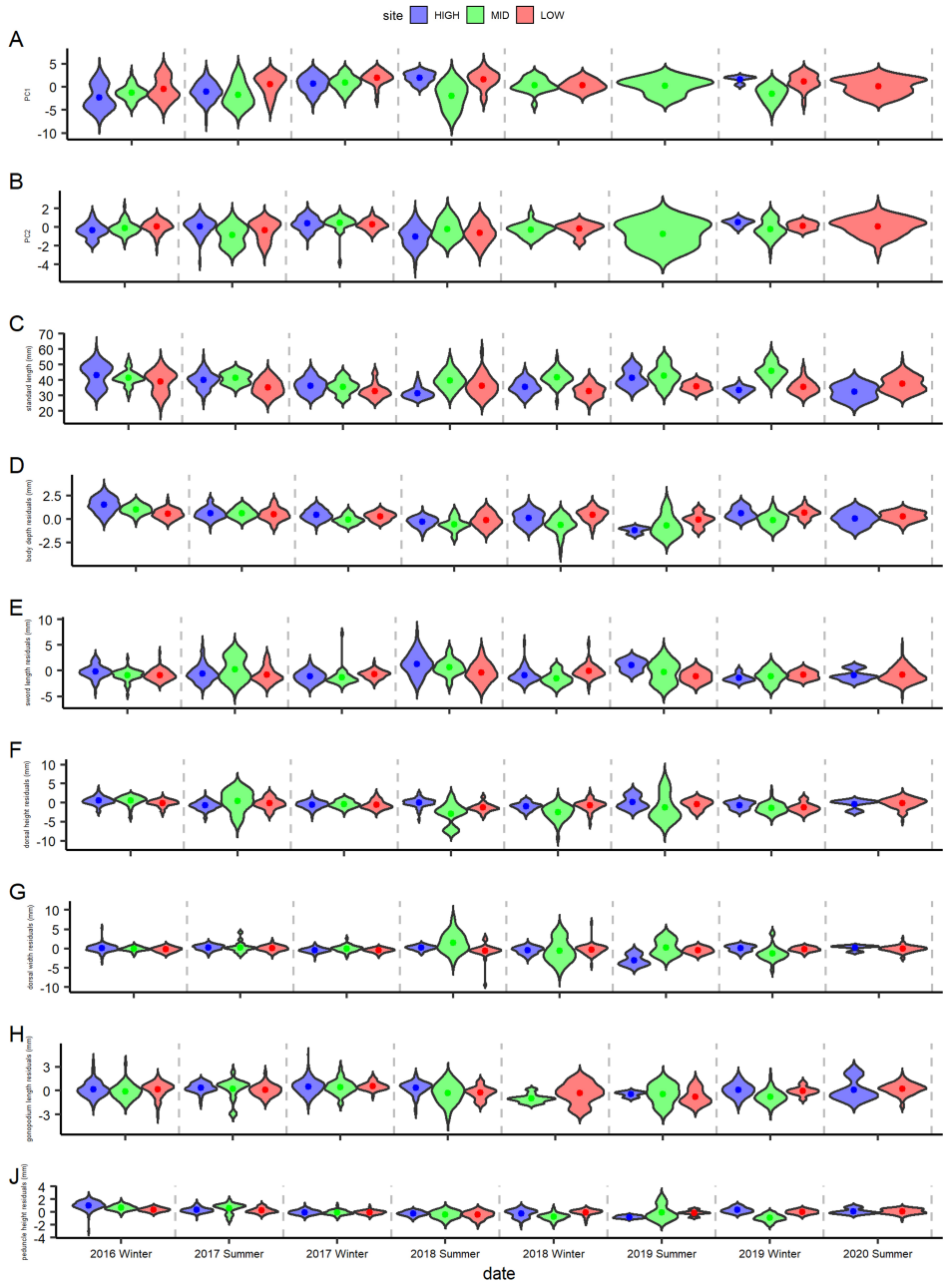

Figure 3.4: Distributions of A) PC1, B) PC2, C) standard length, D) body depth residuals, E) sword length residuals, F) dorsal height residuals, G) dorsal width residuals, H) gonopodium length residuals, and J) peduncle height residuals in the experimental, early generation hybrid populations among all collection dates. Colored points represent population means.

3.4.3 Morphological trait distributions

I conducted one-way ANOVAs followed by post-hoc Tukey HSD test to determine if residual trait values significantly differed between hybrid and parental populations. For PC1 ($F(2,936) = 16.58$, $p < 0.001$) all hybrid sites were significantly different from one another, yet there were no significant differences between populations for PC2 (Figure 3.5.A,B). Individuals from the intermediate hybrid population had the greatest standard length (Figure 3.5.C, $F(2,1388) = 69.91$, $p < 0.001$) trait values among all hybrid sites, followed by the highland and lowland populations, respectively. The intermediate experimental hybrid population expressed smaller body depth ($F(2,1386) = 27.06$, $p < 0.001$), dorsal height ($F(2,1343) = 12.98$, $p < 0.001$), gonopodium length ($F(2,1066) = 37.02$, $p < 0.001$), and peduncle height ($F(2,1383) = 19.89$, $p < 0.001$) residual trait values relative to the lowland and highland hybrid populations. Sword length and dorsal width residual trait values did not significantly differ between populations.

3.4.4 Hybrid index

We calculated the average genome-wide hybrid index for each site for the first 2 years of the experiment and fit linear model with hybrid index as the response variable and time as the predictor variable to determine if hybrid index was significantly increasing or decreasing over time. While no sites significantly differed from a slope of zero, each site trended in the direction that would be expected if ecological selection was driving genome-wide changes in ancestry (Figure 3.6). Over time, the hybrid index in 71% of the replicates in the highland hybrid population are trending towards *X. malinche* ancestry and in the lowland hybrid population, the hybrid index is trending towards *X. birchmanni* ancestry in 63% of the replicates. In the lowland hybrid population, sword length was positively associated with hybrid index ($R^2 = 0.03$, $p = 0.014$). Individuals with more *X. malinche* genome-wide ancestry expressed larger swords. Among replicates, 75% had positive trends.

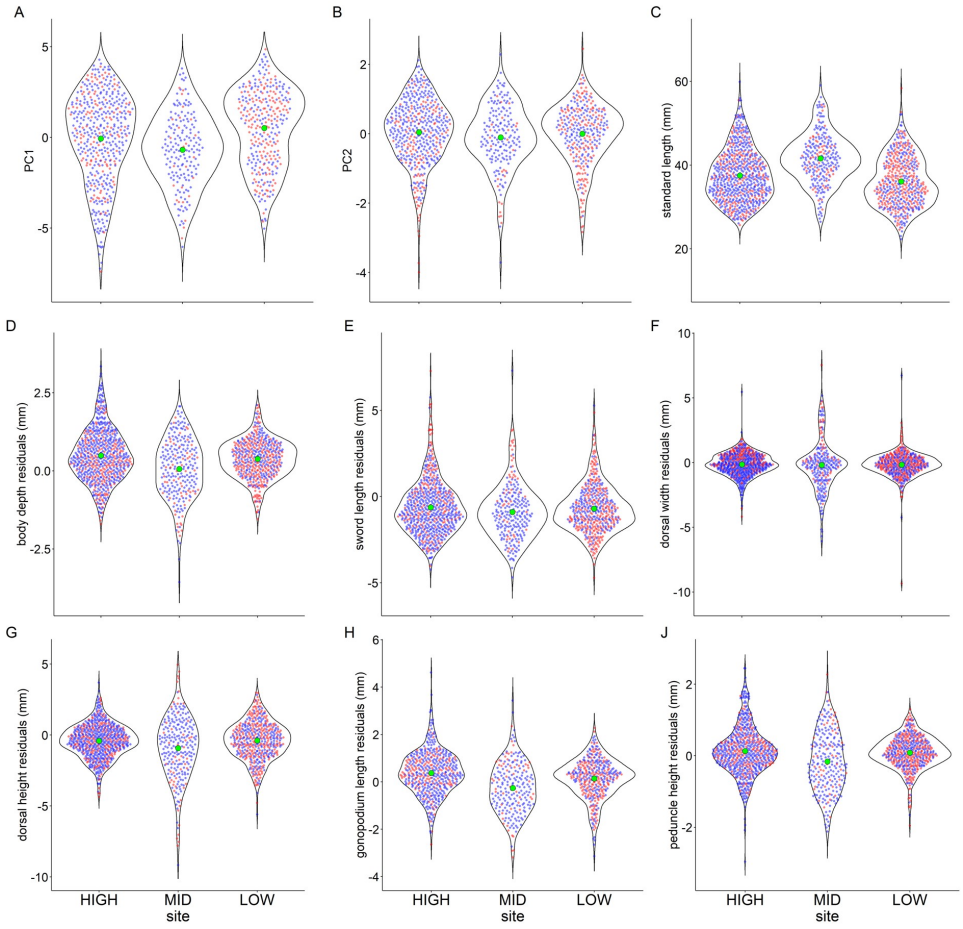

Figure 3.5: Distributions of A) PC1, B) PC2, C) standard length, D) body depth residuals, E) sword length residuals, F) dorsal height residuals, G) dorsal width residuals, H) gonopodium length residuals, and J) peduncle height residuals for the HIGH, MID, and LOW populations among all collection dates combined. Red and blue points represent individuals collected in summer and winter, respectively. Green point represents population mean.

3.4.5 Lifespan

Individuals from the intermediate hybrid population on average had the greatest lifespan of 200 days followed by the highland (167 days) and the lowland population (91 days; Table 3.3, Figure 3.7). The lowland hybrid population had the greatest proportion of individuals that did not survive at least one full season (0.68), whereas the proportion in the intermediate and highland sites were each approximately 0.50. Moreover, as the density of males increased, the lifespan decreased in the intermediate ($R^2 = 0.23$, $p = 0.0013$) and lowland ($R^2 = 0.11$, $p = 0.012$) hybrid populations.

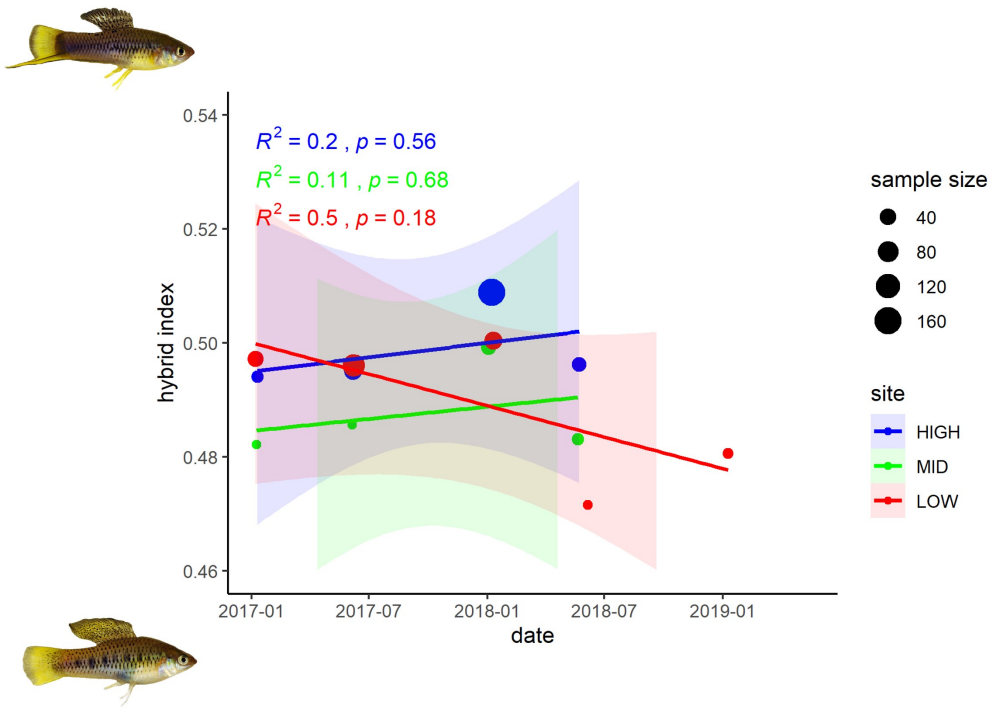

Figure 3.6: Change in hybrid index (i.e. proportion of *X. malinche* ancestry genome-wide) across the three hybrid populations over time.

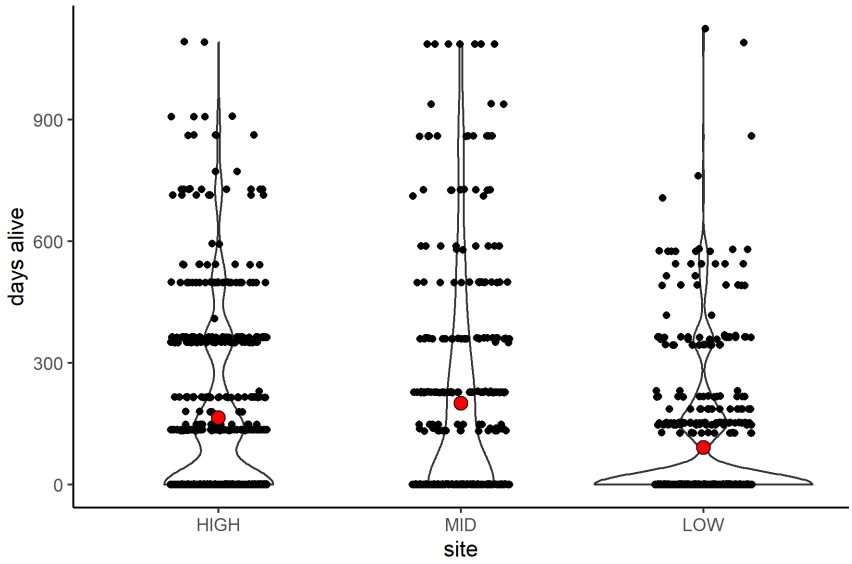

Figure 3.7: Lifespan for individuals among the experimental, early generation hybrid populations.

Longevity	Site		
	LOW	MID	HIGH
< 1 season	348	172	262
>= 1 season	164	170	280
mean (days)	91	201	167

Table 3.3: Number of individuals that were only collected once (i.e. less than one season) versus individuals that survived at least one full season.

3.4.6 Growth rates

I calculated the growth rates of all continuous traits for individuals from each hybrid population that had survived at least one full season (i.e. timespan between collection dates; $\sim$6 months). The intermediate population had larger growth rates for standard length ($F(2,439) = 29.649$, $p < 0.001$), body depth ($F(2,326) = 23.701$, $p < 0.001$), dorsal width ($F(2,490) = 57.567$, $p < 0.001$), dorsal height ($F(2,498) = 31.448$, $p < 0.001$), and peduncle height ($F(2,396) = 21.567$, $p < 0.001$) relative to the highland and lowland populations. Unexpectedly, the lowland hybrid population had the greatest growth rate in sword length followed by the intermediate and highland populations, respectively (Figure B.7.C, $F(2,511) = 8.8398$, $p < 0.001$).

3.5 Discussion

In this study, hybrid phenotypes and genotypes were measured over the course of five years in controlled experimental stock tanks. The primary aim was to monitor hybrid evolution in response to different thermal environments during the incipient stages of evolution. Our primary findings reveal there is more variation in morphology between replicates within hybrid populations than between hybrid populations at different sites. We expect this pattern when ecological selection is not the primary force governing morphological evolution. If selection were primarily responsible, then we would expect two replicates chosen at random within sites to share more principal component space than would any two sites chosen at random. Instead, genetic drift or arbitrary runaway selection may have a greater influence in early generation hybrids with relatively small effective population sizes [60]. However, the lowland population shared slightly more principal component

space with the parental *X. birchmanni* population than did the highland population. While negligible, this pattern is consistent with ecological selection whereby morphology in the lowland hybrid site is evolving to be more like the lowland parental species. The amount of variation in morphology among replicates was relatively the same (i.e. each replicate has a similar ellipse area for PC1 and PC2). This suggests, elevation and thermal environment do not strongly influence the degree of variation present within populations.

In the highland and intermediate hybrid populations, individuals that were collected in winter expressed smaller swords than individuals collected in the summer. Because seasonal fluctuations in water temperature act as a selective force in this system, we expect mature males collected in the summer, but were born the season before and have already experienced winter thermal extremes to skew towards *X. malinche* phenotypes since *X. malinche* are locally adapted to cooler water temperatures [19]. We expect the opposite pattern during winter collections; mature males collected in the winter that were born the season before should skew towards *X. birchmanni* phenotypes since *X. birchmanni* are locally adapted to warmer water temperatures [19].

The intermediate population was the only population to exhibit an increasing trend in trait variance over time, or to have overall greater trait variance relative to the highland and lowland sites. This may be due to weakened constraints at intermediate regions among a hybrid zone transect relative to populations situated closer to pure parental populations. If hybrid populations located closer to either parental species are under greater ecological selective pressure to express traits similar to the parental species it shares a neighboring habitat with, then we would expect these populations to have decreased trait variance as directional selection, followed by stabilizing selection constrains divergence from the optimal trait mean [36].

While no sites had a genome-wide index that significantly differed from a slope of zero over time, each site trended in the direction that would be expected if ecological selection was driving genome-wide changes in ancestry. In the highland hybrid population, there is a positive coefficient that trends towards *X. malinche* ancestry. Conversely, the lowland population exhibits a negative coefficient that trends towards *X. birchmanni* ancestry. The intermediate population, which has

a similar thermal water temperature profile as the highland population also exhibits a positive coefficient towards the highland species, *X. malinche*. If ecological selection takes precedence in driving hybrid genotypic evolution, then we would expect to observe these patterns. However, because we were only able to sequence and genotype the first two years of the experiment, we are limited on what we can conclude with respect to genotypic evolution in early generation hybrids. Sequencing each individual from our dataset would provide insight into how the hybrid index evolve past the first few generations of hybridization and within more admixed genomes where recombination has had time to resolve some genetic incompatibilities [61].

Interestingly, in the lowland hybrid population, sword length was associated with a *X. birchmanni*-like hybrid index in six of the eight stock tank replicates. The sword is a trait expressed in the parental species *X. malinche* but has been lost in *X. birchmanni* [21]. We expect individuals with a greater hybrid index (i.e. more *X. malinche* ancestry) to harbor larger swords. Surprisingly, no other traits were correlated with hybrid index. This suggests phenotypes do not necessarily correspond to genome-wide ancestry patterns, but instead depend on genotypes at localized genomic regions responsible for controlling trait variability, which don't necessarily correlate with hybrid index.

Individuals from the intermediate hybrid population on average had the greatest lifespan followed by the highland and the lowland population, respectively. Moreover, they had larger growth rates for several morphological traits relative to the highland and lowland populations. This finding supports the model that hybrids perform best at intermediate environments relative to the environment of either parental species [12].

An important caveat to note about our study is that selection can only act on the variation present. Since our initial populations were bottlenecked, the variation for which selection can act upon is reduced from the start. This alone can constrain how hybrids subsequently evolve [62]. Additionally, because of limited resources and feasibility, we do not measure realized fitness values or female preference during this study. Still, we are able to gain a tremendous amount of insight into patterns of early generation hybrid evolution from this dataset. As such, the results from this

study support the scenario that variation can be explained more within sites than between sites in admixed populations.

4. UNDERDOMINANCE, INCOMPATIBILITIES, AND PHYSICAL BARRIERS LIMIT ADAPTIVE INTROGRESSION OF LOCI ASSOCIATED WITH THERMAL TOLERANCE IN SWORDTAILS*

4.1 Abstract

Hybrid zones provide conduits for which alleles can introgress from one population to another. This process has the propensity to lead to adaptation in the recipient population if the migrant alleles confer a fitness advantage. Yet studies monitoring changes in allele frequencies at genomic regions associated with ecologically relevant traits over time in hybrid populations is rare. The goal of this project was twofold. First, we sought to identify the genetic architecture of a trait associated with thermal tolerance, critical thermal maximum (CT_{max}). We conducted quantitative trait locus (QTL) mapping on CT_{max} and found a region on chromosome 22 to explain nearly 7% of the variation in hybrids. Moreover, this genomic region exhibited underdominance, whereby heterozygous genotypes had markedly reduced CT_{max} values. Next, we tracked changes in ancestry and patterns of introgression at genomic regions associated with this trait in natural hybrid populations. We repeated our analyses at genomic regions harboring heat-shock protein genes and found ancestry was significantly increasing towards *X. malinche* ancestry with respect to the genome-wide average for a small portion of these genes (<5%) in a lowland, *X. birchmanni*-skewed natural hybrid zone. Geographic cline analysis revealed several loci considered center and/or width outliers among two independent river drainages, one among which expressed a bias for introgression of *X. malinche* alleles downstream. Together, these findings draw attention to the difficulties of discerning the impact hybridization has on trait evolution and adaptive introgression.

*Part of the data reported in this chapter is reprinted with permission from "Genomic insights into variation in thermotolerance between hybridizing swordtail fishes" by Payne, Cheyenne, Richard Bovio, Daniel L. Powell, Theresa R. Gunn, Shreya M. Banerjee, Victoria Grant, Gil G. Rosenthal, and Molly Schumer. 2022. *Molecular Ecology*, Copyright 2022 by John Wiley and Sons.

4.2 Introduction

Hybridization is increasingly recognized as an important evolutionary process [1]. It is often facilitated when genetically diverged species are distributed along an environmental gradient such that a conduit for secondary contact and genetic exchange becomes established. Hybrid zone formation consequently enables alleles to introgress along an environmental gradient. Under conditions like these, hybridization serves as an important evolutionary mechanism by which diversity and adaptation can arise in a process called adaptive introgression [1, 2].

Uncovering the genetic underpinnings of adaptive traits and how they evolve has been a long-standing goal in evolutionary biology. This in turn requires characterizing the genetic architecture of ecologically relevant traits and monitoring changes in allele frequency over time. However, rarely do studies monitor how the allelic frequencies change over many generations typically due to the difficulty of mapping and maintaining experimental hybrid populations in addition to sampling from natural hybrid populations regularly.

A recent, and increasingly recognized source of variation for adaptation to occur arises when closely related species hybridize [2]. If a conduit for genetic exchange exists across closely related species, then alleles can introgress from one population to another [63]. This can facilitate adaptation on a shorter timescale than novel mutation alone [64, 65].

An ecologically relevant trait commonly used to assess how organisms respond to rising global temperatures is thermal tolerance [66, 67, 68]. While several traits encompass an organism's thermal tolerance, I focus on critical thermal maximum (CT_{max}), the maximum temperature a fish can withstand before it loses equilibrium and can no longer maintain balance [69, 70]. Critical thermal maxima is a suitable metric since it is often used in studies to measure the thermal response of organisms [71]. This makes our results generally comparable between studies. Moreover, it serves as a predictor of the temperature individuals can withstand in the short term and can be measured nonlethally. An important note is that CT_{max} measures acute heat stress over the short term, while in the wild, shifts in thermal habitat occur at a slower rate. Because of this, CT_{max} is limited in the predictive power of an organism's actual thermal tolerance in the wild.

Previous work in the *birchmanni-malinche* system has demonstrated the distribution of *X. birchmanni* and *X. malinche* are in part determined by their thermal environment. *X. malinche* lives in cooler (7-25 °C) streams at high elevations, while *X. birchmanni* lives in warmer streams (15-35 °C) at low elevations [19]. As a consequence of local adaption to their respective thermal habitats, CT_{max} and heat-shock protein (*hsp*) gene expression vary between *X. birchmanni* and *X. malinche* in natural populations [19]. Across independent river drainages, there is a gradient of low CT_{max} in the highlands and high CT_{max} in the lowlands. Moreover, when reared in a common garden, *X. birchmanni* maintained thermal equilibrium at warmer water temperatures than *X. malinche*. Together, these results suggest that water temperature, and perhaps more importantly, fluctuations in water temperature are likely primary driving forces underlying individual fitness with respect to environment in this system. However, it's unclear exactly which regions of the genome contribute to interspecific differences in thermal tolerance or how effective these regions introgress and serve as a mechanism for adaptation. Ultimately, this study set the stage for exploring the genetic architecture of a trait under ecological selection and the introgression, or lack thereof, of species-specific alleles across hybrid populations.

To this end, I performed quantitative trait locus (QTL) mapping on intercrossed artificial hybrids generated in semi-natural mesocosm stock tanks to identify regions of the genome associated with thermal tolerance. While the exact loci regulating thermal tolerance are unknow, it nonetheless is in part driven by *hsp* genes [72, 73, 74, 19].

Next, I leverage low-coverage whole genome data from two independent hybrid populations we have sampled for nearly a decade to test how frequencies at *hsp* loci and the QTL for CT_{max} change over time. I expect one of two putative scenarios to occur. Since water temperature is increasing in both populations, selection may favor individuals harboring *X. birchmanni* alleles capable of withstanding a greater CT_{max} across both hybrid zones. Alternatively, selection for the major parent genome may outweigh selection at CT_{max} loci. Ancestry should become more *X. malinche*-like in the highland and more *X. birchmanni*-like in the lowland.

Finally, I use geographic cline analysis across two independent river drainages to detect if these

genomic regions deviated significantly from genome wide expectations and show patterns of introgression [3, 75]. This analysis models the change in allele frequency over a spatial gradient using a sigmoid function (and estimates the cline center and width [13, 76, 77]. Comparing the centers and the widths of multiple loci, along with the average genome ancestry through the geographic gradient, allows the detection of regions with differential introgression [78].

We detected a QTL on the chromosome 22 explaining 6.9% of variation in CT_{max} in hybrids [79]. Individuals that are heterozygous for ancestry at the QTL have reduced CT_{max}. In addition, there was an interaction between chromosome 22 and chromosome 15. Together the combined additive and interaction effects explain up to 14.8% of the variation in CT_{max}. Most *hsp* loci fell outside the genome-wide 95% CI interval, however ancestry did not skew towards one parent more than another in any population. The geographic cline analysis yielded several loci considered center and/or width outliers among the two river drainages.

4.3 Methods

4.3.1 Thermal tolerance trials

Artificial hybrid fry (n=240) from experimental mesocosm stock tanks described in chapter 3 were collected in summer 2019 and reared in a common garden environment with respect to water temperature at the CICHAZ field station. Due to the difficulty of raising a sufficient number of individuals in common garden conditions, our mapping population included individuals from all sites ranging from F2-F4 generations. To measure the variation in CT_{max} in artificial hybrids, thermal tolerance trials were conducted in January 2020 once the fish had developed to maturity. Briefly, we tested CT_{max} by placing the fish (eight per test) in a pot of water set to the same temperature as the stock tanks (20 +- 2 °C) and steadily increasing the temperature at a rate of 0.3 °C per minute until the fish lost equilibrium [70]. As soon as the fish lost equilibrium, I recorded the time and temperature of initial loss of equilibrium and then immediately removed the individual and placed it in an aerated recovery tank. In addition to CT_{max} time and temperature, we recorded the sex and the site and tank of origin for each individual. No mortalities occurred during this assay.

After the fish had recovered equilibrium, they were anesthetized in MS-222, photographed and fin clipped. All fish were placed back in their respective common garden stock tank at the end of the procedures. Trials followed procedures approved in Texas A&M IACUC protocol #2020-081. In addition, we tested a subset of individuals from each stock tank site that were not acclimated to the common garden water temperature to serve as a baseline for comparison. Baseline fish were taken from their respective site, transported back to the CICHAZ field station and immediately tested for CT_{max}. I used an ANOVA to test for the effects of site on CT_{max} and a post-hoc Tukey HSD test to determine pairwise differences.

4.3.2 DNA extraction and library preparation

I applied the multiplex shotgun genotyping (MSG) approach for genome-wide genotyping of hybrids [24, 58]. DNA was extracted using an Agencourt DNAdvance bead-based purification kit (Beckman Coulter) following the manufacturer's instructions except we used half reactions and diluted to a 2.5 ng/μL. Each sample was sheared using the Illumina Tagment DNA TDE1 Enzyme and Buffer Kits, and amplified in a dual-index PCR reaction with the conditions outlined in Table 4.1.

Step	Temp	Time
1	68°C	3 min – Extend Tn5 transposon ends
2	95°C	30 sec
3	95°C	10 sec
4	55°C	30 sec
4	68°C	30 sec
5		Cycle to Step 3 12 times
6	68°C	5 minutes (final extension)
7	4°C	Hold

Table 4.1: Polymerase chain reaction conditions for Illumina tagmentation library preparation.

PCR products were pooled and purified with 18% SPRI magnetic beads (Beckman Coulter). The library was quantified using a Qubit 3.0 fluorometer, run on a Agilent 4200 TapeStation to

assess the library size distribution, and finally sequenced on an Illumina Genome Analyzer (HiSeq 4000 PE 150 bp reads).

4.3.3 Artificial hybrid QTL mapping sample sequencing and genotyping

Raw sequence reads were mapped to each of the parental species' genome using the program ancestryinfer [80]. This program records the number of reads matching each parental allele at ancestry-informative sites between each parental genome. This information is subsequently used in the program AncestryHMM [81]. This program is a local ancestry inference program that uses a hidden Markov model to generate posterior probabilities of ancestry states genome-wide. This analysis yields approximately 700,000 ancestry informative posterior probabilities. For downstream analysis, we converted posterior probability estimates greater than 0.9 for any particular genotype into hard genotype calls. Estimates less than 0.9 were converted to NAs.

4.3.4 CT_{max} QTL mapping analysis

We used QTL mapping to identify regions of the genome that are associated with variation in thermotolerance. We performed QTL mapping with R/qtl [82] to identify associations between genotypes at ancestry-informative markers across the genome and the CT_{max} phenotype. For computational efficiency, we thinned to approximately 30,000 markers. Loci where less than 80% of samples were genotyped were filtered. Samples that had greater than 25% of markers genotyped as NA were filtered. Lastly, markers were evaluated for segregation distortion. After filtering, 144 samples and 29,042 markers remained.

To select an appropriate model for mapping in R/qtl, we used the R step function to calculate AIC for models incorporating a suite of possible covariates. These included tank origin, hybrid index, heterozygosity, and sex. The model with the lowest AIC score yielded tank origin as a significant covariate. We retained tank origin and hybrid index as covariates to the model. While hybrid index was not included in the model with the lowest AIC score, we have experienced artifacts when it is left out. For this reason, we included it in downstream analyses. We performed a genome-wide scan using the scanone function. We used a single-QTL model using the Haley-

Knott regression method with tank origin and hybrid index as covariates [83]. The 5% (LOD = 4.72) and 10% (LOD = 4.33) LOD significance thresholds (p = 0.05) were estimated based on 1,000 permutations where CT_{max} phenotypes were shuffled onto genotypes and a QTL scan conducted 1,000 times to create a null distribution of associations expected by chance. To identify interacting QTL, we performed a second scan using the same method, but added genotypes at the chromosome 22 QTL peak as an interaction term in the model (5% and 10% LOD thresholds of 9.63 and 8.96, respectively).

To obtain estimates of the effect size of the detected QTL, we used the drop-one-term analysis from fitting a multiple QTL model with the R/qtl function `fitqtl`. Because of the Beavis effect, whereby effect size estimates are often inflated due to low statistical power [84], we also performed ABC simulations to obtain a confidence interval of possible effect sizes for the main effect QTL on chromosome 22.

4.3.5 Heat-shock protein and CT_{max} QTL evolution through time

The *X. birchmanni* reference genome is not annotated. Therefore, to obtain the coordinates of *hsp* genes, I searched for 'heat shock protein' AND 'xiphophorus maculatus' in NCBI. This resulted in 95 *hsp* genes. I blasted these regions against the *X. birchmanni* reference genome, manually removed false positives, and extracted the *X. birchmanni* genomic coordinates. I used a custom script to calculate the allele frequency for each *hsp* gene among two independent hybrid populations, Acuapa (ACUA; 450 m, $\sim$25% *X. malinche* ancestry), and Tlatemaco (TLMC; 440 m; $\sim$75% *X. malinche* ancestry), across time. We have sampled and sequenced whole genome data from ACUA between 2006 to 2018, and from TLMC between 2009 to 2017. Together, these two independent hybrid zones provide a time series dataset whereby the change in ancestry can be monitored over time.

I fit a linear model to change in ancestry of each *hsp* gene, the CT_{max} region, and the genome-wide hybrid index for both populations and extracted the coefficients and FDR adjusted p-values. I compare the regression coefficients of the genome-wide hybrid index to each *hsp* locus with an ANOVA to determine if *hsp* loci are significantly changing with respect to the genome-wide

average.

For each collection across both rivers, I examine the number of *hsp* loci that have skewed ancestry towards either of the parental species (i.e. ancestry at *hsp* locus falls outside 95% confidence interval of the genome-wide hybrid index). I use Wilcoxon signed rank test to determine if loci are skewed in one direction or another. I repeat this same analysis in the experimental stock tanks to identify patterns of selection as well. The most admixed individuals we have sampled and sequenced to date are F3 and a few F4 individuals from the highland and lowland site.

4.3.6 Genomic cline analysis

I used the Metropolis–Hastings algorithm in the R package hzar [85], to fit the whole genome hybrid index, the allele frequencies of each *hsp* locus, as well as the $\mathrm{CT}_{\max}$ QTL region to a geographic cline model [77]. To fit a cline, I used three models [86]. Model I estimated only the center γ and width ω of the cline, assumed no tails (θ the rate in which the cline tail decays, and β, the size of the cline, fixed to one and zero, respectively), and included fixed ends (P_{min} and $P_{\max}$ fixed to zero and one, respectively). Model II estimated γ, ω, and P_{min} and $P_{\max}$ from the data and assumed no tails. Model III was the same as Model II but with tail estimates and β allowed to vary. Each model parameter was estimated using three independent chain runs using 100,000 MCMC steps after a burn-in of 10,000 steps. The model with the lowest Akaike information criterion (AIC) was selected. For each locus, hzar computes the estimated value and 95% confidence intervals for each parameter, as well as the whole genome hybrid index from the Huazalingo and Pochula rivers. Centers and widths for any locus were considered coincident (same γ) and concordant (same ω) if their parameters overlap with the 95% confidence intervals of the genome-wide hybrid index. Loci were considered outliers if they did not overlap with the 95% confidence intervals of the genome-wide hybrid index cline γ and ω.

4.4 Results

4.4.1 Thermal tolerance trials

The common garden tanks the fish were reared in are below ground 500 L tanks. The water temperature among the common garden tanks were consistent with one another throughout the course of the experiment (Figure C.1). Data from underwater temperature loggers recorded water temperatures ranging from 11.334 to 27.075 °C. Critical thermal maxima values ranged from 32.118 to 37.824 °C, which corresponds with CT_{max} ranges published in previous studies (Figure 4.1; [19]). Origin site of each sample had a significant effect on CT_{max} values ($F = 59.62$, $p < 0.001$). Each site significantly differed from one another (Tukey HSD, p-values < 0.001). Hybrids originating from the highland site had the greatest CT_{max} values, following by the lowland and intermediate population, respectively. Critical thermal maxima values for unacclimated samples from the high and lowland population that were immediately tested were not significantly different, however samples from the lowland population trended towards higher CT_{max} values (Two sample t-test; $t = 1.6508$, $p = 0.053$).

4.4.2 CT_{max} QTL mapping

We detected a single QTL associated with CT_{max} at a 10% false discovery rate threshold (Figure 4.2.A). The interval of the QTL spans ∼2.5 Mbs on chromosome 22. Surprisingly, the QTL was not associated with species-specific differences in CT_{max}. Instead, individuals heterozygous in this region were associated with an average reduction in CT_{max} of 0.3 °C (Figure 4.3.A). The observed pattern is consistent with underdominance. Individuals homozygous for *X. birchmanni* or *X. malinche* ancestry at the interval on chromosome 22 exhibit similar CT_{max} values, while individuals heterozygous in ancestry have reduced CT_{max} on average. We estimate the effect size of this QTL to explain approximately 6.9% of the variation.

To determine if there were any interacting QTL, we repeated the same analysis but added the genotype at the chromosome 22 QTL as an interaction term. We found a second QTL associated with CT_{max} at a 10% false discovery rate threshold (Figure 4.2.B). The interval of this QTL spans

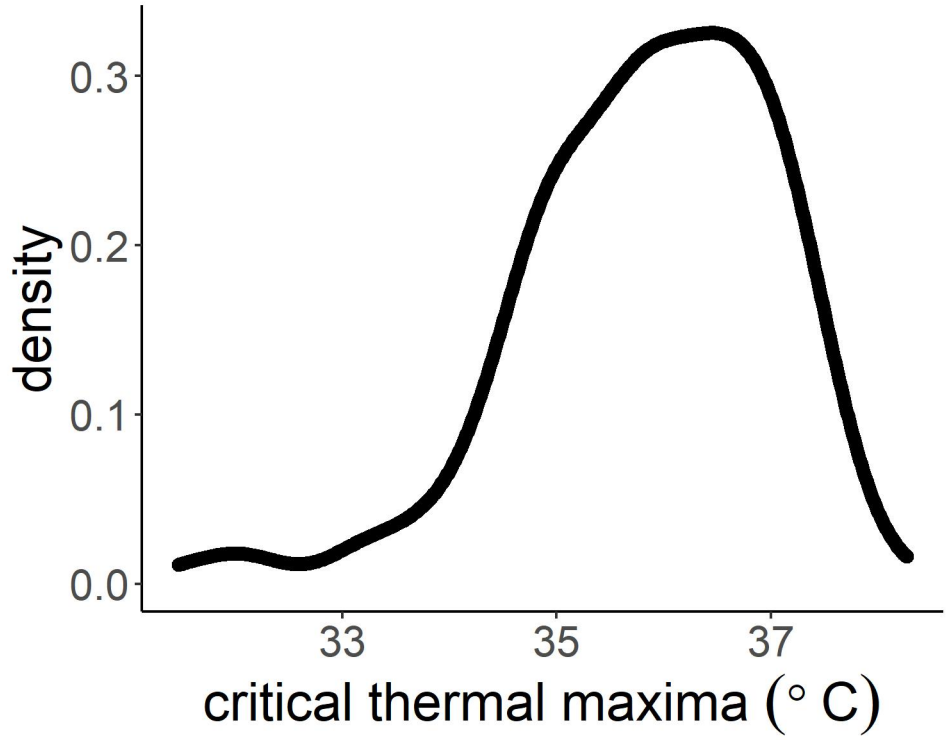

Figure 4.1: Variation in CT_{max} among all intercrossed hybrids used for QTL analysis.

$\sim$2.1 Mbs on chromosome 15. Individuals heterozygous at the chromosome 22 QTL and either heterozygous or homozygous *X. malinche* at the chromosome 15 QTL have reduced CT_{max} (-0.4 °C) on average, but individuals homozygous *X. birchmanni* at the chromosome 15 QTL have an inflated CT_{max} (+0.5 °C, Figure 4.3.B). We estimate that the combined additive and interaction effects of the chromosome 22 and 15 QTL explain $\sim$14.8% of the total variation in CT_{max} in the hybrids.

4.4.3 Ancestry patterns in natural hybrid populations at regions implicated in thermal tolerance

A total of 95 *hsp* loci were examined in addition to the CT_{max} locus. The genomic region of the CT_{max} locus did not overlap with the genomic regions of any *hsp* genes. A majority of *hsp* loci fell outside the 95% CI of the genome-wide hybrid index for each collection date across both rivers (Table 4.2). However, the number of loci that skewed towards either of the parental species

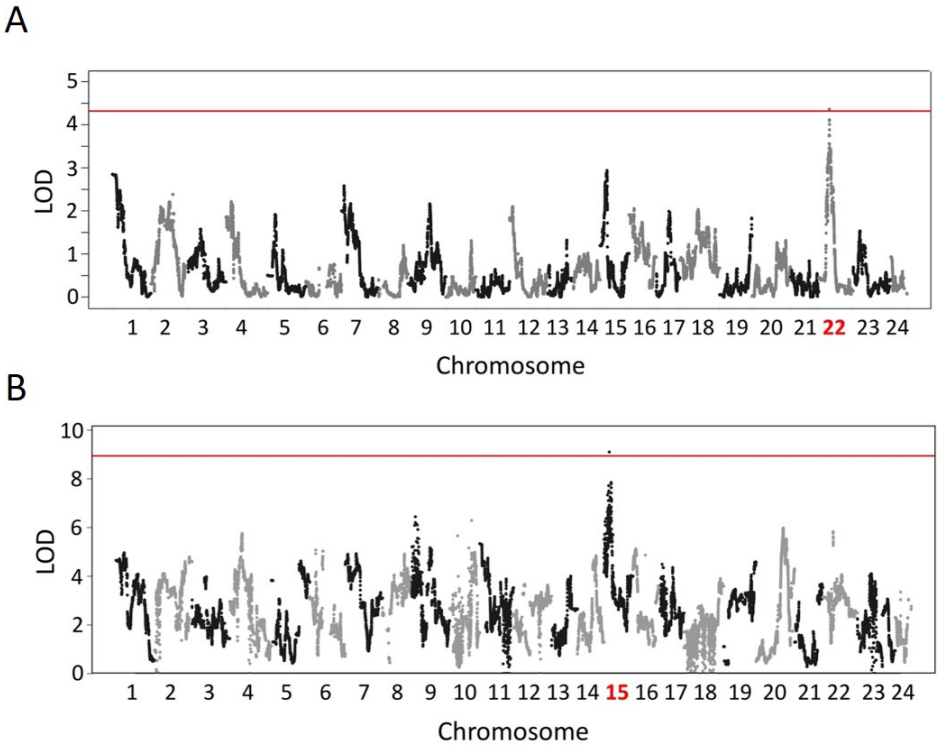

Figure 4.2: A) QTL associated with CT_{max} detected on chromosome 22 at a 10% false discovery rate threshold (red line). B) A second QTL scan that included the region on chromosome 22 as an interaction term reveals a putative interacting region on chromosome 15 at a 10% false discovery rate threshold (red line).

at any collection dates did not differ significantly for either population (one-tailed Wilcoxon sign test; p-values > 0.05).

I repeated the same analysis in the experimental stock tanks to identify patterns of selection as well. The most admixed individuals we have sampled and sequenced to date are F3 and a few F4 individuals from the highland and lowland site. While the number of loci that skewed towards either parental species was not significant for either stock tank site, loci trended towards the majority parent ancestry (i.e. *X. malinche* ancestry) in the highland site (25 loci *X. birchmanni*-skew, 35 loci no skew, and 36 loci *X. malinche*-skew; one-tailed Wilcoxon test, $p = 0.052$; Figure 4.4). There was a significant difference in the genome-wide hybrid index between the lowland and highland sites (one-tailed Student's t-test, $t = 133.92$, $p < 0.001$). In addition, there was significant

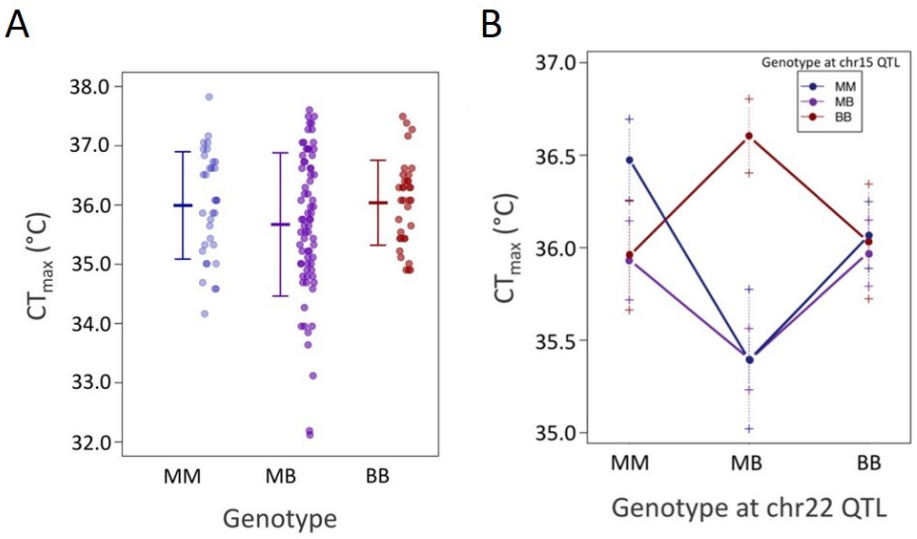

Figure 4.3: A) CT_{max} as a function of genotype at the peak associated with CT_{max} on chromosome 22 among common garden hybrids. Individuals that were heterozygous at this marker had a 0.3 C reduction in CT_{max} compared to individuals that were homozygous for either of the parental species. Points represent the CT_{max} of individual hybrids. B) Interaction between the peak associated marker of the chromosome 22 QTL (on the x-axis) and the peak associated marker of the chromosome 15 QTL (in the legend). Bars and whiskers show the mean and 1 standard error.

difference in the number of *hsp* loci that had a greater allele frequency in the highland population compared to the lowland population (75 loci have greater *X. malinche* ancestry in highland site; two-tailed Wilcoxon signed rank test, $p < 0.001$; Figure 4.4).

I fit a linear model of the change in ancestry of each *hsp* gene, the CT_{max} region, and the genome-wide hybrid index for both natural populations (Figure 4.5). I use the regression coefficients and associated p-values from fitted models to test whether the ancestry is significantly increasing or decreasing over time towards or away the genome-wide average. After correcting for multiple comparisons there were 7 loci that deviate significantly from the genome-wide average coefficient in the ACUA population (*hspb7, LOC102225424, LOC102222412, LOC102221741, LOC102221189, LOC111609415, LOC102225689*, Figure C.2.A). The rate of change in ancestry is significantly increasing towards *X. malinche* ancestry with respect to the genome-wide average for each of these loci except *LOC111609415, LOC102225689*. For these, the ancestry is decreas-

Population	Year	*X. birchmanni*-skew	No skew	*X. malinche*-skew
TLMC	2009	31	34	31
	2010	39	9	48
	2012	42	4	50
	2013	43	4	49
	2015	42	0	54
	2017	43	0	53
ACUA	2006	39	27	30
	2008	48	1	35
	2013	54	2	40
	2015	53	2	41
	2018	52	0	44

Table 4.2: Distribution of ancestry for each *hsp* locus across all collection dates in the ACUA and TLMC natural hybrid populations.

ing with respect to the genome-wide average. In the TLMC population, there were no loci with regression coefficients that different significantly from the genome-wide expectation after correcting for multiple comparisons (Figure C.2.B).

4.4.4 Geographic cline center and widths deviating from genome wide expectations

Loci considered center outliers have the center of their clines shifted towards one of the parental species more than would be expected by chance. This indicates introgression of an allele from one end of the transect to the other. Clines with steep slopes will have more narrow widths. This indicates selection against hybrid genotypes, while wider widths can be the result of heterozygote advantage or weak selective pressures along the environmental gradient. The center of the Huazalingo genome-wide cline was 775.79 m (lower CI: 761.42 m, upper CI: 800.45 m) and the width was 104.32 m (lower CI: 8.71 m, upper CI: 167.95 m; Figure 4.6). The center of the Pochula genome-wide cline was 425.80 m (lower CI: 407.09 m, upper CI: 442.55 m) and the width was 159.07 m (lower CI: 108.53 m, upper CI: 230.86 m). From the 96 loci analyzed in Huazalingo, the majority show coincident centers, with 1 being non-coincident (*X. birchmanni*-skewed center outlier) and 3 non-concordant (width outliers all with narrower widths; Table 4.3). In Pochula, most loci are also coincident, with 12 being center outliers (9 *X. malinche*-skewed and 3 *X. birch-*

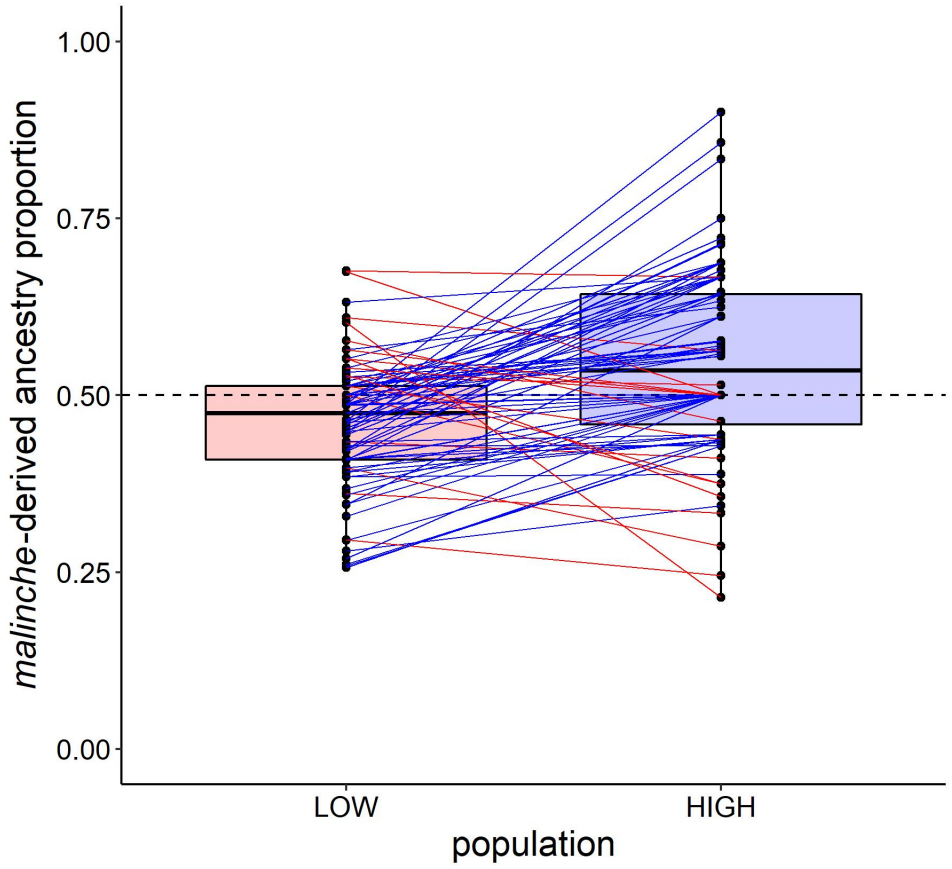

Figure 4.4: Distribution of *hsp* ancestry between the lowland and highland experimental stock tanks. Each point represents the ancestry proportion of a single *hsp* gene. Lines connect the same gene between populations. Blue lines represent a greater ancestry proportion in the highland site and red lines represent a greater ancestry proportion in the lowland site.

manni-skewed) and 9 width outliers, all of which had narrower clines with the exception of one (Table 4.4). There were 2 genes (*hikeshi* and *hspa8*) in the Pochula cline that are considered both center and width outliers and 1 locus (*dnajc16l*) in the Huazalingo cline that was considered both a center and width outlier. The *hikeshi* gene encodes an evolutionarily conserved nuclear transport receptor that mediates heat-shock-induced nuclear import of *hsp70s* [87]. There were no genes that were outliers in both river drainages. Only *hsp* loci were outliers; the CT_{max} QTL did not significantly differ from the genome-wide expectations in either river drainage (Figure 4.6).

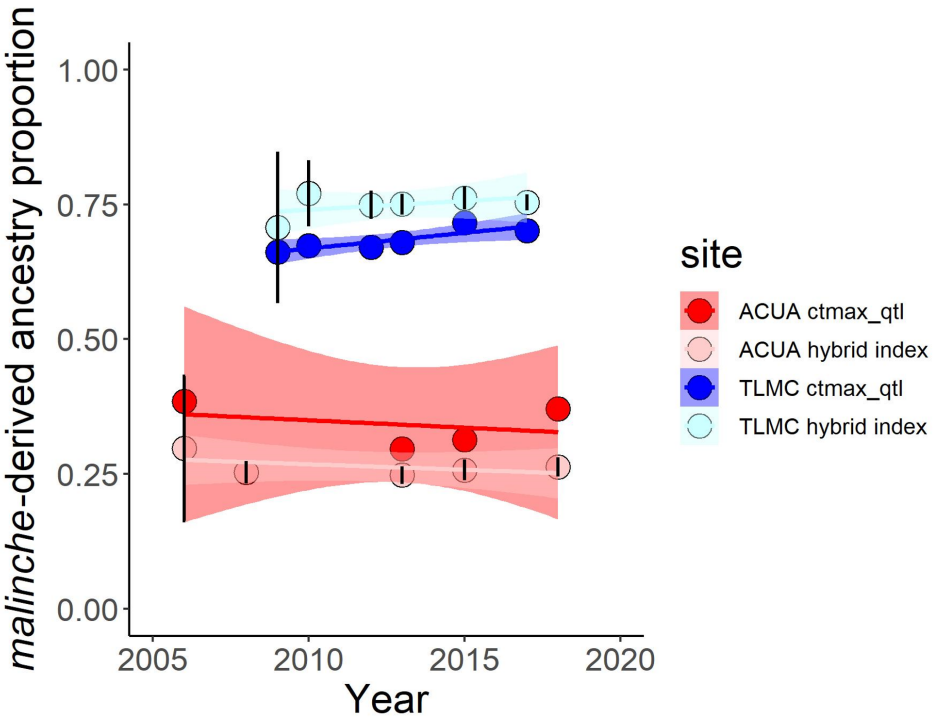

Figure 4.5: Genome-wide hybrid index (light blue and light red) and ancestry proportion at the QTL for CT_{max} (dark blue and dark red) at two natural hybrid populations, TLMC and ACUA. TLMC is skewed towards *X. malinche* ancestry and ACUA is skewed towards *X. birchmanni* ancestry. Trend lines are coefficients of a linear model with ancestry as a function of time.

4.5 Discussion

In this study I made use of several whole genome datasets to explore whether genomic regions considered to be putative targets of selection were introgressing or changing in ancestry over time. While non-significant, baseline CT_{max} trials follow the expected pattern – lowland derived hybrids have greater CT_{max} than highland derived hybrids. These findings are consistent with studies measuring CT_{max} in unacclimated hybrids derived from different elevations (unpublished data). However, common garden CT_{max} trials did not follow the expected pattern. Interestingly, fish derived from the highland site had the greatest CT_{max} values, followed by the lowland and midland site, respectively. One explanation of this pattern is that selection coefficient at loci contributing to variation in CT_{max} is not great enough to cause shifts in ancestry over the course of time we

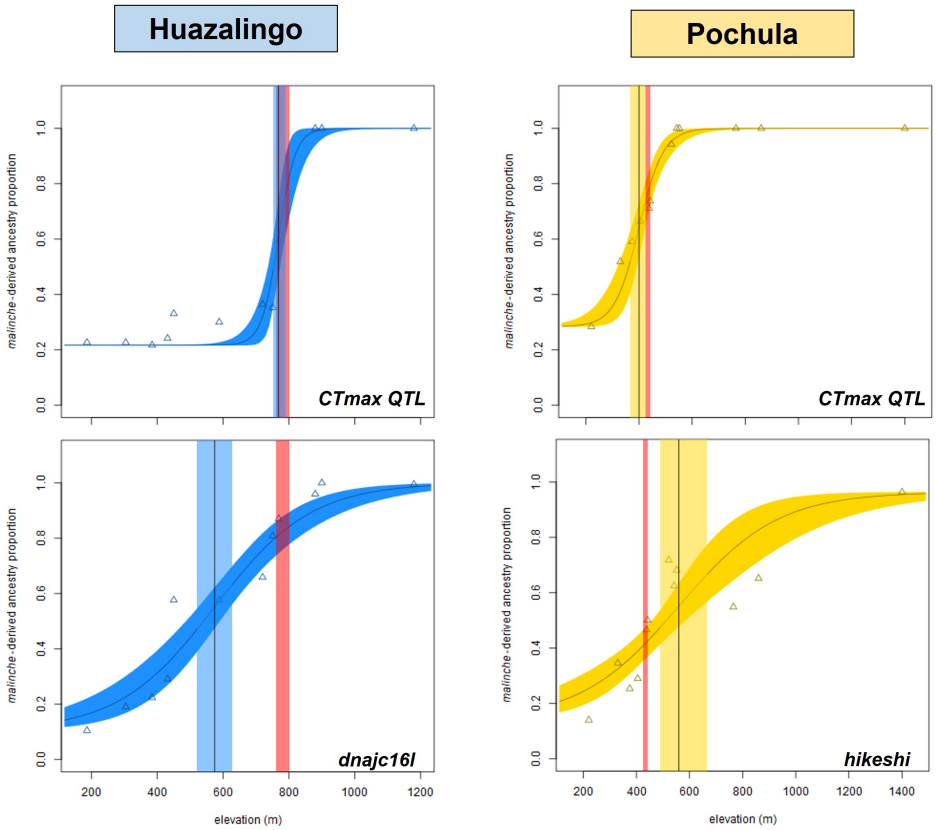

Figure 4.6: Geographic clines for select loci in the Huazalingo and Pochula cline. Loci are specified in the bottom right corner of each cline. Shaded vertical bars represent the 95% confidence interval of the center of the cline for the genome-wide hybrid index (red) and the locus of interest (blue and yellow). Black veritcal line represents the estimated center of the cline for the locus of interest.

have been monitoring. In addition, the effective population size that the fish were collected from was relatively small (∼50 or less), which will naturally decrease the strength of selection. Nevertheless, we expect a single QTL scan to return the QTL with the greatest effect. We expect traits controlled by one or a handful of genes to be revealed as a single large effect QTL during this scan. Whereas polygenic traits should result in a moderate or small effect QTL. Since the single scan yielded a moderate effect (5-15%) and marginally significant QTL (at the 10% FDR significance threshold), this indicates the trait is polygenic and that this would be the only QTL we be able to mine from the dataset; a consequence of being underpowered. The multiple scan reduces power further, though produced a qualitatively similar result (i.e. one peak significant at the 10% FDR

Center outliers	Direction of introgression	Width outliers	Relative width change
dnajc16l	downstream	dnajc16l	wider
		dnajc13	wider
		hspa12b	wider

Table 4.3: Genes with geographic clines that deviate significantly from the genome-wide hybrid index in the Huazalingo cline. Loci in bold are both center and width outliers. Downstream = *X. malinche* alleles introgressing downstream. Upstream = *X. birchmanni* alleles introgressing upstream.

Center outliers	Direction of introgression	Width outliers	Relative width change
LOC102225424*	upstream	hspb1	narrow
LOC102224270	upstream	hspa9	narrow
LOC102222412*	upstream	hspa14	narrow
LOC102221741*	upstream	dnajc21	narrow
LOC102221189*	upstream	hspb11	narrow
LOC111609415*	upstream	usp19	narrow
hspd1	downstream	Sacs	narrow
LOC102225458	downstream	hikeshi	wider
dnajb1b	downstream	hspa8	narrow
hikeshi	upstream		
hspa8	upstream		
dnajc14	upstream		

Table 4.4: Genes with geographic clines that deviate significantly from the genome-wide hybrid index in the Pochula cline. Loci in bold are both center and width outliers. Loci with stars are outliers in the ACUA time series dataset.

threshold). One important limitation to QTL analysis is that it is impossible to determine the exact number of genes that control a polygenic trait. With more power and larger datasets, a close approximation can be made, but ultimately it is not feasible to recover every tiny effect region. A further limitation of our dataset is that we cannot distinguish between true underdominance and pseudo-underdominance generated by closely linked genes [88, 89]. The QTL for CT_{max} exhibits underdominance. Underdominance is an unstable equilibrium and can lead to the fixation of alleles [90, 91, 92], however scenarios where underdominance can be stable do exist, for example, in a spatially extended population by a selection–migration equilibrium [93]. Underdominance

constitutes a barrier to introgression since individuals that harbor heterozygous ancestry at QTL region on chromosome 22 will have markedly reduced thermotolerance. These findings infer the selective pressures driving adaptive introgression of *X. birchmanni* alleles may not be sufficient to offset the cost of hybridization and subsequently lead to higher thermotolerance in highland populations. Such consequences can limit the success of genetic rescue. We also note that we did not focus on cold tolerance in this study since cold tolerance does not significantly differ between parental species reared in a common garden [19]. Regardless, studying the genetic architecture and evolution of cold tolerance may provide useful insight into thermal adaptation in hybrid systems.

Natural hybrids from the ACUA and TLMC populations derive the majority of their genomes from *X. birchmanni* and *X. malinche*, respectively, but both reside at *X. birchmanni* typical elevations. I analyzed how ancestry at *hsp* loci, as well as the CT_{max} QTL, changes over time relative to the genome-wide average ancestry in two independent hybrid populations. While a majority of *hsp* genes had mean ancestry that fell outside the hybrid index 95% CI, in neither population did they skew towards one parent ancestry over another, inferring selection is not driving ancestry changes in the same direction across all loci. However, 7 of 96 loci deviated significantly in the rate of change in ancestry over time in the ACUA, but not TLMC, population. A majority of these were increasing in *X. malinche* ancestry over time (5/7). ACUA is a lowland, *X. birchmanni*-skewed hybrid zone. In general, selection should favor the major parent ancestry genome-wide to offset the cost of genetic incompatibilities [94]. Particularly in the ACUA population, we expected an increase in *X. birchmanni* ancestry given water temperatures are increasing in natural hybrid sites and *X. birchmanni* alleles should confer greater CT_{max} values. One explanation is that the suite of *hsp* loci I investigated are not responsible for governing thermotolerance, or at least the cumulative effect size is relatively small and genetic drift has more influence. Alternatively, selection may too weak to drive significant differences in the direction we expect. Incoming migrants from upstream with greater proportions of *X. malinche* ancestry can drive the frequency of the *hsp* alleles to increase. Aside from the underdominant CT_{max} QTL, one explanation for the lack of introgression of *X. birchmanni* alleles into *X. malinche* populations is due to the geography these species inhabit.

These streams contain many physical barriers to gene flow including waterfalls and dried up pools that prevent *X. birchmanni* for migrating further upstream. This study demonstrates the complexity of mapping and identifying the genetic architecture of polygenic traits.

Geographic cline analysis revealed several loci that were considered width and center outliers in both river drainages. However, there were no loci that exhibited the same pattern of introgression between the two drainages. In addition, there were more outliers in the Pochula river drainage compared to the Huazalingo drainage. Interestingly, most of the outliers in the Pochula cline were skewed upstream (i.e. increase in *X. malinche* ancestry along the cline). Since it's more difficult for alleles to introgress upstream as a result of physical barriers in the river, this finding is not unexpected. However, we still expect increases in *X. birchmanni* ancestry if ecological selection favors increased thermotolerance. Together, these findings draw attention to the difficulties of discerning the impact hybridization has on trait evolution and adaptive introgression.

5. CONSPECIFIC SPERM PRECEDENCE IN NATURALLY-HYBRIDIZING SWORDTAILS

5.1 Abstract

Assortative mating is of primary importance to the origin and maintenance of reproductive isolation. When hybrids have reduced fitness, selection favors the evolution of mechanisms that bias fertilization towards conspecifics. A vast body of research has focused on premating barriers to hybridization, particularly mate choice. However, theory suggests that mate choice is a weak barrier to gene flow under the best circumstances. Postmating-prezygotic mechanisms have received less focus in speciation studies but may play an equally important role in maintaining reproductive isolation between hybridizing species. For this project, I take advantage of a unique system whereby assortative mating maintains isolation in structured populations, despite any evidence of behavioral mating preferences in the chemical or visual modality. I propose that cryptic female choice may play a role in preventing ongoing hybridization and maintaining population structure. Uncovering whether these mechanisms are at play in recently diverged sister species is pivotal to understanding patterns that contribute to speciation.

5.2 Introduction

When populations become isolated and begin to diverge in allopatry, independent substitutions may accumulate among the genomic background of each lineage. Upon secondary contact, attempts to hybridize between parental genomes will expose genetic incompatibilities and lead to reduced hybrid fitness. As a result, reinforcement will drive the evolution of premating isolation barriers as a response to postzygotic selection against hybrids [95]. Until recently, most studies in speciation biology focused on premating and postzygotic isolation barriers as the primary barriers preventing hybrid production and gene flows. However, when females mate with multiple conspecific and heterospecific males, postmating-prezygotic barriers can reduce hybrid production and evolve in a process similar to reinforcement [96, 97]. Females perform multiple bouts of mate choice when they mate with multiple conspecific and heterospecific males, but when pre-

mating barriers (e.g. mate choice) are incomplete or break down, postmating-prezygotic barriers may be experienced to a severe degree to overcome the costs of mating, maintain isolation, and prevent gene flow with heterospecifics [98, 99, 100, 101]. These types of barriers include isolation mechanisms that act between the postmating and prezygotic stages of mate choice. The evolution of conspecific sperm precedence, the disproportionate use of conspecific sperm by females when conspecific and heterospecific sperm are competing simultaneously, can be an effective barrier to reproduction [98]. This is often achieved by sperm competition [102] and cryptic female choice [103, 104, 105].

Sperm competition occurs when sperm from more than one male compete within the reproductive tract of the female to fertilize the ova. In species where females engage in multiple sexual encounters and competition between sperm from more than one male exists, selection can favor certain sperm and ejaculate characteristics. Traits most influenced by this competition include sperm swimming velocity, morphology, viability, and longevity [106, 107, 108]. In guppies (*Poecilia reticula*) and green swordtails (*X. helleri*), males with faster swimming sperm sire a greater proportion of offspring than their competitors [109, 110]. Similarly, in *X. nigrensis*, males with more viable sperm sire a greater proportion of the offspring [111]. Sperm swimming speed can vary within and among ejaculates of the same male [108]. While associations between sperm traits and competitive success have been demonstrated, sperm quality and the specific factors contributing to that success remains difficult to measure and reports in the literature are mixed. For instance, faster sperm seem to increase fertilization success in many taxa [107], but in others, slower sperm have increased longevity and thus a better chance of fertilization [108]. Generally, which combination of traits are under selection and contributing the greatest effect towards fertilization success will be highly context-dependent and taxa specific.

Cryptic female choice within the reproductive tract provides another opportunity for selection to take place. Seminal fluid components, sperm-egg surface interactions, and the ovarian fluid present in the female reproductive tract provide mechanisms by which females can discriminate against sperm between multiple males [112]. Here, females control competitive fertilization suc-

cess by influencing the time of insemination, sperm transfer, sperm storage/dumping, and sperm-egg attraction [113]. In guppies, females use ovarian fluid to discriminate against closely related individuals [114]. This mechanism may be most important when females cannot prevent copulation, mate with suboptimal males, or in broadcast spawners where premating barriers are weak [115, 101]. However, in house sparrows, reproductive fluid components do not effect sperm performance between conspecific and heterospecific males which could account for their tendency to hybridize [100]. Assortative mating [116] and mate choice [117, 118] are by themselves insufficient to maintain premating isolation.

I take advantage of a unique system whereby assortative mating maintains isolation in structured populations, despite any evidence of behavioral mating preferences in the chemical or visual modality [24]. Individuals in the *Xiphophorus* genus occupy shallow pools and streams which can become isolated from other pools up and downstream for a period of time. In regions of species overlap (i.e. hybrid zones) where females may become isolated from a larger pool of potential mates, they may take on conspecific and heterospecific matings. In *Xiphophorus*, females are internal fertilizers. Males pass bundles of sperm (spermatozeugmata) to the female reproductive tract via a modified anal fin called the gonopodium. Bundles dissociate once in the reproductive tract and can be stored for several months (Constantz 1989). Gestation lasts approximately 25 days. Female *X. birchmanni* exhibit reproductive skew for certain males over others in broods sired by more than one conspecific male [23, 119]. Researchers have used guppies (*Poecilia reticulata*) and swordtails (*Xiphophorus spp.*) to reveal high levels of sperm competition and cryptic female choice within species [109, 110, 114, 120, 121, 122]. Although these studies have shed valuable insight into intraspecific postmating-prezygotic sexual selection, whether females disproportionately favor conspecific sperm during mating interactions between species has remained relatively overlooked, despite a history of extensive hybridization [123].

Here, I investigate the postmating-prezygotic barriers to reproduction in recently diverged sister species *X. birchmanni* and *X. malinche*. First, I evaluate sperm performance metrics to test for asymmetries in predictors of sperm competition between species. Second, I quantify fertilization

bias in females that have been artificially inseminated by sperm from a conspecific and heterospecific male.

If selection is acting against hybrid phenotypes, then females are expected to exhibit conspecific sperm precedence such that a disproportional amount of conspecific sperm fertilizes the female's eggs [97]. If conspecific sperm precedence mediates hybridization in this way, then I predict a greater proportion of embryos will be fertilized by the conspecific male. If there happens to be no effect, it suggests biased fertilization is not responsible for maintaining isolation and that other mechanisms should be considered.

5.3 Methods

5.3.1 Sperm bundle concentration

The species average sperm concentration per bundle had not been reported in these species. Therefore, I needed to know the concentration of sperm per bundle in order to inseminate females with equal proportions of sperm from competing males. I used the same 10 *X. birchmanni* and 10 *X. malinche* males in this experiment. Sperm bundles were stripped from anesthetized males, plated in sperm extender solution (INRA 96), and photographed under a dissection microscope. All bundles were pipetted into a tube to a final volume of 50 μL. The concentration of sperm/μL was determined by counting cells under a hematocytometer and dividing by the number of bundles collected. The number of bundles collected was determined by counting the number of bundles in the photographs (Figure D.2).

5.3.2 Artificial insemination

All males and females were tagged, photographed and fin clipped before sperm extractions. Sperm was stripped from a male and transferred to a microscope slide with sperm extender solution where sperm remained quiescent. To activate the sperm, I transfered 2 μL containing 100 bundles to 16 μL sperm extender solution. This process was repeated for the second male, whereby 100 bundles were transferred to the same tube. I used a Drummond sequencing pipette (Sigma-Aldrich) to inseminate virgin females with 3 μL of the sperm mixture. Before inseminating, I first

punctured the gonoduct sinus with a sequencing pipette tip so that I can easily enter the gonoduct and inseminate sperm. All fish were anesthesized in MS-222 before handling. Embryos from inseminated females will be dissected out and preserved in ethanol for genotyping 25 days after insemination.

5.3.3 Sperm physiology

Sperm from each male was placed on a standard Leja 12 micron chamber slide containing sperm extender solution. I measured the following sperm physiological traits: average path velocity (VAP), straight path velocity (VSP), curvilinear velocity (VCL), straightness (VSL/VAP * 100), total motility, and progressive motility on a computer-assisted sperm analyzer (Figure D.3). Total motility is the ratio of motile cells to the total cell concentration expressed as percentage. Progressive motility is the number of cells that move with path velocity greater than medium VAP cut-off and having straightness greater than 75%. VAP, VCL, and VSL are positively correlated in many live-bearing fish [109, 110, 122]. To assess whether these variables were correlated in this system, I constructed correlation matrices. Each trait was highly correlated with one another (Figure D.1, R^2 values range from 0.94 to 1). For this reason, I use VAP for the following analyses.

5.3.4 DNA extraction

DNA from dissected embryos was extracted using an in-house protocol. Briefly, $\sim$10 mg of fin tissue was added to 600 μL of cell lysis solution (0.1 M Tris, 0.1 M EDTA, 1% SDS, pH 7.1) with 8 μL of proteinase K (20 mg/mL). All samples were incubated at 55 °C overnight on a plate agitator. Samples were removed from incubator and cooled at room temperature for 15 minutes. I added 250 μL of protein precipitation solution (7.5 M NH_4AoC), vortexed, and spun down at 13,000 rpm for 5 minutes. The resulting supernatant was added to 600 μL of ice cold 100% isopropanol. Samples were mixed by inverting and spun down at 13,000 rpm for 10 minutes. The supernatant was discarded and 70% EtOH was added to each sample. Samples were spun down at 13,000 rpm for 10 minutes. The EtOH was discarded and the samples air dried for 15 minutes. I added 30 μL of TE (1 mM EDTA, 10MM Tris, pH 8.0) and allowed the DNA to rehydrate at 4 °C overnight.

Concentration of each sample was determined by a Qubit fluorometer 3.0 (Applied Biosystems) and diluted to 5ng/μL. Samples were stored at -20 °C.

5.3.5 Library preparation, sequencing, and genotyping

Three nuclear markers were amplified and sequenced. These markers are species-specific and have been used to diagnosis heterozygotes from homozygotes in preview studies (Culumber et al 2012). A single PCR reaction was conducted to amplify all markers simultaneously. The reaction contained 5 μL 5X Q5 buffer (New England Biolabs), 2 μL gDNA (5 ng/μL), 0.5 μL 10 mM dNTPs, 2.5 μL 10 μM primers, 0.25 μL Q5 Taq polymerase, and 14.75 μL dH_20. Primers are a mixture of all 6 primers (3 forward and 3 reverse) at 10 μM each. PCR amplification was performed with cycling conditions as follows: 98 °C for 1 minute, then 32 cycles of 98 °C for 10 s, 58 °C for 20 s and 72 °C for 30 s, and finally 72 °C for 5 minutes. Samples were cleaned using Ampure XP beads and eluted in elution buffer (EB) solution (10 mM Tris, pH 9.4; Qiagen).

The second PCR was similar, except primers included Illumina adapter sequences added at the 3‘ ends. Samples were clean as before. A final PCR was performed, again using the same conditions, except we used indexing primers compatible with the Illumina primers used above. Samples were cleaned for a final time as described above. Samples were pooled and sequenced on an Illumina MiSeq Nano Sequencer (1 M 2X150 reads). Raw sequence reads were assessed for quality using FastQC. Reads were mapped to the *X. birchmanni* reference genome using Burrows-Wheeler Aligner (BWA; [59]) and BAM files were manually scored has homozygous or heterozygous on Integrative Genomics Viewer (IGV; [124]).

5.3.6 Logistic Regression Analysis

To determine the relationships between male sperm traits and fertilization success, I used a generalized linear model with a quasibinomial error distribution and a logit link function. A quasibinomial error distribution was used because the data were overdispersed. In addition, each sample was weighted by the number of offspring in order to control for uneven brood sizes among samples. For each replicate, each male was arbitrarily labelled as ’A’ or ’B’. The response variable

was determined as the number of fertilized embryos by male 'B' divided by the total number of offspring sired in each brood. Predictor variables included species, VAP, and which ejaculate was collected first. Sperm swimming velocity was calculated as the difference in trait values between the two putative sires (trait B-trait A). I constructed two reduced subsequent models, one including only species and sperm swimming velocity as predictor variables, and the other only including species.

I used ANOVA to check if the additional variables contribute to the predictive ability of the model. Because the use of a quasibinomial error distribution does not produce an AIC value that I could use to select the best model, I performed a likelihood ratio test (where the null model corresponds to the reduced model) by calculating the probability for observing a chi-squared distributed test statistic (i.e. the change in deviance) as extreme or more extreme. Adding additional terms did not improve the model. I chose to report the model with the fewest predictor variables. Therefore, the final model only included species as a predictor variable.

5.4 Results

5.4.1 Sperm bundle concentration and physiology

There was no evidence to suggest that the variance in bundle concentration is significantly different for the two species (Levene's test, $p = 0.054$). Moreover, the distribution of the data are not significantly different from a normal distribution (Shapiro-Wilk normality test, $W = 0.92$, $p = 0.363$; $W = 0.95$, $p = 0.671$ for *X. birchmanni* and *X. malinche*, respectively). The concentration of sperm per bundle was not significantly different between the two species (Figure 5.1, Student's two sample t-test, $t(16) = -1.5628$, $p = 0.1377$).

There were no significant differences in sperm physiological trait values between *X. birchmanni* and *X. malinche*. (Figure D.5; MANOVA, motile: $F = 2.8875$, $p = 0.0968$; progressive: $F = 1.2643$, $p = 0.2674$; straightness: $F = 0.1282$, $p = 0.7221$; VCL: $F = 0.1259$, $p = 0.7246$; VSL: $F = 0.2232$, $p = 0.6391$; VAP: $F = 0.4513$, $p = 0.5055$). Sperm velocity variables VCL, and VSL are highly correlated with VAP (both r-values > 0.94, $p < 0.001$). I used VAP for all remaining analyses to

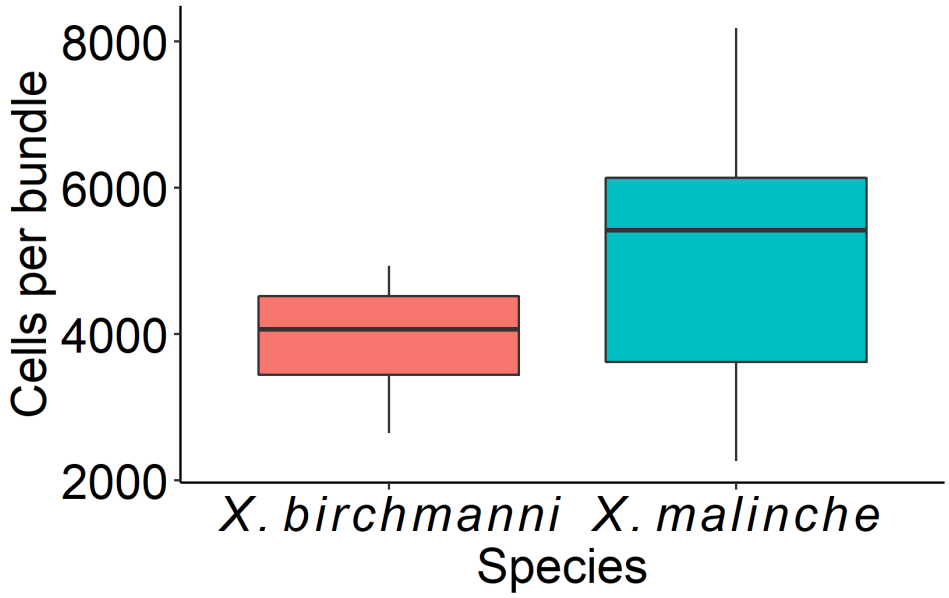

Figure 5.1: Number of cells per sperm bundle between *X. birchmanni* and *X. malinche*.

reduce the number of predictor variables and avoid multiple comparisons.

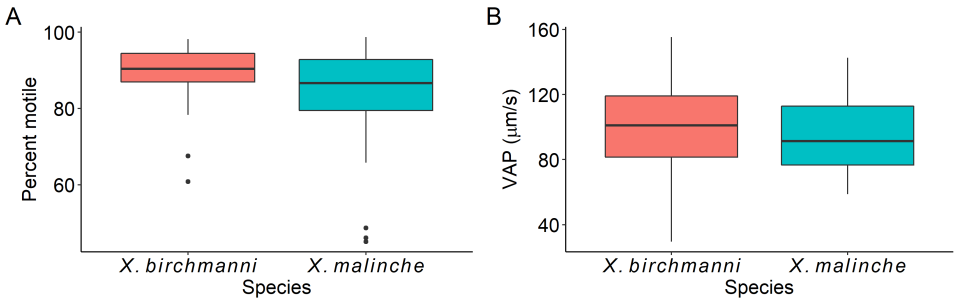

Figure 5.2: A) Percent motile and B) average path velocity (VAP) between *X. birchmanni* and *X. malinche* males used for artificial insemination.

5.4.2 Artificial insemination

A total of 23 artificial inseminations were conducted. Of these, 9 females had fertilized embryos for a 40% success rate. The null hypothesis is that each male has an equal chance of fertilizing eggs when sperm compete within the female reproductive tract. Among the 9 successful

artificial inseminations, all females showed a bias towards conspecific males (Figure 5.3; mean proportion *X. birchmanni* = 0.85, sd = 0.16, one-tailed binomial sign test, p = 0.002). Species had a significant effect on determining the proportion of offspring sired by male 'B' (Logistic regression, t = -3.319, p = 0.0451). Among individuals used for artificial insemination, there was no difference in sperm performance metrics VAP or motility between species, with numerical trends favoring *X. malinche*. There were no significant differences in sperm trait values between males that successfully fertilized and those that did not for either species.

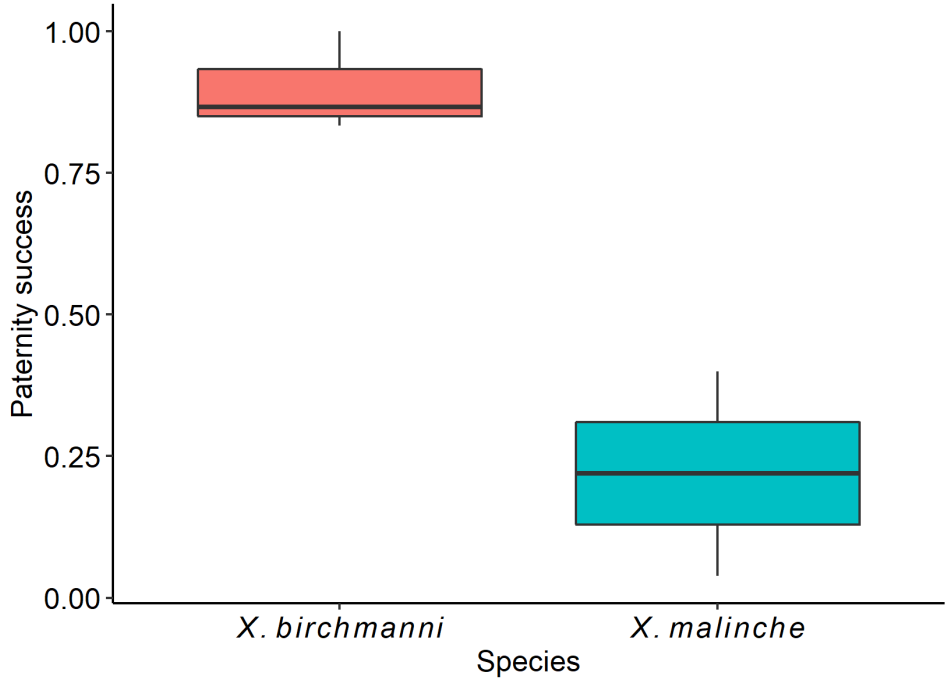

Figure 5.3: Paternity success of *X. birchmanni* versus *X. malinche* males in the artificial insemination assay.

5.5 Discussion

Postmating-prezygotic sexual selection such as conspecific sperm precedence can have important biological implications by facilitating species divergence via a reinforcement-like process [96, 97, 98]. If hybrids are less fit and premating isolation is weak or incomplete, selection can

favor discrimination against heterospecifics within the reproductive tract of a female and lead them to mate assortatively [125, 126, 127]. In cases like these, fitness is higher for females that favor conspecific sperm when mated with conspecific and heterospecific males [128]. Even in the absence of selection, postmating-prezygotic preferences for conspecifics can arise simply due to divergence of female and male reproductive traits between lineages, particularly as a consequence of sexual conflict [129].

This project was motivated by recent findings that show near-perfect reproductive isolation between genetically divergent hybrid clusters in the swordtail fish *X. birchmanni* and *X. malinche* [24]. While the two parental species have strong conspecific preferences based on pheromone profiles, hybrids have no detectable behavioral preference for homotypic mates. This suggests cryptic choice may be the mechanism preventing gene flow. Both the premating and postzygotic barriers between *X. birchmanni* and *X. malinche* have been well characterized [21, 22, 29, 32], but whether females disproportionately favor conspecific sperm during mating interactions between these species remains unknown. This is particularly important as premating isolating mechanisms are frequently susceptible to disruption of communication channels or other environmental context, yet ongoing hybridization is extremely rare [18, 19, 24]. The fertilization success of heterospecific males is not always low, but typically decreases considerably when in competition with conspecific gametes [130]. In *Xiphophorus*, many species are capable of producing viable and fertile hybrid offspring [131].

Here, I investigate the postmating-prezygotic barriers to reproduction in recently diverged sister species *X. birchmanni* and *X. malinche*. My primary aim was to quantify fertilization bias in females that have been inseminated by equal proportions of sperm from conspecific and heterospecific males. I found none of the sperm physiological traits to significantly differ between the two species. I did not expect to see large differences in these traits considering they are closely related (99.9% genome-wide similarity) and exhibit similar mating strategies where males compete among each other to maximize reproductive fitness by mating with as many females as possible. Most notably, the two species did not have an equal probability of fertilizing eggs when competing

within the reproductive tract simultaneously. Conspecific males had significantly greater fertilization success. This pattern held regardless of the species the sperm was collected from first. Since females mate multiply, this may constitute an important barrier to gene flow between these species. Moreover, it demonstrates postmating-prezygotic sexual selection has the propensity to attenuate hybridization.

This work is the first step to grasping the way postmating-prezygotic processes mediate how individuals avoid fitness costs of hybridization in an environment where premating signals are unreliable. Future studies, such as conducting the reciprocal experiment to test for asymmetries in conspecific sperm precedence would be necessary to understand the full impact of postmating-prezygotic selective pressures in this system. Moreover, studies investigating sperm physiology within ovarian fluid would shed novel insight into the degree to which females mediate fertilization success. Finally, to better understand how postmating-prezygotic selection is associated with hybridization, a comparative analysis between sperm traits within the entire *Xiphophorus* genus could be correlated with rates of hybridization among lineages.

6. CONCLUSIONS

Uncovering how hybrid populations evolve is critical to understanding the speciation process. This requires measuring many aspects of a hybridizing system, including genotype frequencies, morphological changes, mating interactions, environmental selective pressures, and population demography. Perhaps more important is how variation in these factors contribute to the evolution of hybrid populations and the evolutionary consequences it has on speciation. In my dissertation, I investigate how hybrid populations evolve and how barriers to reproduction, or the lack thereof, mediate hybridization in nature.

In the second and third chapters of my dissertation, I monitor how morphological characteristics change in natural and experimental hybrid populations, respectively. In Chapter II, I measure morphological variation in male traits between three independent natural hybrid populations. Two of the hybrid populations are located in the lowlands at similar elevations yet differ in genome-wide hybrid index. The third hybrid population is a structured population located at higher elevations that consist of a *X. malinche*-skewed cluster and an *X. birchmanni*-skewed cluster of individuals that are reproductively isolated. I used a combination of univariate and multivariate analyses to reveal the lowland hybrid populations were more similar in morphology compared to the highland hybrid population. Moreover, they are more similar to the lowland parental species, *X. birchmanni*, than were individuals from the AGZC population. However, this signal is obscured due to population structure. When I partition individuals in the AGZC population into *X. malinche*-skewed and *X. birchmanni*-skewed, I found the *X. birchmanni*-skewed individuals to be more similar to *X. birchmanni* morphotypes. Male *X. malinche*-skewed individuals were more like *X. malinche* morphotypes. A major question I am testing is whether elevation and thermal environment are driving morphological evolution or if morphology depends more heavily on genome wide hybrid index. My findings suggest in the lowland populations, environment drives morphology to resemble the lowland parental species. In the highland hybrid population, morphology aligns with genome-wide hybrid index. These findings reveal independent hybrid populations can exhibit sim-

ilar morphometric combinations of traits despite genome-wide ancestry composition, but that it is context dependent. In other cases, morphology aligns with genome-wide ancestry.

I pair Chapter II closely with Chapter III where I explore evolutionary change in experimental early generation hybrid populations. In addition, I make use of whole genome data in subset of my samples to monitor changes in hybrid index over time. The primary aim was to monitor hybrid evolution in response to different thermal environments during the incipient stages of evolution. We seeded eight replicate populations of F1 generation hybrids at three different elevations. Our primary findings reveal there is more variation in morphology between replicates within hybrid populations than between hybrid populations at different sites. We expect the breadth of variation to be larger in these early generation hybrids relative to the natural populations whereby selection and genetic drift have had time to fix alleles. Empirical studies reveal phenotypic variation can decrease within only a few generations in early generation hybrids [132]. While non-significant, standard length traits variance decreased in the lowland and highland populations ($p = 0.065$, $p = 0.052$, respectively) and increased in the intermediate hybrid population ($p = 0.052$) over the 4-5 generations we monitored. We expect hybrid populations closer to the parental species habitats to experience selective pressures that shape trait evolution towards the respective parental morphology [36]. In early generation hybrids, this will manifest as divergent selection towards the optimal trait values followed by stabilizing selection around the mean. Increase in variance at the intermediate population may be due to weakened constraints at intermediate habitats. Season influenced sword variability over time in two of the three experimental hybrid populations. This effect is likely driven by seasonal fluctuations in water temperature that selects for *X. birchmanni* alleles during the summer and *X. malinche* alleles during the winter [19]. The genome-wide hybrid index among populations trended in the direction expected if ecological selection were driving changes in ancestry. In only the lowland hybrid population was sword length positively associated with a hybrid index. Surprisingly, no other traits were correlated with hybrid index. This suggests phenotypes do not necessarily correspond to genome-wide ancestry patterns, but instead depend on genotypes at localized genomic regions responsible for controlling trait variability.

In Chapter IV, I aim to investigate putative loci under thermal selection and detect whether allele frequencies at these regions significantly deviate from the genome-wide expectation over time in natural hybrid populations. I use a combination QTL analysis and genomic cline analysis to investigate the genomic architecture underlying thermal adaptation and how these genomic regions evolve in hybrid populations. We detected a QTL on the chromosome 22 explaining 6.9% of variation in CT_{max} in hybrids. Individuals that are heterozygous for ancestry at the QTL have reduced CT_{max}. Due to the underdominant nature of this locus, it constitutes a barrier to introgression since individuals that harbor heterozygous ancestry at QTL region on chromosome 22 will have markedly reduced thermotolerance. As such, adaptive introgression of *X. birchmanni* alleles may not be sufficient to offset the cost of hybridization. Such consequences can limit the success of genetic rescue of cold-tolerant, highland populations exposed to warming waters.

Ancestry at *hsp* genes distributed across the genome did not skew towards one parent more than another in any population. Less than 5% of *hsp* loci examined significantly increased in *X. malinche* frequency in the ACUA hybrid population. This was contrary to what we predicted. If water temperatures are increasing and *X. birchmanni* alleles confer a greater thermal advantage than we expect an increase in *X. birchmanni hsp* alleles. There are several explanations to this pattern. First, the suite of loci I investigated may not be responsible for controlling a large enough proportion of variance in thermotolerance. Alternatively, selection may simply be too weak to drive significant changes in the direction we expect. Instead, genetic drift may take precedence. Incoming migrants from upstream with greater proportions of *X. malinche* ancestry can drive the frequency of the *hsp* alleles to increase as well. Due to the geography these species inhabit, it is more feasible for alleles to flow downstream since the river contains many physical barriers that hinder alleles to introgress upstream.

The geographic cline analysis yielded several loci considered center and/or width outliers among the two river drainages. Interestingly, most of the outliers in the Pochula cline were skewed upstream (i.e. increase in *X. malinche* ancestry along the cline). Since it's more difficult for alleles to introgress upstream as a result of physical barriers in the river, this finding is not unex-

pected. However, we still expect increases in *X. birchmanni* ancestry if ecological selection favors increased thermotolerance.

In my penultimate chapter, I investigate the role of postmating-prezygotic sexual selection has on mediating fertilization bias between hybridizing species. Specifically, I tested whether females bias fertilization towards conspecific males when inseminated with equal proportions of conspecific and heterospecific sperm and whether sperm physiological traits correlate with fertilization success. I found conspecific males to have a greater probability of fertilization success relative to their heterospecific competitors. Since females mate multiply, this may constitute an important barrier to gene flow between these species. Moreover, it demonstrates postmating-prezygotic sexual selection has the propensity to attenuate hybridization.

Together, this dissertation illuminates the complexity of hybrid evolution and despite similarities between independent hybrid zones, heterogeneity in how these populations evolve clearly exist. This highlights the need for further investigation of hybrid zone evolution in order to fully understand how hybridization mediates the speciation process.

REFERENCES

[1] R. Abbott, D. Albach, S. Ansell, J. W. Arntzen, S. J. Baird, N. Bierne, J. Boughman, A. Brelsford, C. A. Buerkle, R. Buggs, *et al.*, "Hybridization and speciation," *Journal of evolutionary biology*, vol. 26, no. 2, pp. 229–246, 2013.

[2] O. Seehausen, "Hybridization and adaptive radiation," *Trends in ecology & evolution*, vol. 19, no. 4, pp. 198–207, 2004.

[3] B. A. Payseur and L. H. Rieseberg, "A genomic perspective on hybridization and speciation," *Molecular ecology*, vol. 25, no. 11, pp. 2337–2360, 2016.

[4] A. M. Kearns, M. Restani, I. Szabo, A. Schrøder-Nielsen, J. A. Kim, H. M. Richardson, J. M. Marzluff, R. C. Fleischer, A. Johnsen, and K. E. Omland, "Genomic evidence of speciation reversal in ravens," *Nature Communications*, vol. 9, no. 1, pp. 1–13, 2018.

[5] J. Mallet, "Hybrid speciation," *Nature*, vol. 446, no. 7133, pp. 279–283, 2007.

[6] M. Schumer, G. G. Rosenthal, and P. Andolfatto, "How common is homoploid hybrid speciation?," *Evolution*, vol. 68, no. 6, pp. 1553–1560, 2014.

[7] J. A. Coyne, H. A. Orr, *et al.*, *Speciation*, vol. 37. Sinauer associates Sunderland, MA, 2004.

[8] R. C. Lewontin and L. Birch, "Hybridization as a source of variation for adaptation to new environments," *Evolution*, pp. 315–336, 1966.

[9] M. A. Bell and M. P. Travis, "Hybridization, transgressive segregation, genetic covariation, and adaptive radiation," *Trends in ecology & evolution*, vol. 20, no. 7, pp. 358–361, 2005.

[10] N. F. Parnell, C. D. Hulsey, and J. T. Streelman, "Hybridization produces novelty when the mapping of form to function is many to one," *BMC Evolutionary Biology*, vol. 8, no. 1, pp. 1–11, 2008.

[11] M. Tobler and E. W. Carson, "Environmental variation, hybridization, and phenotypic diversification in cuatro ciénegas pupfishes," *Journal of evolutionary biology*, vol. 23, no. 7, pp. 1475–1489, 2010.

[12] W. S. Moore, "An evaluation of narrow hybrid zones in vertebrates," *The Quarterly Review of Biology*, vol. 52, no. 3, pp. 263–277, 1977.

[13] N. H. Barton and G. M. Hewitt, "Analysis of hybrid zones," *Annual review of Ecology and Systematics*, pp. 113–148, 1985.

[14] G. M. Hewitt, "Hybrid zones-natural laboratories for evolutionary studies," *Trends in ecology & evolution*, vol. 3, no. 7, pp. 158–167, 1988.

[15] N. H. Barton and G. M. Hewitt, "Adaptation, speciation and hybrid zones," *Nature*, vol. 341, no. 6242, pp. 497–503, 1989.

[16] C. M. Curry, "An integrated framework for hybrid zone models," *Evolutionary Biology*, vol. 42, no. 3, pp. 359–365, 2015.

[17] S. Sankararaman, S. Mallick, M. Dannemann, K. Prüfer, J. Kelso, S. Pääbo, N. Patterson, and D. Reich, "The genomic landscape of neanderthal ancestry in present-day humans," *Nature*, vol. 507, no. 7492, pp. 354–357, 2014.

[18] Z. Culumber, H. Fisher, M. Tobler, M. Mateos, P. Barber, M. Sorenson, and G. Rosenthal, "Replicated hybrid zones of xiphophorus swordtails along an elevational gradient," *Molecular ecology*, vol. 20, no. 2, pp. 342–356, 2011.

[19] Z. Culumber, D. Shepard, S. Coleman, G. Rosenthal, and M. Tobler, "Physiological adaptation along environmental gradients and replicated hybrid zone structure in swordtails (teleostei: Xiphophorus)," *Journal of evolutionary biology*, vol. 25, no. 9, pp. 1800–1814, 2012.

[20] H. S. Fisher, B. B. Wong, and G. G. Rosenthal, "Alteration of the chemical environment disrupts communication in a freshwater fish," *Proceedings of the Royal Society B: Biological Sciences*, vol. 273, no. 1591, pp. 1187–1193, 2006.

[21] B. B. Wong and G. G. Rosenthal, "Female disdain for swords in a swordtail fish," *The American Naturalist*, vol. 167, no. 1, pp. 136–140, 2006.

[22] M. Schumer, R. Cui, D. L. Powell, R. Dresner, G. G. Rosenthal, and P. Andolfatto, "High-resolution mapping reveals hundreds of genetic incompatibilities in hybridizing fish species," *Elife*, vol. 3, p. e02535, 2014.

[23] K. A. Paczolt, C. N. Passow, P. J. Delclos, H. K. Kindsvater, A. G. Jones, and G. G. Rosenthal, "Multiple mating and reproductive skew in parental and introgressed females of the live-bearing fish xiphophorus birchmanni," *Journal of Heredity*, vol. 106, no. 1, pp. 57–66, 2015.

[24] M. Schumer, D. L. Powell, P. J. Delclós, M. Squire, R. Cui, P. Andolfatto, and G. G. Rosenthal, "Assortative mating and persistent reproductive isolation in hybrids," *Proceedings of the National Academy of Sciences*, vol. 114, no. 41, pp. 10936–10941, 2017.

[25] G. I. Jofre Rodriguez, *Patterns of Introgression Across Clinal Hybrid Zones*. PhD thesis, 2020.

[26] N. H. Barton, "The role of hybridization in evolution," *Molecular ecology*, vol. 10, no. 3, pp. 551–568, 2001.

[27] R. Dillon and J. Manzi, "Genetics and shell morphology in a hybrid zone between the hard clams mercenaria mercenaria and m. campechiensis," *Marine Biology*, vol. 100, no. 2, pp. 217–222, 1989.

[28] T. Lamb and J. C. Avise, "Morphological variability in genetically defined categories of anuran hybrids," *Evolution*, vol. 41, no. 1, pp. 157–165, 1987.

[29] H. S. Fisher, S. J. Mascuch, and G. G. Rosenthal, "Multivariate male traits misalign with multivariate female preferences in the swordtail fish, xiphophorus birchmanni," *Animal Behaviour*, vol. 78, no. 2, pp. 265–269, 2009.

[30] A. R. Beaumont, G. Turner, A. R. Wood, and D. O. Skibinski, "Hybridisations between mytilus edulis and mytilus galloprovincialis and performance of pure species and hybrid

veliger larvae at different temperatures," *Journal of Experimental Marine Biology and Ecology*, vol. 302, no. 2, pp. 177–188, 2004.

[31] D. L. Powell, C. Payne, S. M. Banerjee, M. Keegan, E. Bashkirova, R. Cui, P. Andolfatto, G. G. Rosenthal, and M. Schumer, "The genetic architecture of variation in the sexually selected sword ornament and its evolution in hybrid populations," *Current Biology*, vol. 31, no. 5, pp. 923–935, 2021.

[32] G. G. Rosenthal, X. F. de la Rosa Reyna, S. Kazianis, M. J. Stephens, D. C. Morizot, M. J. Ryan, and F. J. García de León, "Dissolution of sexual signal complexes in a hybrid zone between the swordtails xiphophorus birchmanni and xiphophorus malinche (poeciliidae)," *Copeia*, vol. 2003, no. 2, pp. 299–307, 2003.

[33] M. D. Abràmoff, P. J. Magalhães, and S. J. Ram, "Image processing with imagej," *Biophotonics international*, vol. 11, no. 7, pp. 36–42, 2004.

[34] R. J. Hyndman and Y. Khandakar, "Automatic time series forecasting: the forecast package for r," *Journal of statistical software*, vol. 27, pp. 1–22, 2008.

[35] L. H. Rieseberg, J. F. Wendel, *et al.*, "Introgression and its consequences in plants," *Hybrid zones and the evolutionary process*, vol. 70, p. 109, 1993.

[36] T. Majtyka, B. Borczyk, M. Ogielska, and M. Stöck, "Morphometry of two cryptic tree frog species at their hybrid zone reveals neither intermediate nor transgressive morphotypes," *Ecology and evolution*, vol. 12, no. 1, p. e8527, 2022.

[37] S. E. Carney, K. A. Gardner, and L. H. Rieseberg, "Evolutionary changes over the fifty-year history of a hybrid population of sunflowers (helianthus)," *Evolution*, vol. 54, no. 2, pp. 462–474, 2000.

[38] L. A. Cirino and C. W. Miller, "Seasonal effects on the population, morphology and reproductive behavior of narnia femorata (hemiptera: Coreidae)," *Insects*, vol. 8, no. 1, p. 13, 2017.

[39] N. G. Forger and S. M. Breedlove, "Seasonal variation in mammalian striated muscle mass and motoneuron morphology," *Journal of neurobiology*, vol. 18, no. 2, pp. 155–165, 1987.

[40] T. Schmoll, O. Kleven, and M. Rusche, "Individual phenotypic plasticity explains seasonal variation in sperm morphology in a passerine bird," *Evolutionary Ecology Research*, vol. 19, no. 5, pp. 547–560, 2018.

[41] R. I. Bailey, F. Eroukhmanoff, and G.-P. SæTRE, "Hybridization and genome evolution ii: Mechanisms of species divergence and their effects on evolution in hybrids," *Current Zoology*, vol. 59, no. 5, pp. 675–685, 2013.

[42] M. L. Arnold and N. H. Martin, "Hybrid fitness across time and habitats," *Trends in Ecology & Evolution*, vol. 25, no. 9, pp. 530–536, 2010.

[43] C. D. Jiggins, C. Salazar, M. Linares, and J. Mavarez, "Hybrid trait speciation and heliconius butterflies," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 363, no. 1506, pp. 3047–3054, 2008.

[44] M. Schrader, R. C. Fuller, and J. Travis, "Differences in offspring size predict the direction of isolation asymmetry between populations of a placental fish," *Biology Letters*, vol. 9, no. 5, p. 20130327, 2013.

[45] J. M. Good, M. D. Dean, and M. W. Nachman, "A complex genetic basis to x-linked hybrid male sterility between two species of house mice," *Genetics*, vol. 179, no. 4, pp. 2213–2228, 2008.

[46] M. Turelli and L. C. Moyle, "Asymmetric postmating isolation: Darwin's corollary to haldane's rule," *Genetics*, vol. 176, no. 2, pp. 1059–1088, 2007.

[47] M. Turelli and H. A. Orr, "The dominance theory of haldane's rule.," *Genetics*, vol. 140, no. 1, pp. 389–402, 1995.

[48] R. J. Pereira, F. S. Barreto, and R. S. Burton, "Ecological novelty by hybridization: experimental evidence for increased thermal tolerance by transgressive segregation in tigriopus californicus," *Evolution*, vol. 68, no. 1, pp. 204–215, 2014.

[49] H. S. Fisher and G. G. Rosenthal, "Male swordtails court with an audience in mind," *Biology Letters*, vol. 3, no. 1, pp. 5–7, 2007.

[50] A. L. Basolo and B. C. Trainor, "The conformation of a female preference for a composite male trait in green swordtails," *Animal Behaviour*, vol. 63, no. 3, pp. 469–474, 2002.

[51] A. L. Basolo, "Female preference for male sword length in the green swordtail, xiphophorus helleri (pisces: Poeciliidae)," *Animal Behaviour*, vol. 40, no. 2, pp. 332–338, 1990.

[52] J. B. Johnson and A. L. Basolo, "Predator exposure alters female mate choice in the green swordtail," *Behavioral Ecology*, vol. 14, no. 5, pp. 619–625, 2003.

[53] J. J. Kingston, G. G. Rosenthal, and M. J. Ryan, "The role of sexual selection in maintaining a colour polymorphism in the pygmy swordtail, xiphophorus pygmaeus," *Animal Behaviour*, vol. 65, no. 4, pp. 735–743, 2003.

[54] G. G. Rosenthal, C. S. Evans, and W. L. Miller, "Female preference for dynamic traits in the green swordtail, xiphophorus helleri," *Animal Behaviour*, vol. 51, no. 4, pp. 811–820, 1996.

[55] M. J. Ryan and W. E. Wagner Jr, "Asymmetries in mating preferences between species: female swordtails prefer heterospecific males," *Science*, vol. 236, no. 4801, pp. 595–597, 1987.

[56] H. Kobayashi and T. Iwamatsu, "Fine structure of the storage micropocket of spermatozoa in the ovary of the guppy poecilia reticulata," *Zoological science*, vol. 19, no. 5, pp. 545–555, 2002.

[57] A. López-Sepulcre, S. P. Gordon, I. G. Paterson, P. Bentzen, and D. N. Reznick, "Beyond lifetime reproductive success: the posthumous reproductive dynamics of male trinidadian guppies," *Proceedings of the Royal Society B: Biological Sciences*, vol. 280, no. 1763, p. 20131116, 2013.

[58] P. Andolfatto, D. Davison, D. Erezyilmaz, T. T. Hu, J. Mast, T. Sunayama-Morita, and D. L. Stern, "Multiplexed shotgun genotyping for rapid and efficient genetic mapping," *Genome research*, vol. 21, no. 4, pp. 610–617, 2011.

[59] H. Li and R. Durbin, "Fast and accurate short read alignment with burrows–wheeler transform," *bioinformatics*, vol. 25, no. 14, pp. 1754–1760, 2009.

[60] N. Pekkala, "Fitness and viability of small populations: the effects of genetic drift, inbreeding, and interpopulation hybridization," *Jyväskylä studies in biological and environmental science*, no. 237, 2012.

[61] C. A. Buerkle and L. H. Rieseberg, "The rate of genome stabilization in homoploid hybrid species," *Evolution: International Journal of Organic Evolution*, vol. 62, no. 2, pp. 266–275, 2008.

[62] M. Nei, T. Maruyama, and R. Chakraborty, "The bottleneck effect and genetic variability in populations," *Evolution*, pp. 1–10, 1975.

[63] J. A. Hamilton and J. M. Miller, "Adaptive introgression as a resource for management and genetic conservation in a changing climate," *Conservation Biology*, vol. 30, no. 1, pp. 33–41, 2016.

[64] E. M. Oziolor, N. M. Reid, S. Yair, K. M. Lee, S. Guberman VerPloeg, P. C. Bruns, J. R. Shaw, A. Whitehead, and C. W. Matson, "Adaptive introgression enables evolutionary rescue from extreme environmental pollution," *Science*, vol. 364, no. 6439, pp. 455–457, 2019.

[65] M. R. Jones, L. S. Mills, P. C. Alves, C. M. Callahan, J. M. Alves, D. J. Lafferty, F. M. Jiggins, J. D. Jensen, J. Melo-Ferreira, and J. M. Good, "Adaptive introgression underlies polymorphic seasonal camouflage in snowshoe hares," *Science*, vol. 360, no. 6395, pp. 1355–1358, 2018.

[66] B. Y. Ofori, A. J. Stow, J. B. Baumgartner, and L. J. Beaumont, "Influence of adaptive capacity on the outcome of climate change vulnerability assessment," *Scientific reports*, vol. 7, no. 1, pp. 1–12, 2017.

[67] S. Foo and M. Byrne, "Acclimatization and adaptive capacity of marine species in a changing ocean," *Advances in marine biology*, vol. 74, pp. 69–116, 2016.

[68] P. Calosi, D. T. Bilton, and J. I. Spicer, "Thermal tolerance, acclimatory capacity and vulnerability to global climate change," *Biology letters*, vol. 4, no. 1, pp. 99–102, 2008.

[69] R. B. Cowles, C. M. Bogert, *et al.*, "A preliminary study of the thermal requirements of desert reptiles. bulletin of the amnh; v. 83, article 5," 1944.

[70] C. D. Becker and R. G. Genoway, "Evaluation of the critical thermal maximum for determining thermal tolerance of freshwater fish," *Environmental Biology of Fishes*, vol. 4, no. 3, pp. 245–256, 1979.

[71] W. I. Lutterschmidt and V. H. Hutchison, "The critical thermal maximum: history and critique," *Canadian Journal of Zoology*, vol. 75, no. 10, pp. 1561–1574, 1997.

[72] M. E. Feder and G. E. Hofmann, "Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology," *Annual review of physiology*, vol. 61, no. 1, pp. 243–282, 1999.

[73] N. A. Fangue, M. Hofmeister, and P. M. Schulte, "Intraspecific variation in thermal tolerance and heat shock protein gene expression in common killifish, fundulus heteroclitus," *Journal of Experimental Biology*, vol. 209, no. 15, pp. 2859–2872, 2006.

[74] F. E. Kayhan and B. S. Duman, "Heat shock protein genes in fish," *Turkish Journal of Fisheries and Aquatic Sciences*, vol. 10, no. 2, 2010.

[75] R. G. Harrison and E. L. Larson, "Hybridization, introgression, and the nature of species boundaries," *Journal of Heredity*, vol. 105, no. S1, pp. 795–809, 2014.

[76] L. Gay, P.-A. Crochet, D. A. Bell, and T. Lenormand, "Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models," *Evolution: International Journal of Organic Evolution*, vol. 62, no. 11, pp. 2789–2806, 2008.

[77] J. M. Szymura and N. H. Barton, "Genetic analysis of a hybrid zone between the fire-bellied toads, bombina bombina and b. variegata, near cracow in southern poland," *Evolution*, vol. 40, no. 6, pp. 1141–1159, 1986.

[78] S. Stankowski, J. M. Sobel, and M. A. Streisfeld, "Geographic cline analysis as a tool for studying genome-wide variation: a case study of pollinator-mediated divergence in a monkeyflower," *Molecular Ecology*, vol. 26, no. 1, pp. 107–122, 2017.

[79] C. Payne, S. Bovio, D. Powell, T. Gunn, S. Banerjee, V. Grant, G. G. Rosenthal, and M. Schumer, "Genomic insights into variation in thermotolerance between hybridizing swordtail fishes," *Molecular Ecology*, 2021.

[80] M. Schumer, D. L. Powell, and R. Corbett-Detig, "Versatile simulations of admixture and accurate local ancestry inference with mixnmatch and ancestryinfer," *Molecular ecology resources*, vol. 20, no. 4, pp. 1141–1151, 2020.

[81] R. Corbett-Detig and R. Nielsen, "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*, vol. 13, no. 1, p. e1006529, 2017.

[82] K. W. Broman, H. Wu, Ś. Sen, and G. A. Churchill, "R/qtl: Qtl mapping in experimental crosses," *bioinformatics*, vol. 19, no. 7, pp. 889–890, 2003.

[83] C. S. Haley and S. A. Knott, "A simple regression method for mapping quantitative trait loci in line crosses using flanking markers," *Heredity*, vol. 69, no. 4, pp. 315–324, 1992.

[84] S. Xu, "Theoretical basis of the beavis effect," *Genetics*, vol. 165, no. 4, pp. 2259–2268, 2003.

[85] E. P. Derryberry, G. E. Derryberry, J. M. Maley, and R. T. Brumfield, "Hzar: hybrid zone analysis using an r software package," *Molecular ecology resources*, vol. 14, no. 3, pp. 652–663, 2014.

[86] R. T. Brumfield, R. W. Jernigan, D. B. McDonald, and M. J. Braun, "Evolutionary implications of divergent clines in an avian (manacus: Aves) hybrid zone," *Evolution*, vol. 55, no. 10, pp. 2070–2087, 2001.

[87] S. Kose, M. Furuta, and N. Imamoto, "Hikeshi, a nuclear import carrier for hsp70s, protects cells from heat shock-induced nuclear damage," *Cell*, vol. 149, no. 3, pp. 578–589, 2012.

[88] J. W. Gowen *et al.*, "Heterosis.," *Heterosis.*, 1952.

[89] C. W. Stuber and J. Janick, "Heterosis in plant breeding," *Plant breeding reviews*, vol. 12, pp. 227–251, 2010.

[90] D. P. Doolittle, *Population Genetics:: Basic Principles*, vol. 16. Springer Science & Business Media, 2012.

[91] P. M. Altrock, A. Traulsen, and F. A. Reed, "Stability properties of underdominance in finite subdivided populations," *PLoS computational biology*, vol. 7, no. 11, p. e1002260, 2011.

[92] Á. J. Láruson and F. A. Reed, "Stability of underdominant genetic polymorphisms in population networks," *Journal of Theoretical Biology*, vol. 390, pp. 156–163, 2016.

[93] D. S. Wilson and M. Turelli, "Stable underdominance and the evolutionary invasion of empty niches," *The American Naturalist*, vol. 127, no. 6, pp. 835–850, 1986.

[94] M. Schumer, R. Cui, G. G. Rosenthal, and P. Andolfatto, "Reproductive isolation of hybrid populations driven by genetic incompatibilities," *PLoS genetics*, vol. 11, no. 3, p. e1005041, 2015.

[95] T. Dobzhansky, "Genetic nature of species differences," *The American Naturalist*, vol. 71, no. 735, pp. 404–420, 1937.

[96] J. L. Marshall, M. L. Arnold, and D. J. Howard, "Reinforcement: the road not taken," *Trends in Ecology & Evolution*, vol. 17, no. 12, pp. 558–563, 2002.

[97] P. D. Lorch and M. R. Servedio, "The evolution of conspecific gamete precedence and its effect on reinforcement," *Journal of evolutionary biology*, vol. 20, no. 3, pp. 937–949, 2007.

[98] D. J. Howard, P. G. Gregory, J. Chu, and M. L. Cain, "Conspecific sperm precedence is an effective barrier to hybridization between closely related species," *Evolution*, vol. 52, no. 2, pp. 511–516, 1998.

[99] S. Immler, S. Pitnick, G. A. Parker, K. L. Durrant, S. Lüpold, S. Calhim, and T. R. Birkhead, "Resolving variation in the reproductive tradeoff between sperm size and number," *Proceedings of the National Academy of Sciences*, vol. 108, no. 13, pp. 5325–5330, 2011.

[100] E. R. Cramer, M. Ålund, S. E. McFarlane, A. Johnsen, and A. Qvarnström, "Females discriminate against heterospecific sperm in a natural hybrid zone," *Evolution*, vol. 70, no. 8, pp. 1844–1855, 2016.

[101] M. Ålund née Podevin, *Sex, Sperm and Speciation: On sexual selection and fertility in hybridizing flycatchers*. PhD thesis, Acta Universitatis Upsaliensis, 2017.

[102] G. A. Parker, "Sperm competition and its evolutionary consequences in the insects," *Biological reviews*, vol. 45, no. 4, pp. 525–567, 1970.

[103] R. Thornhill, "Cryptic female choice and its implications in the scorpionfly harpobittacus nigriceps," *The American Naturalist*, vol. 122, no. 6, pp. 765–788, 1983.

[104] W. Eberhard, *Female control: sexual selection by cryptic female choice*, vol. 17. Princeton University Press, 1996.

[105] W. G. Eberhard, "Postcopulatory sexual selection: Darwin's omission and its consequences," *Proceedings of the National Academy of Sciences*, vol. 106, no. supplement_1, pp. 10025–10032, 2009.

[106] R. R. Snook, "Sperm in competition: not playing by the numbers," *Trends in ecology & evolution*, vol. 20, no. 1, pp. 46–53, 2005.

[107] T. Pizzari and G. A. Parker, "Sperm competition and sperm phenotype," *Sperm biology*, pp. 207–245, 2009.

[108] J. L. Fitzpatrick and S. Lüpold, "Sexual selection and the evolution of sperm quality," *Molecular Human Reproduction*, vol. 20, no. 12, pp. 1180–1189, 2014.

[109] C. Boschetto, C. Gasparini, and A. Pilastro, "Sperm number and velocity affect sperm competition success in the guppy (poecilia reticulata)," *Behavioral Ecology and Sociobiology*, vol. 65, no. 4, pp. 813–821, 2011.

[110] C. Gasparini, L. W. Simmons, M. Beveridge, and J. P. Evans, "Sperm swimming velocity predicts competitive fertilization success in the green swordtail xiphophorus helleri," *PloS one*, vol. 5, no. 8, p. e12146, 2010.

[111] C. C. Smith, "Opposing effects of sperm viability and velocity on the outcome of sperm competition," *Behavioral Ecology*, vol. 23, no. 4, pp. 820–826, 2012.

[112] J. A. Zeh and D. W. Zeh, "The evolution of polyandry ii: post–copulatory defenses against genetic incompatibility," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 264, no. 1378, pp. 69–75, 1997.

[113] R. C. Firman, C. Gasparini, M. K. Manier, and T. Pizzari, "Postmating female control: 20 years of cryptic female choice," *Trends in Ecology & Evolution*, vol. 32, no. 5, pp. 368–382, 2017.

[114] C. Gasparini, G. Andreatta, and A. Pilastro, "Ovarian fluid of receptive females enhances sperm velocity," *Naturwissenschaften*, vol. 99, no. 5, pp. 417–420, 2012.

[115] T. Birkhead, "Cryptic female choice: criteria for establishing female sperm choice," *Evolution*, vol. 52, no. 4, pp. 1212–1218, 1998.

[116] D. E. Irwin, "Assortative mating in hybrid zones is remarkably ineffective in promoting speciation," *The American Naturalist*, vol. 195, no. 6, pp. E150–E167, 2020.

[117] M. R. Servedio and R. Bürger, "The counterintuitive role of sexual selection in species maintenance and speciation," *Proceedings of the National Academy of Sciences*, vol. 111, no. 22, pp. 8113–8118, 2014.

[118] G. G. Rosenthal, *Mate choice: the evolution of sexual decision making from microbes to humans*. Princeton University Press, 2017.

[119] M. K. Squire, *Mate Choice and Multiple Paternity in the Xiphophorus Malinche/X. Birchmanni Hybrid System*. PhD thesis, 2015.

[120] C. Gasparini, I. Marino, C. Boschetto, and A. Pilastro, "Effect of male age on sperm traits and sperm competition success in the guppy (poecilia reticulata)," *Journal of Evolutionary Biology*, vol. 23, no. 1, pp. 124–135, 2010.

[121] C. Gasparini, E. Daymond, and J. P. Evans, "Extreme fertilization bias towards freshly inseminated sperm in a species exhibiting prolonged female sperm storage," *Royal Society Open Science*, vol. 5, no. 3, p. 172195, 2018.

[122] C. Smith and M. Ryan, "Evolution of sperm quality but not quantity in the internally fertilized fish xiphophorus nigrensis," *Journal of evolutionary biology*, vol. 23, no. 8, pp. 1759–1771, 2010.

[123] R. Cui, M. Schumer, K. Kruesi, R. Walter, P. Andolfatto, and G. G. Rosenthal, "Phylogenomics reveals extensive reticulate evolution in xiphophorus fishes," *Evolution*, vol. 67, no. 8, pp. 2166–2179, 2013.

[124] J. T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, and J. P. Mesirov, "Integrative genomics viewer," *Nature biotechnology*, vol. 29, no. 1, pp. 24–26, 2011.

[125] M. Lövy, R. Šumbera, G. Heth, and E. Nevo, "Presumed ecological speciation in blind mole rats: does soil type influence mate preferences?," *Ethology Ecology & Evolution*, vol. 32, no. 1, pp. 46–59, 2020.

[126] C. Simon, J. Tang, S. Dalwadi, G. Staley, J. Deniega, and T. R. Unnasch, "Genetic evidence for assortative mating between 13-year cicadas and sympatric "17-year cicadas with 13-year life cycles" provides support for allochronic speciation," *Evolution*, vol. 54, no. 4, pp. 1326–1336, 2000.

[127] T. Malausa, M.-T. Bethenod, A. Bontemps, D. Bourguet, J.-M. Cornuet, and S. Ponsard, "Assortative mating in sympatric host races of the european corn borer," *Science*, vol. 308, no. 5719, pp. 258–260, 2005.

[128] T. R. Birkhead and T. Pizzari, "Postcopulatory sexual selection," *Nature reviews genetics*, vol. 3, no. 4, pp. 262–273, 2002.

[129] S. Pitnick and D. J. Hosken, "Postcopulatory sexual selection," *Evolutionary behavioral ecology*, pp. 379–399, 2010.

[130] M. L. Cain, V. Andreasen, and D. J. Howard, "Reinforcing selection is effective under a relatively broad set of conditions in a mosaic hybrid zone," *Evolution*, vol. 53, no. 5, pp. 1343–1353, 1999.

[131] M. Schartl, "Evolution of xmrk: an oncogene, but also a speciation gene?," *Bioessays*, vol. 30, no. 9, pp. 822–832, 2008.

[132] G. M. Walter, T. J. Richards, M. J. Wilkinson, M. W. Blows, J. D. Aguirre, and D. Ortiz-Barrientos, "Loss of ecologically important genetic variation in late generation hybrids reveals links between adaptation and speciation," *Evolution letters*, vol. 4, no. 4, pp. 302–316, 2020.

APPENDIX A

CHAPTER II

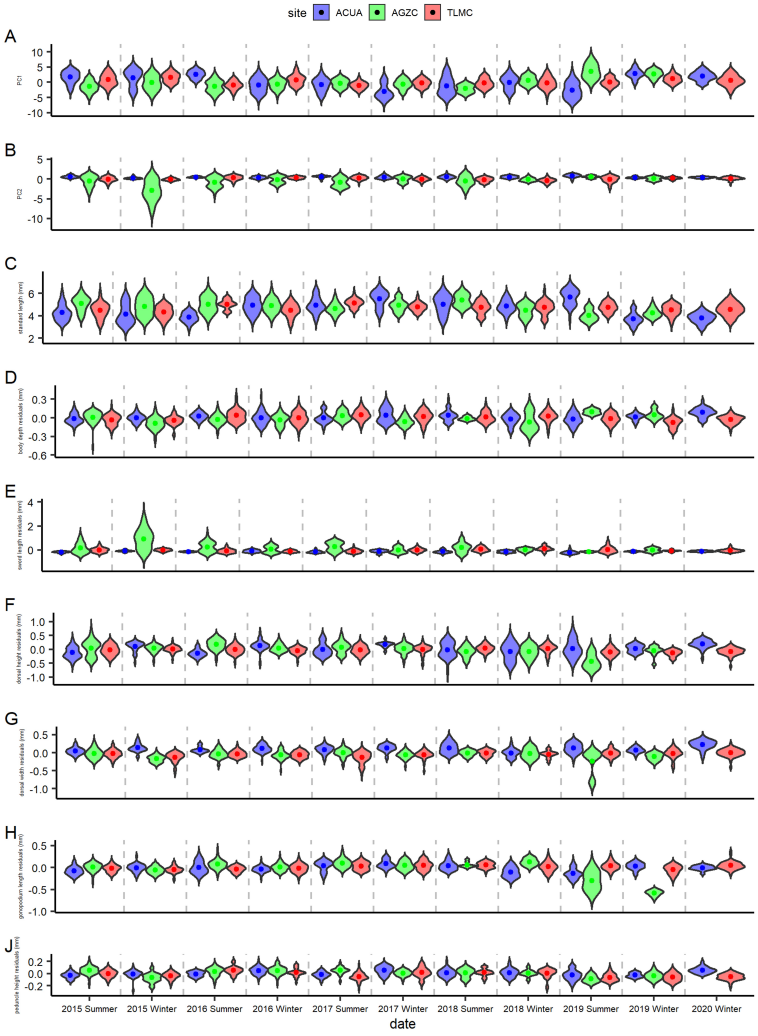

Figure A.1: Distributions of A) PC1, B) PC2, C) standard length, D) body depth residuals, E) sword length residuals, F) dorsal height residuals, G) dorsal width residuals, H) gonopodium length residuals, and J) peduncle height residuals in the natural hybrid populations among all collection dates. Colored points represent population means.

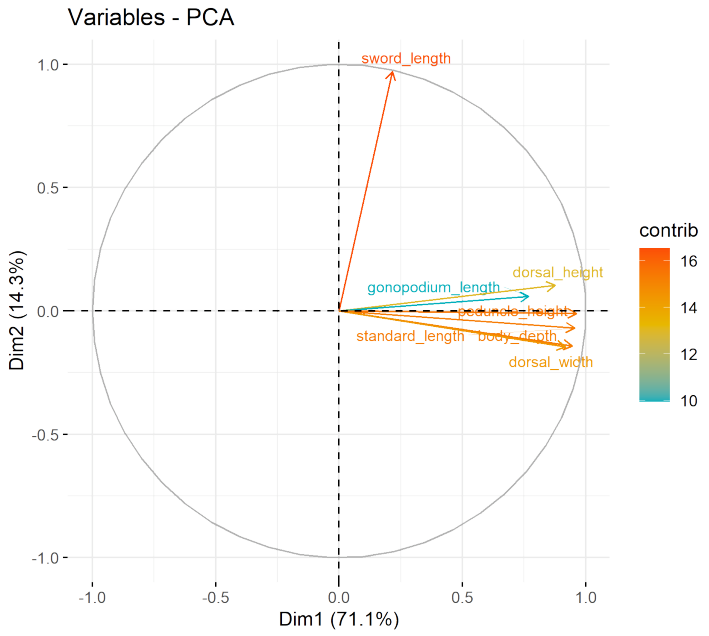

Figure A.2: Contributions of each trait among the first two dimensions for the PCA plot of the natural hybrid sites and the parental species populations.

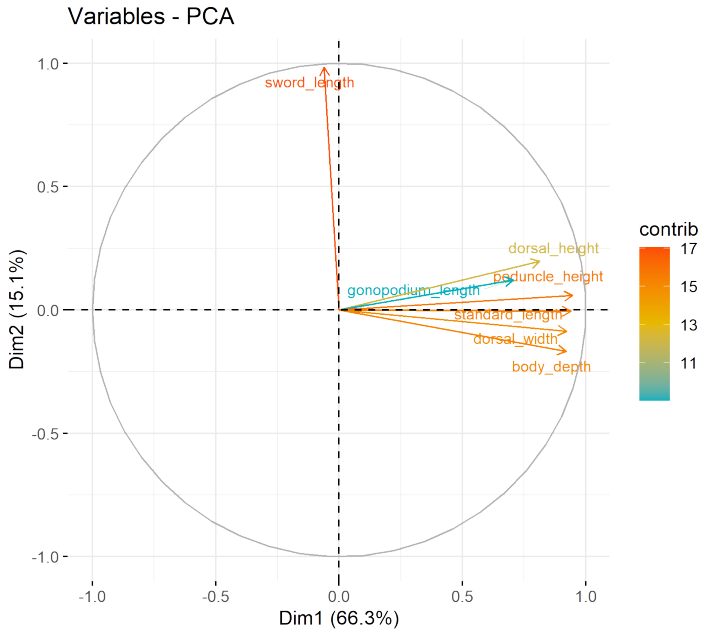

Figure A.3: Contributions of each trait among the first two dimensions for the PCA plot of the AGZC hybrid clusters and the parental species populations.

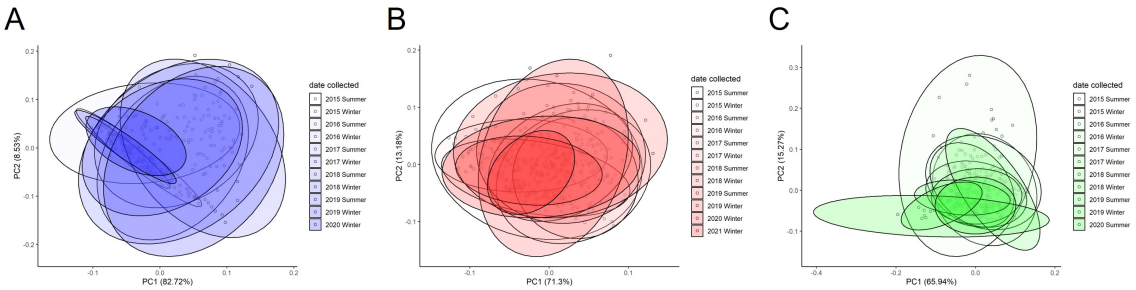

Figure A.4: PCA of A) ACUA, B) TLMC and C) AGZC over time.

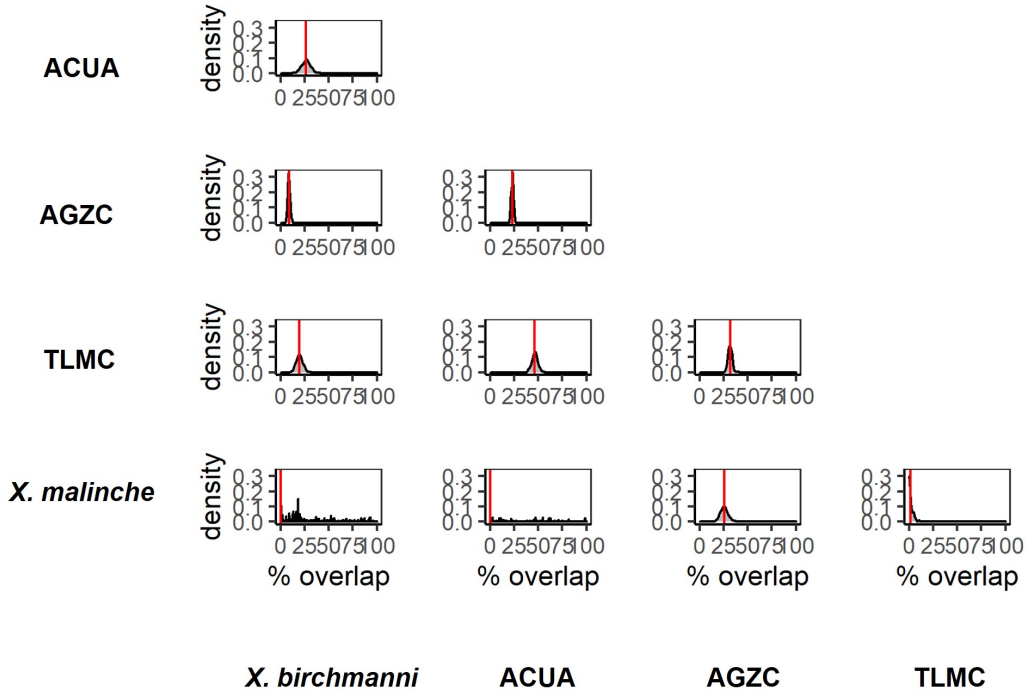

Figure A.5: Pairwise posterior ellipse overlap distributions between the natural hybrid populations and the parental species.

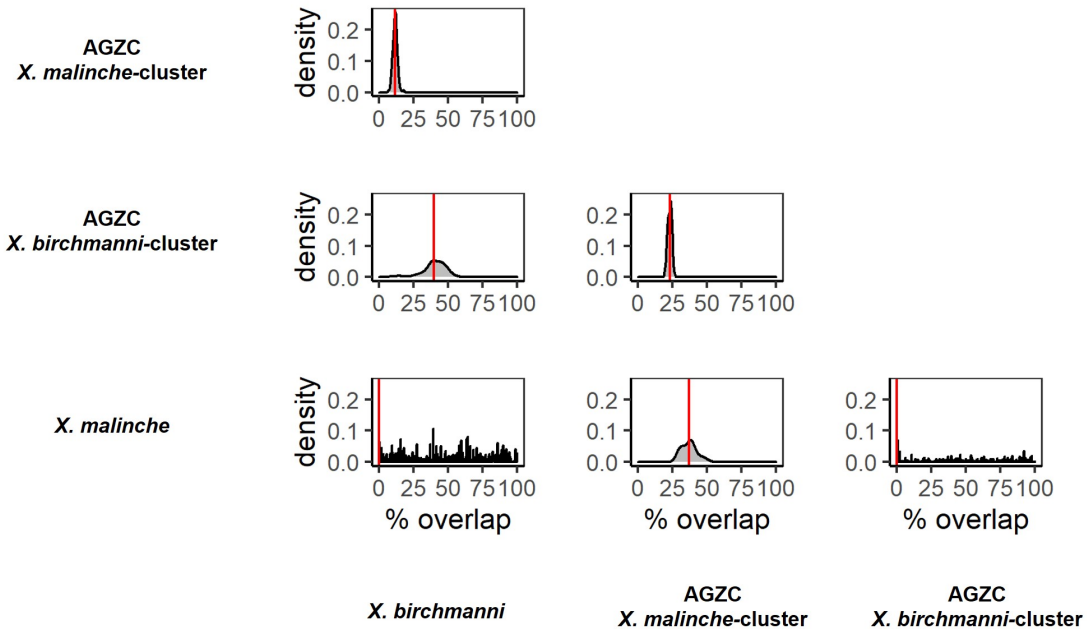

Figure A.6: Pairwise posterior ellipse overlap distributions between the AGZC hybrid clusters and the parental species.

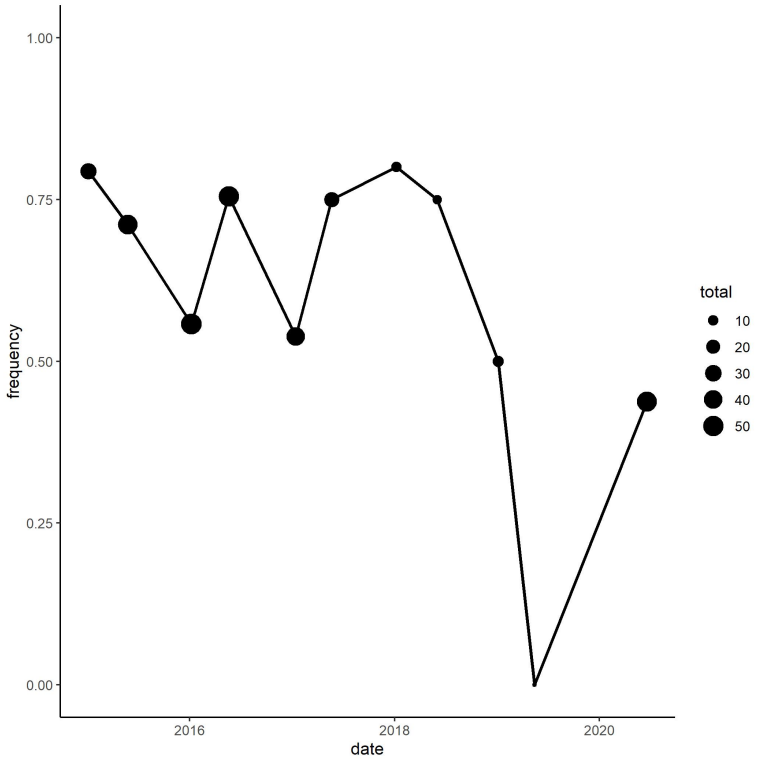

Figure A.7: Frequency of the *X. birchmanni* and *X. malinche* skewed clusters in the AGZC hybrid population over time.

APPENDIX B

CHAPTER III

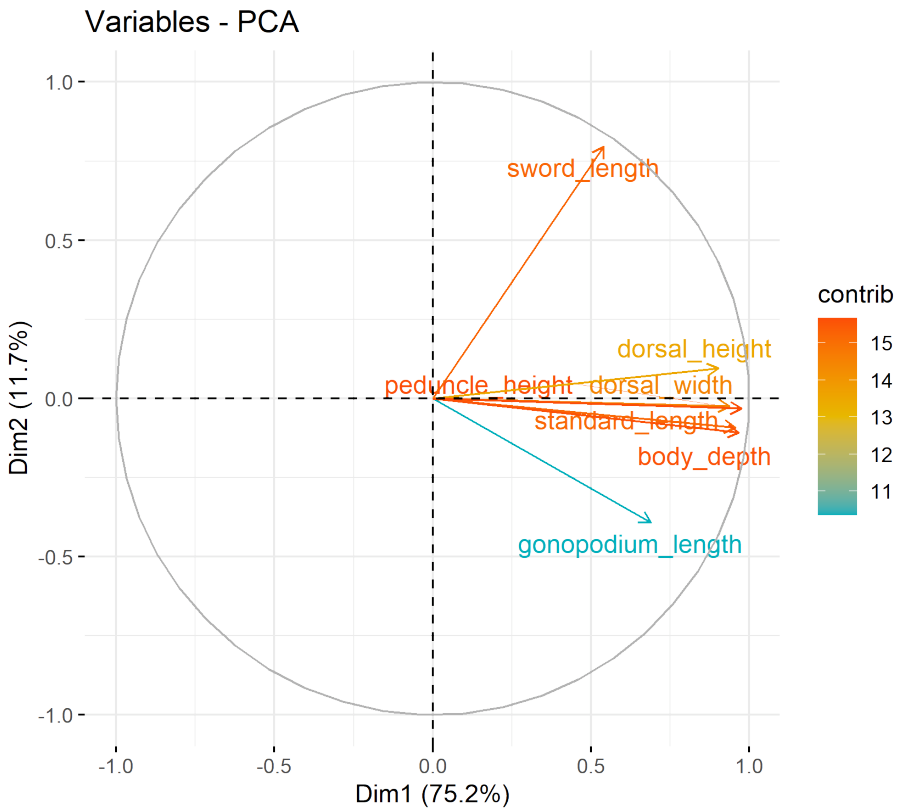

Figure B.1: Contributions of each trait among the first two dimensions for the PCA plot of the experimental, early generation hybrid sites and the parental species populations.

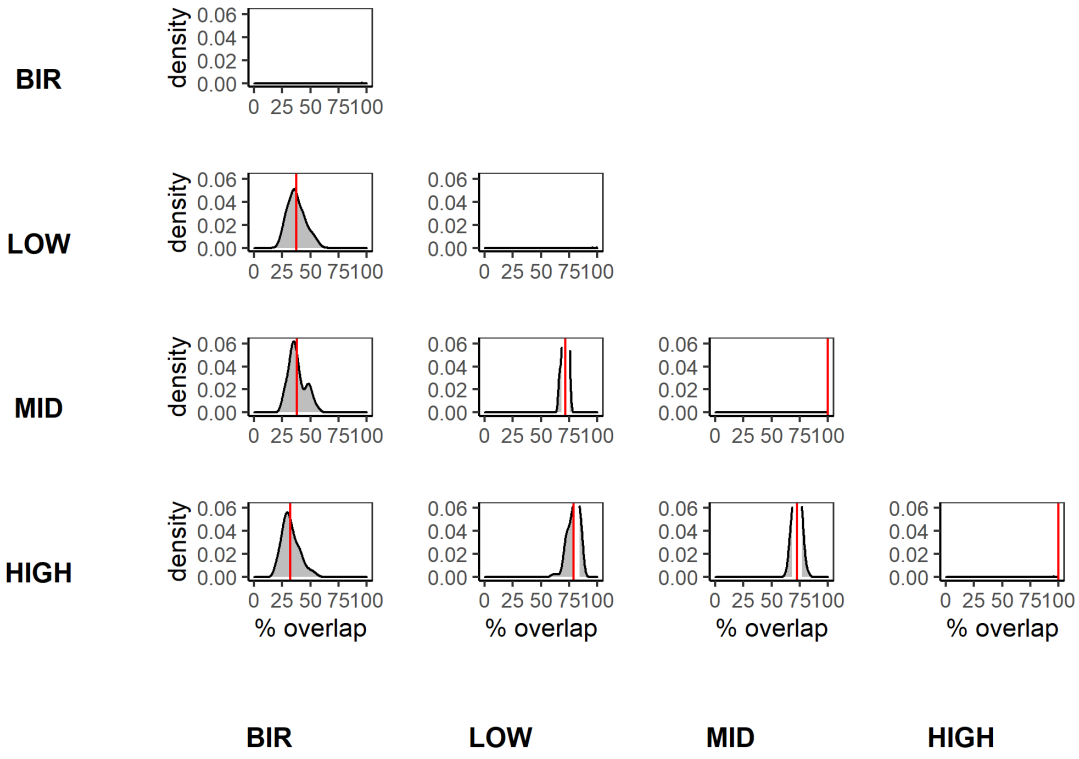

Figure B.2: Pairwise posterior ellipse overlap distributions between the experimental, early generation hybrid populations and the parental species, *X. birchmanni*.

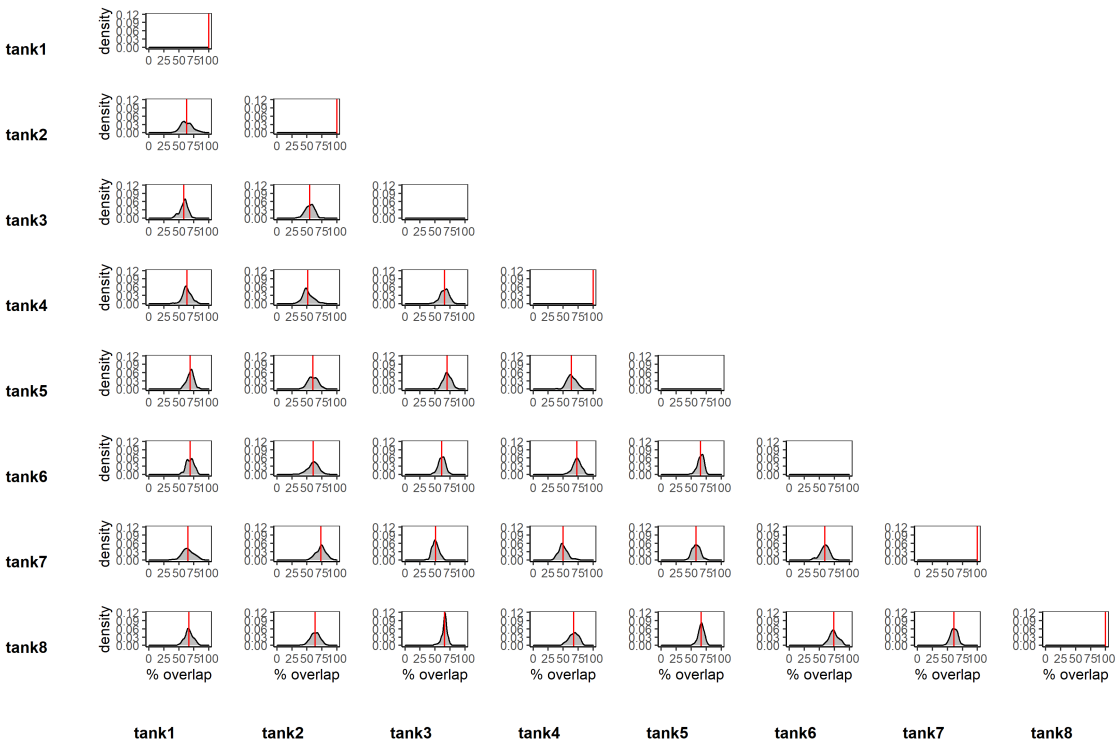

Figure B.3: Pairwise posterior ellipse overlap distributions between replicates within the HIGH experimental, early generation population.

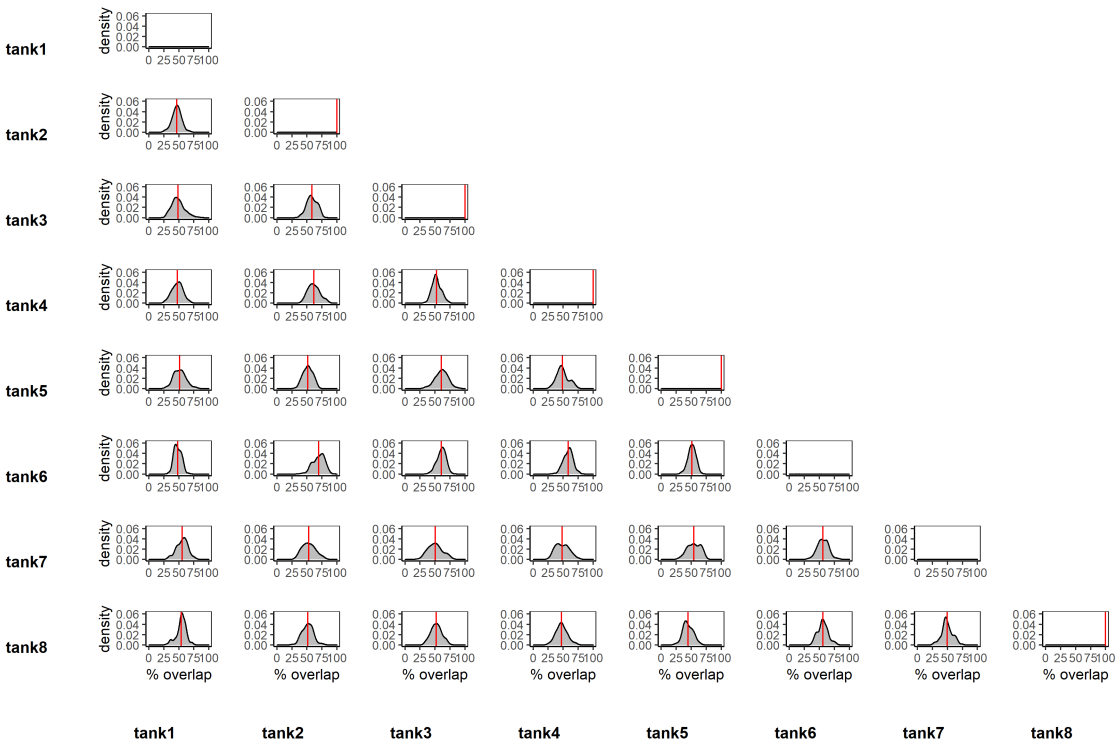

Figure B.4: Pairwise posterior ellipse overlap distributions between replicates within the MID experimental, early generation population.

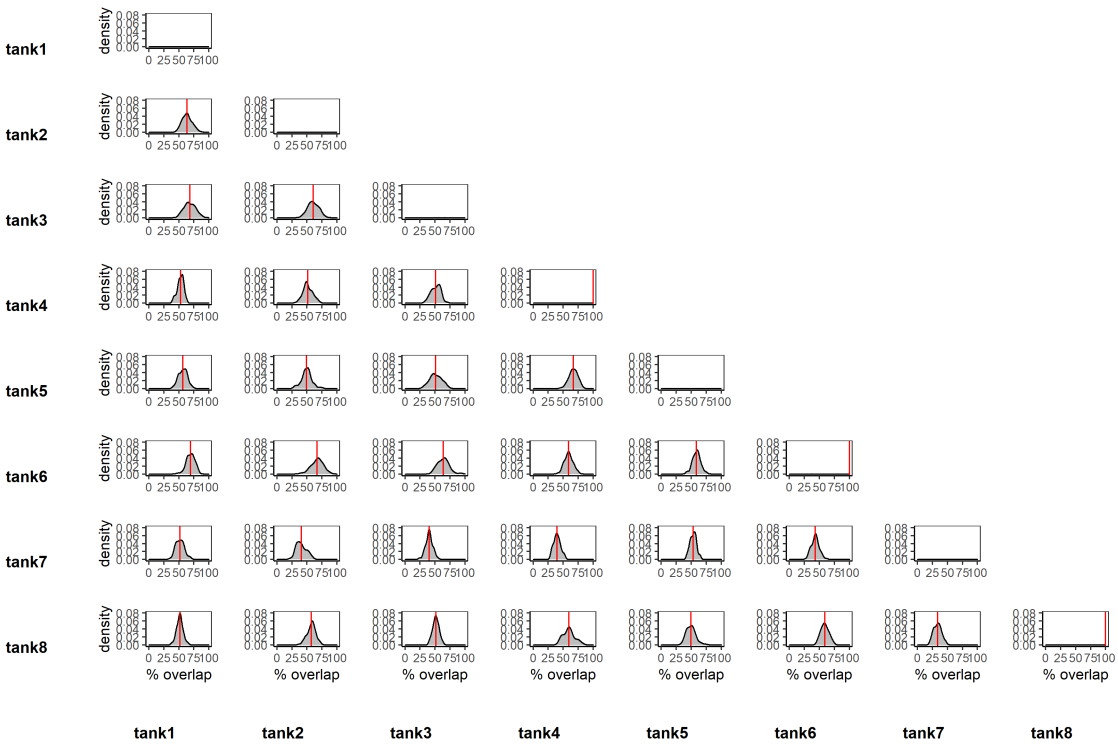

Figure B.5: Pairwise posterior ellipse overlap distributions between replicates within the LOW experimental, early generation population.

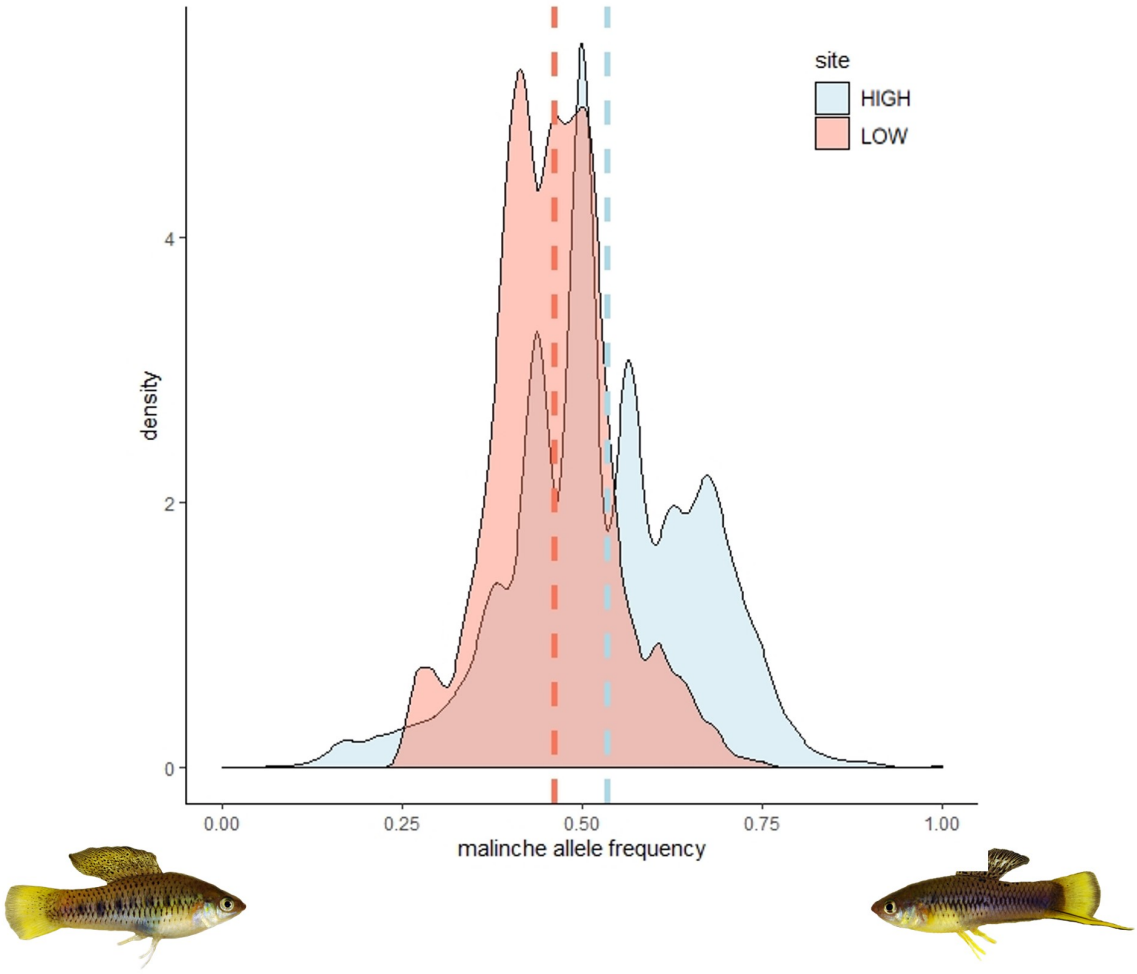

Figure B.6: Distibution of *X. malinche* ancestry for all ancestry informative markers among the most admixed individuals sequenced from the HIGH and LOW experimental, early generation hybrid populations. Dotted lines represent mean ancestry proportion among all individuals.

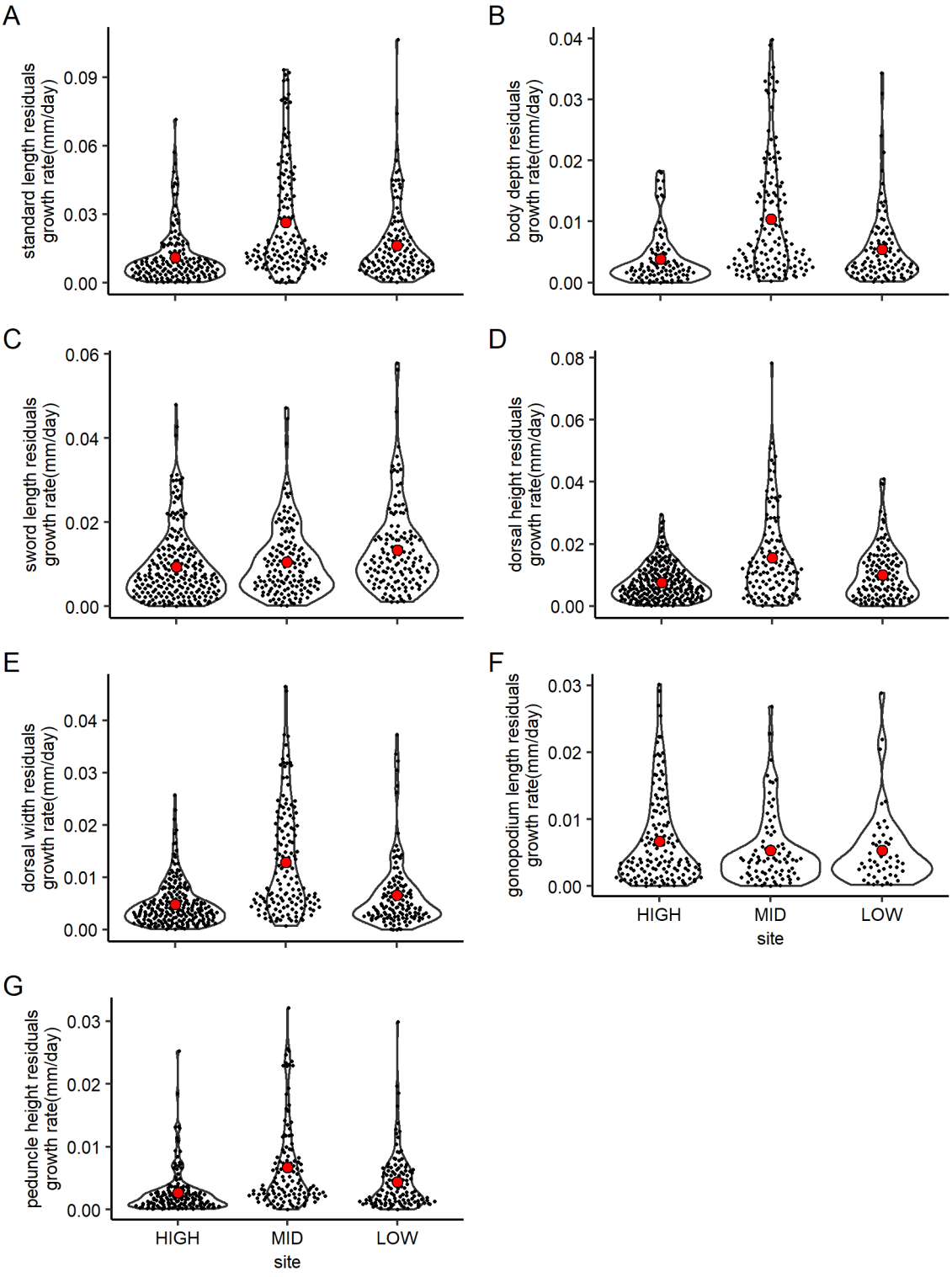

Figure B.7: Distribution of growth rates for A) stanard length, B) body depth residuals, C) sword length residuals, D) dorsal height residuals, E) dorsal width residuals, F) gonopodium length residuals, and G) peduncle height residuals in the experimental, early generation hybrid populations. Red points represent trait mean.

APPENDIX C

CHAPTER IV

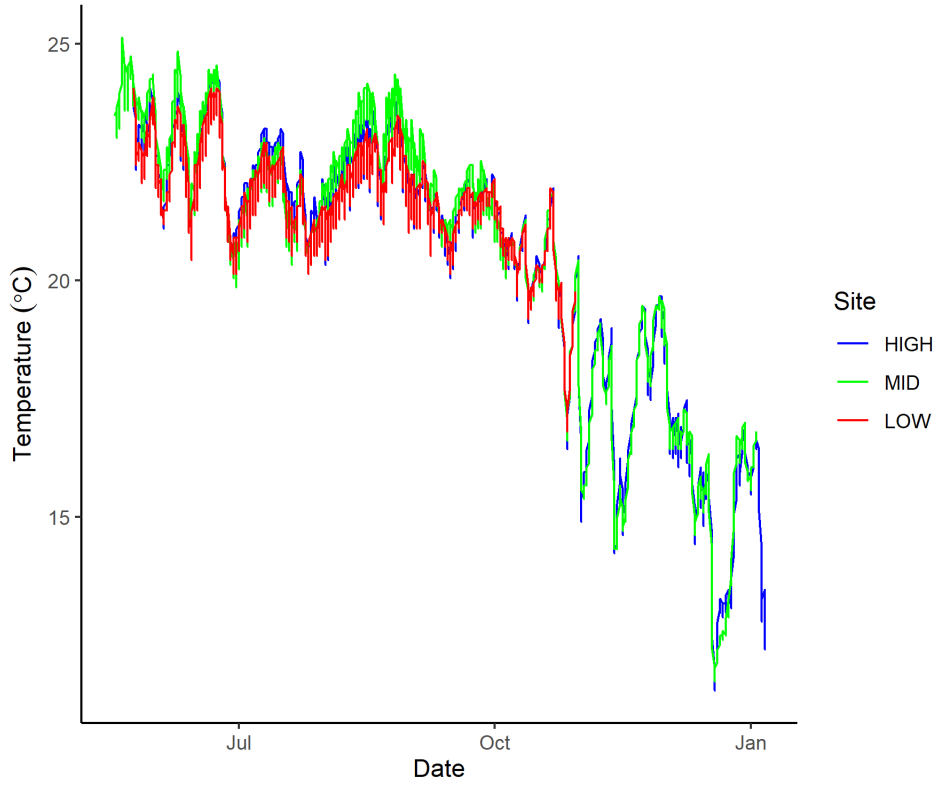

Figure C.1: Temperature profile of common garden tanks separated by site of origin.

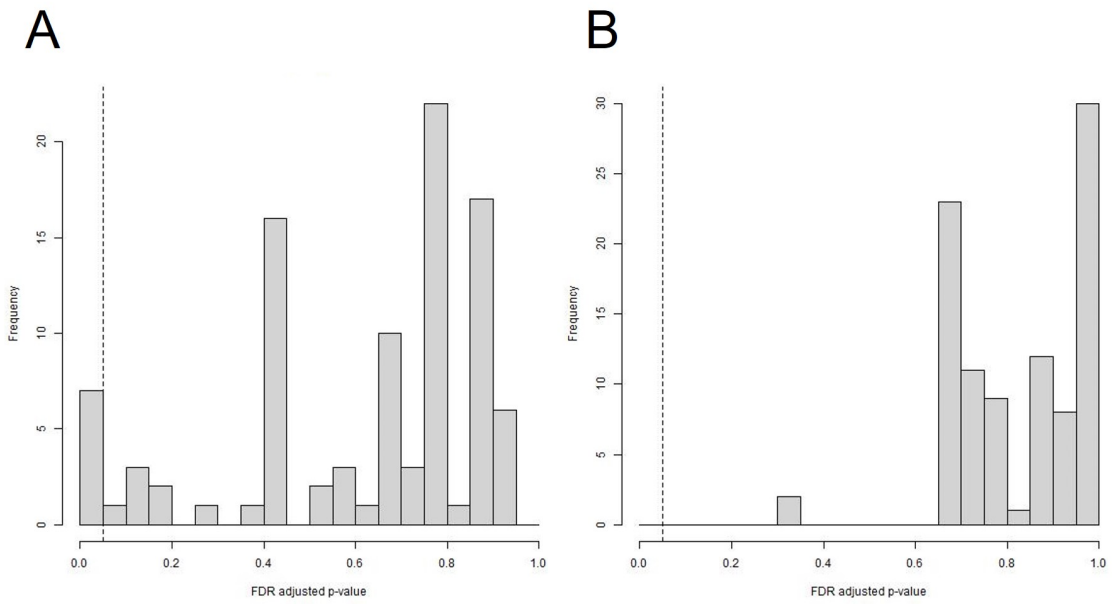

Figure C.2: Distribution of FDR adjusted p-values among natural hybrid populations A) ACUA and B) TLMC indicating whether the change in ancestry at *hsp* loci was significantly different from the genome-wide expectation.

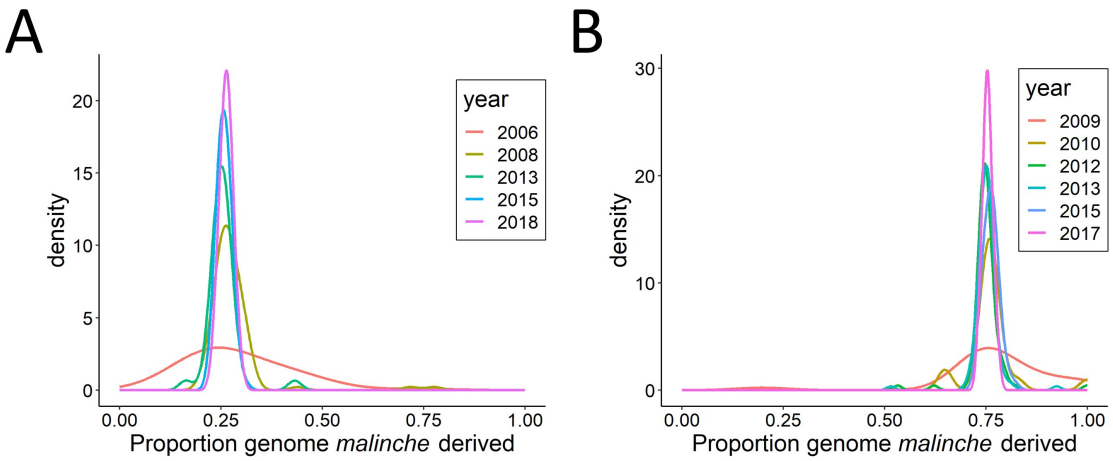

Figure C.3: Distribution of genome-wide hybrid index among natural hybrid populations A) ACUA and B) TLMC over time.

APPENDIX D

CHAPTER V

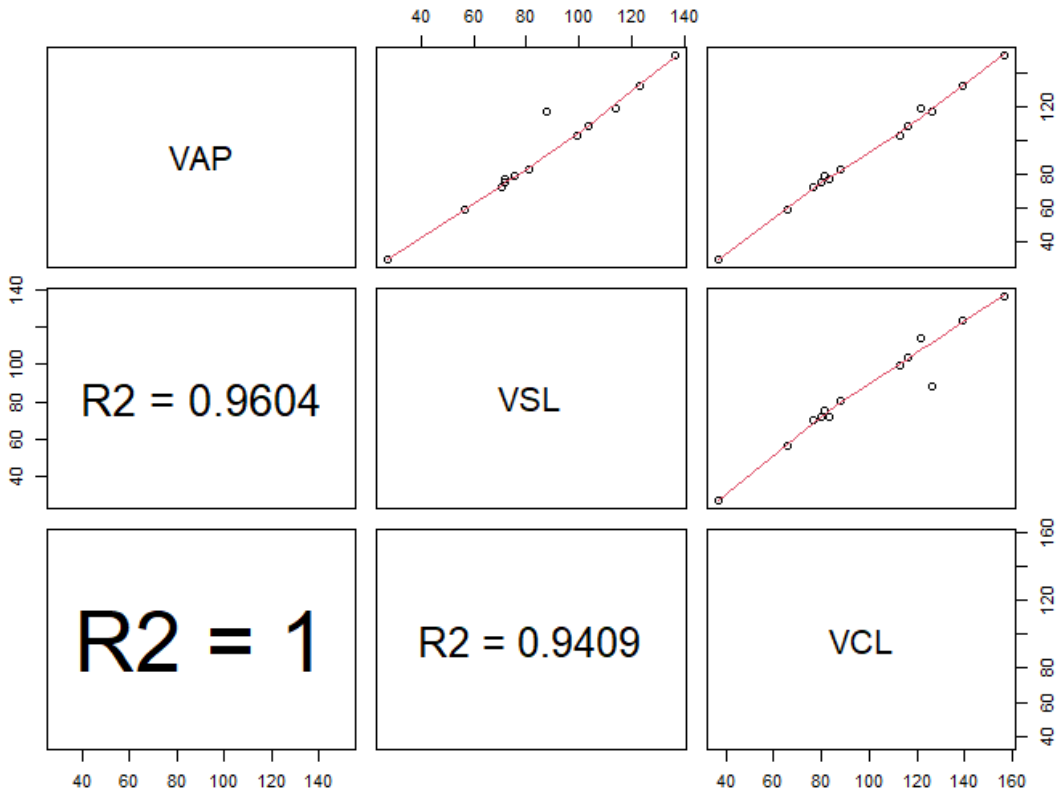

Figure D.1: Correlation matrix of sperm physiology velocity metrics.

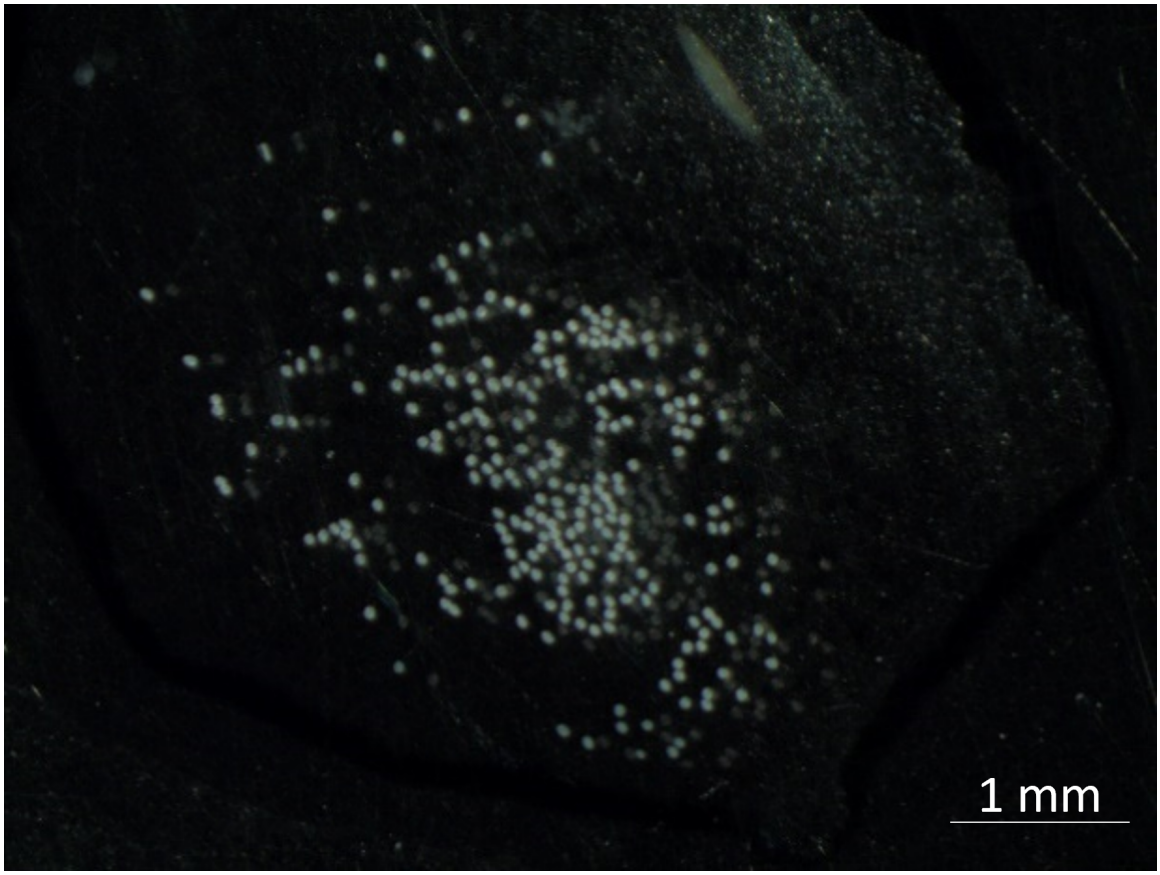

Figure D.2: Example photo of sperm bundles under a dissection scope. Each white dot represents a sperm bundle.

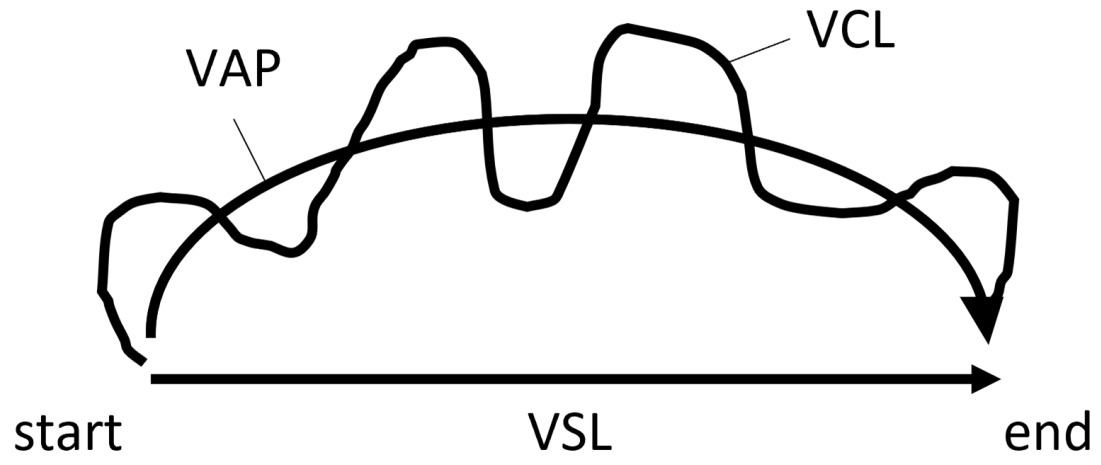

Figure D.3: Diagram of sperm physiology velocity metrics. VAP = average path velocity, VCL = curvilinear velocity, VSL = straight path velocity.

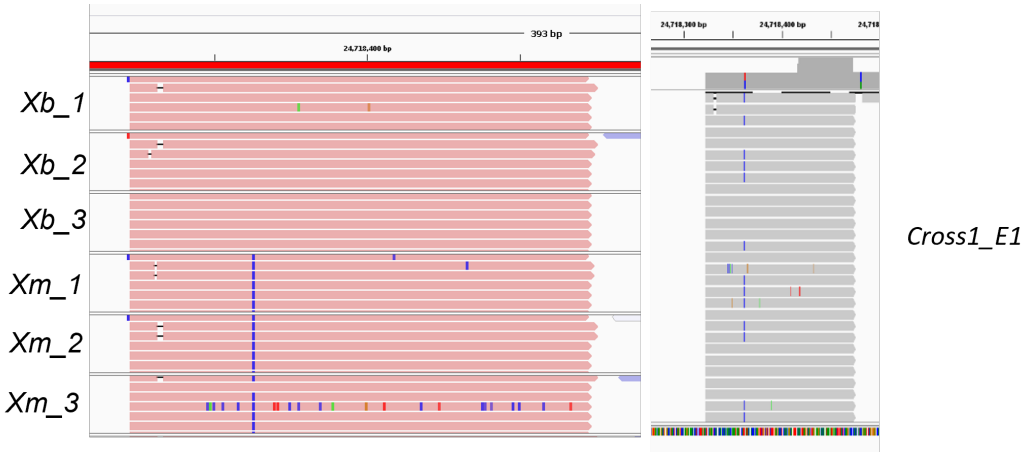

Figure D.4: Example of *X. birchmanni* and *X. malinche* genotypes at DNA Ligase I. The three *X. malinche* samples have a species-specific SNP. On the right is an example of a dissected embryo that was scored as a heterozygote at this locus.

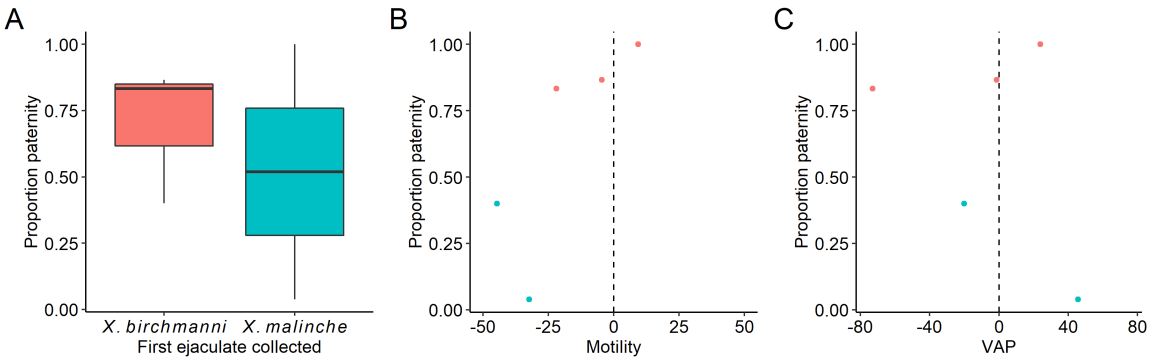

Figure D.5: Proportion of paternity sired as a function of A) which ejaculate was collected first, B) percent motility, and C) average path velocity.