

CAUSES AND IMPLICATIONS OF VARIATION IN REPRODUCTIVE ISOLATION
ACROSS *MIMULUS* MONKEYFLOWERS

by

MATTHEW C. FARNITANO

(Under the Direction of Andrea L. Sweigart)

ABSTRACT

Understanding the nature of species boundaries, including the mechanisms underlying their formation, maintenance, and breakdown, has long been a goal of evolutionary biology. Reproductive isolation between lineages is astoundingly variable across space, across time, across lineages, and across genomic loci, but we still have much to learn about the evolutionary, environmental, and genetic factors that shape this variation. Here, we use multiple species groups of *Mimulus* monkeyflowers as models to investigate the causes and consequences of variation in reproductive isolation at different scales. First, we characterize genomic patterns of divergence and components of postmating reproductive isolation within an understudied group of species in *Mimulus* section *Eunanus*. We find near-complete postmating isolation across all species pairs, driven by a combination of pollen-pistil incompatibility, hybrid seed inviability, and hybrid sterility. Consistent with these patterns, we find substantial genetic divergence and cryptic diversity, but signatures of ancient hybridization are still detectable. Next, we shift our focus to *Mimulus guttatus* and *Mimulus nasutus*, a recently diverged species pair with ongoing hybridization. We investigate spatial and temporal heterogeneity in patterns of hybrid ancestry across a decade of sampling, using genome-wide ancestry inference combined with measures of

prematuring isolation. We find that hybrid ancestry is pervasive, structured at incredibly small spatial scales, and stable across time. Phenological isolation fluctuates widely across years, likely driven by changes in water availability, with impacts on gene flow between admixture groups. Finally, we generate genomic data from a second area ~1000km away to examine parallelism across the genome in hybrid ancestry, and to ask how well reproductive barrier loci from controlled crosses can predict introgression outcomes in wild populations. We see differences in the extent of hybridization, but substantial parallelism in hybrid ancestry patterns across distant geographic areas. Known reproductive barrier loci are poor predictors of hybrid ancestry patterns, highlighting the importance of multiple complementary approaches to understanding drivers of reproductive isolation. Overall, these three studies demonstrate how careful characterization of reproductive barriers across multiple axes of variation can inform our understanding of the speciation process as it plays out in complex environments.

INDEX WORDS: Speciation; Hybridization; Reproductive isolation; Introgression; *Mimulus*

CAUSES AND IMPLICATIONS OF VARIATION IN REPRODUCTIVE ISOLATION
ACROSS *MIMULUS* MONKEYFLOWERS

by

MATTHEW C. FARNITANO

B.S. Biology, Duke University, 2016

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2024

© 2024

Matthew C. Farnitano

All Rights Reserved

CAUSES AND IMPLICATIONS OF VARIATION IN REPRODUCTIVE ISOLATION
ACROSS *MIMULUS* MONKEYFLOWERS

by

MATTHEW C. FARNITANO

Major Professor:	Andrea L. Sweigart
Committee:	Jill T. Anderson
	Casey M. Bergman
	Kelly A. Dyer
	Molly Schumer

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
December 2024

DEDICATION

For my daughter Elena.

ACKNOWLEDGEMENTS

I cannot begin to express sufficient gratitude for the countless mentors, colleagues, friends, and family members who have supported me throughout my graduate career and made this dissertation possible. To Andrea Sweigart, you have been incredibly supportive, boosting my confidence and championing me at every turn, while helping guide my meandering ideas and project concepts in fruitful and productive directions. You have also been understanding and flexible as I navigated the complexities of graduate school, pandemic chaos, and parenthood. Furthermore, your deep, clear thinking and knowledge of the classic literature has been an inspiration and guiding beacon for me as I develop my scientific mind.

Sam Mantel, Gabbie Sandstedt, and Makenzie Whitener, you were the friendly faces that made me feel welcome in the lab and at Georgia. You continued to be immensely supportive throughout our shared time in the lab, cheering on my successes and offering sage wisdom to keep me on track. Each of your absences has been felt as you all moved on to bigger things. Alex Sotola, in your time with our lab, you brought a positive energy and enthusiasm that I will not forget, especially during our field work excursions together.

Natalie Gonzalez, Logan Scott, and Eleanore Ritter, you have been the pep squad that steered me through the last miles of this journey, and you have quickly become my lifelong friends. Aside from providing levity, camaraderie, and pearls of insight, you have been invaluable logistically: caring for Elena, pet-sitting, checking on greenhouse plants, cutting pipette tips, baking delicious treats, and overall being present and available whenever you were needed.

The graduate students of the UGA Genetics department have been endless sources of insight, expertise, clever questions, fresh ideas, extra lab supplies, and friendly chats. I want to shout out in particular a few members of the evolution community: Matthew Treaster, Audrey Ward, Paul Ginsberg, Inam Jameel, Theresa Erlenbach, Pablo Mendieta, Daniel Shaw, Ben Long, Sam Day, Meghan Brady, Rebecca Piri... the list goes on.

Undergraduate researchers in our lab have put in many hours planting seeds, transplanting seedlings, collecting leaves, extracting DNA, entering data, and more, without which my dissertation research would not be possible. Sara Hill, Autumn Knight, Liza Lesley, Sherwin Shirazi, Parker Helms, Melat Mekonnen, and Erin Howard all contributed directly to this work.

To my dissertation committee: Jill Anderson, Casey Bergman, Kelly Dyer, and Molly Schumer, since day one you have encouraged and enabled me to fulfill my potential. Through formal and informal discussions, career advice and project feedback, you have steered me on a smooth path and set me up for success. I also want to thank the Genetics department staff, and in particular Susan White, who always has the administrative hiccups dealt with before you even know they exist.

A special thanks to Robin Hopkins and John Willis, mentors who set me on the academic path I am on. You got me to fall in love with wildflower evolution, showed me the beauty of evolutionary genetics, and convinced me to take on graduate school.

My parents, Lisa and Chris, have had my back since before I was born, and are the reason I had the confidence and ambition to pursue a PhD. You have always encouraged me to find a path that I am passionate and excited about, and to pursue it to the fullest. You have been a great

help with childcare and other life complications, and you are amazing grandparents to Elena. Likewise to my own grandparents, Tillie and Elliott, who are always my number one fans.

And now to my wife, Annie. Everyone deserves a partner as unconditionally supportive and unabashedly devoted as you. You have enthusiastically jumped into every challenge alongside me: moving to Georgia, planning and re-planning a pandemic wedding, buying a house while I was across the country doing fieldwork, giving birth and learning to raise our child together, being a single parent while I left for conferences, and helping me balance my time to finish this dissertation. No matter how a day of work has gone, I always look forward to seeing you at the end of it. You have earned this doctorate as much as I have.

Finally, to my wonderful daughter Elena. You have been my light and inspiration, my daily joy, and my reminder of the magic of life. While I can't say you've made this dissertation easier to complete, you have certainly made it worth doing. My hope is that you will go on to pursue challenging and fulfilling projects in your own life, and that you will look to me for a little bit of inspiration when you do.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
LIST OF TABLES	xi
LIST OF FIGURES	xiii
CHAPTER	
I INTRODUCTION	1
References	11
II STRONG POSTMATING REPRODUCTIVE ISOLATION IN <i>MIMULUS</i>	
SECTION <i>EUNANUS</i>	25
Abstract	26
Introduction	27
Materials and Methods	31
Results	42
Discussion	47
References	53
Tables	63
Figures	64
III FLUCTUATING REPRODUCTIVE ISOLATION AND STABLE ANCESTRY	
STRUCTURE IN A FINE-SCALED MOSAIC OF HYBRIDIZING <i>MIMULUS</i>	
MONKEYFLOWERS	71

Abstract	72
Introduction	73
Materials and Methods	78
Results	89
Discussion	98
References	104
Tables	120
Figures	122

IV CORRELATED GENOMIC PATTERNS OF INTROGRESSION ACROSS GEOGRAPHIC SPACE DESPITE CONTRASTING HYBRIDIZATION

HISTORIES	128
Abstract	129
Introduction	130
Materials and Methods	134
Results	145
Discussion	154
References	164
Tables	172
Figures	173

V CONCLUSION AND FUTURE DIRECTIONS	184
References	191

APPENDICES

A SUPPLEMENTARY TABLES AND FIGURES FROM CHAPTER II	195
--	-----

B	SUPPLEMENTARY TABLES AND FIGURES FROM CHAPTER III	214
C	SUPPLEMENTARY TABLES AND FIGURES FROM CHAPTER IV	227

LIST OF TABLES

	Page
Table 2.1: Summary of postmating reproductive isolation (RI)	63
Table 3.1: Mating system estimation	120
Table 3.2: Measurements of premating reproductive isolation	121
Table 4.1: Correlations between ancestry frequencies and gene density or recombination	172
Table 4.2: Summary of overlap between outlier windows across sample groups	172
Table S2.1: Population location information and number of samples used per population	195
Table S2.2: Sequencing and coverage information	196
Table S2.3: Pedigree information for <i>M. brevipes</i> and <i>M. fremontii</i> F1 families	197
Table S2.4: Nucleotide diversity and divergence values at synonymous sites	198
Table S2.5: ABBA-BABA test results for the ‘complete’ SNP dataset	199
Table S2.6: ABBA-BABA test results from four randomly downsampled datasets	200
Table S2.7: Window-based df and diversity metrics for introgression tests	201
Table S2.8: Crossing success and seed viability for each cross type	202
Table S2.9: Tetrazolium staining results	203
Table S2.10: Average seed measurements and sample sizes by cross type	204
Table S2.11: Pollen viability and counts with sample sizes	204
Table S2.12: Statistical model structures and AIC values compared to null model equivalents	205
Table S3.1: Summary of maternal and offspring samples by year, stream, and cohort	215
Table S3.2: Reference panel lines used for SNP creation	216

Table S3.3: Distribution of samples from each plot in each PCA cluster	217
Table S3.4: Statistical model output summaries	218
Table S3.5: Summary of selfing estimation from BORICE	220
Table S3.6: Summary of sibship estimation from BORICE	221
Table S4.1: Summary of collection locations used in this study	228
Table S4.2: Summary of sample sizes by location and admixture cohort	229
Table S4.3: Outlier window statistics for each sample group	230
Table S4.4: Outlier segment statistics for each sample group	231
Table S4.5: Comparison of outlier windows and outlier segments for each sample group.....	232

LIST OF FIGURES

	Page
Figure 2.1: Species and populations sampled in this study	64
Figure 2.2: Phylogenetic relationships and divergence between sampled species in <i>Mimulus</i> section <i>Eunanus</i>	65
Figure 2.3: Signatures of historical introgression between species in <i>Mimulus</i> section <i>Eunanus</i> .	66
Figure 2.4: Introgression outlier signals are distributed throughout the genome	67
Figure 2.5: Postmating prezygotic isolation between at least one species pair	68
Figure 2.6: Strong hybrid seed inviability between multiple species pairs	69
Figure 2.7: Strong hybrid male sterility in hybrids between <i>M. brevipes</i> and <i>M. fremontii</i>	70
Figure 3.1: Sampling locations and local variation in precipitation	122
Figure 3.2: Directional admixture shapes population structure in replicate streams	123
Figure 3.3: Complete chloroplast capture of a <i>M. nasutus</i> haplotype in sympatric <i>M. guttatus</i> .	124
Figure 3.4: Fine-scaled spatial and phenological structure across years	125
Figure 3.5: Assortative mating agrees with patterns of reproductive isolation	127
Figure 4.1: Geographic distribution of collection locations and admixture proportions	173
Figure 4.2: Distributions of ancestry frequencies across the genome for eight sample groups...	175
Figure 4.3: Correlations in ancestry frequencies across the genome between sample groups	177
Figure 4.4: Overlap in ancestry outlier windows across sample groups.....	178
Figure 4.5: Distributions of ancestry frequencies around reproductive barrier candidate loci....	179
Figure 4.6: Ancestry frequencies at fertility QTL and TRD loci.....	181

Figure 4.7: Overlap between ancestry outliers and fertility QTL or TRD loci.....	183
Figure S2.1: Comparison of phylogenetic methods.....	206
Figure S2.2: F_{branch} results for ‘downsampled’ iterations.....	208
Figure S2.3: TWISST gene tree discordance summaries	209
Figure S2.4: Summary of best-supported introgression events from multiple methods.....	211
Figure S2.5: Seed counts for intra- and interspecific crosses	212
Figure S2.6: Crossing success in BxF hybrids is additive on maternal and paternal sides	212
Figure S2.7: Seed measurements for all intra- and interspecific cross types.....	213
Figure S3.1: Correlation between hybrid index and NGSAdmix structure results.....	222
Figure S3.2: Ancestry heterozygosity vs. hybrid index for maternal and offspring plants	222
Figure S3.3: PC5 identification of <i>M. sookensis</i>	223
Figure S3.4: Complete mitochondrial capture of a <i>M. nasutus</i> haplotype in sympatric <i>M.</i> <i>guttatus</i>	224
Figure S3.5: Hybrid index at plot resolution for all three streams.....	225
Figure S3.6: Selfing rate vs. hybrid index	226
Figure S4.1: Ancestry heterozygosity in southern samples	233
Figure S4.2: Distributions of ancestry frequencies across chromosomes 03, 04, 05, and 06.....	234
Figure S4.3: Distributions of ancestry frequencies across chromosomes 09, 10, 11, and 12.....	235
Figure S4.4: Marker-by-marker ancestry correlations.....	236
Figure S4.5: Overlap in ancestry outlier segments across sample groups.....	237

CHAPTER I

INTRODUCTION

In Darwin's *On the Origin of Species* (1859), he proposed a common descent of all forms of life. But despite its title, the book left many unanswered questions about how species actually originate. It was not until the modern synthesis of evolution, heredity, and molecular genetics that scientists such as Bateson (1909), Dobzhansky (1937), Muller (1942), Mayr (1942), and Stebbins (1950) began to outline the genetic mechanisms by which a single lineage could split into multiple independent species. The foundations of species formation, they argued, lay in reproductive isolation: the inability of individuals from two distinct groups to mate and form viable, successful hybrids.

Reproductive isolation is a broad term that can refer to any intrinsic or extrinsic mechanism that prevents successful reproduction. We often enumerate two primary categories: premating isolation prevents individuals from forming hybrids at all, while postmating isolation prevents hybrids from surviving or successfully reproducing once formed. Premating barriers include geographic separation at either the landscape or microhabitat scale (ecogeographic isolation: Osmolovsky et al., 2022; Sobel, 2014), differences in the timing of reproduction (phenological isolation: Hood et al., 2019; Sianta & Kay, 2021), incompatible mating behaviors or mate preferences (behavioral isolation: Selz et al., 2014; Sullivan, 1995), physical inability to mate (mechanical isolation: Barnard et al., 2017; Richmond et al., 2011), and especially for plants, variation in pollinator preferences (pollinator isolation: Hopkins & Rausher, 2012; Schemske & Bradshaw, 1999) or self-fertilization mechanisms (mating system isolation: N. H.

Martin & Willis, 2007; Sianta et al., 2024). Postmating barriers include failures to fertilize after mating (cryptic female choice or postmating-prezygotic incompatibility: Feller et al., 2024; Firman et al., 2017), death during early development (hybrid inviability: Oneal et al., 2016; Presgraves et al., 2003), failure to thrive due to intrinsic incompatibilities (hybrid lethality or necrosis: Wright et al., 2013; Zuellig & Sweigart, 2018b), a poor match to the external environment (hybrid maladaptation: Melo et al., 2014; Richards & Ortiz-Barrientos, 2016) or failure of reproductive organs or gametes (hybrid sterility: Masly et al., 2006; Sweigart et al., 2006). Some hybrid incompatibilities are not obvious in the first generation, but appear in second- or later-generation hybrids (hybrid breakdown: Li et al., 1997; Stelkens et al., 2015).

Genetic mechanisms of isolation

Premating barriers may be completely extrinsic geographic barriers (Barraclough & Vogler, 2000), but often have a substantial genetic component, which is usually driven by direct environmental or sexual selection (DeMarche et al., 2015; Deutsch, 1997; Grossenbacher & Stanton, 2014; Hopkins & Rausher, 2012; Svensson et al., 2006), but could also arise due to drift or indirect selective mechanisms (Gavrilets & Boake, 1998). Many premating isolating barriers are charismatic traits that have captured the imaginations and attention of scientists: floral size and shape in plants (Hopkins & Rausher, 2012; Schemske & Bradshaw, 1999; Wessinger et al., 2023) or mating behaviors and decorative signals in animals (Arbuthnott, 2009; Irion & Nüsslein-Volhard, 2019; Selz et al., 2014). As a result, we know a great deal about the genetics underlying these traits, which often consist of a few genetic loci with large effects (Hopkins & Rausher, 2011; Irion & Nüsslein-Volhard, 2019; Schemske & Bradshaw, 1999; Streisfeld & Rausher, 2009). However, this finding may be biased – more charismatic and easily-studied traits may be more likely to be driven by large-effect loci. With a combination of direct selection,

large-effect loci, and consequences early in the reproductive cycle, premating barriers are thought to generally arise more quickly than postmating barriers, and may constitute the first steps on the road to speciation (Christie et al., 2022; Feder et al., 1994). However, they are also frequently mediated by interactions with the external environment, and therefore may be less stable sources of isolation in the long term as environmental conditions change (Freeman et al., 2022; Gilman & Behm, 2011; Irwin, 2020).

Postmating reproductive barriers are primarily described by the Dobzhansky-Muller model of epistatic incompatibility, in which two independently evolving lineages accumulate changes that are neutral or beneficial on their own, but have detrimental consequences when brought together in hybrid genomes (Dobzhansky, 1937; Muller, 1942; Orr & Turelli, 2001). This model provides a mechanism for how postmating isolation can arise despite its cost to hybrid fitness. Incompatibilities of this form are likely to arise during periods of limited contact after premating isolation is already in place, because selection can eliminate them if there is sufficient migration between groups (Behm et al., 2010; Xiong & Mallet, 2022). Once established, they are frequently independent of the environment, and so may act as a more permanent barrier promoting the ‘completion’ of the speciation process (Barton, 2020). While incompatibilities are often thought of as occurring through neutral processes, accumulating evidence suggests that many of the alleles involved in these incompatibilities arise in response to direct or indirect positive selection (Sweigart & Flagel, 2015; Wright et al., 2013). In particular, various forms of genomic conflict have been implicated in the evolution of postmating incompatibility (Brandvain & Haig, 2005; Crespi & Nosil, 2013; Murlas Cosmides & Tooby, 1981). For example, in both plants and placental mammals, conflict between maternal and paternal alleles over the optimal provisioning of the embryo with maternal nutrients can drive

developmental changes that result in hybrid incompatibilities (Brandvain & Haig, 2005; Soliman & Coughlan, 2024). Such conflict-based processes might drive more rapid evolution of postmating reproductive barriers, even allowing them to precede strong premating isolation (Coughlan et al., 2020; Coughlan & Matute, 2020; Sandstedt & Sweigart, 2022).

Patterns of reproductive isolation

A substantial body of literature has developed that attempts to address how quickly, in what order, and with what genetic mechanisms reproductive barriers tend to appear across a diverse sampling of species (Christie et al., 2022; Coyne & Orr, 1997; Lowry et al., 2008; Matute & Cooper, 2021). A few patterns have emerged: premating isolation often precedes postmating isolation, especially in sympatry (Christie et al., 2022; Coyne & Orr, 1997; Mendelson, 2003); sterility in species with sex chromosomes is often associated with the heterogametic sex and closely tied to genetic changes involving the sex chromosomes (Haldane, 1922; Macholán et al., 2007; Schilthuizen et al., 2011; Trier et al., 2014); the extent of pairwise incompatibilities is associated with overall genetic divergence and appears to accumulate at a faster-than-linear rate (Christie & Strauss, 2018; Guerrero et al., 2017; Matute et al., 2010; Moyle & Nakazato, 2010); and incompatibilities between the nuclear and organellar genomes are a common occurrence (Fishman & Willis, 2006; Gobron et al., 2013; Lee et al., 2008; Moran et al., 2024). However, underneath these general trends is an incredible diversity of outcomes. Incompatibilities arise at different rates in different groups (Malone & Fontenot, 2008; Moyle et al., 2004), and postmating barriers can sometimes emerge more quickly than premating barriers (Coughlan et al., 2020; Scopece et al., 2008). While we think of partial reproductive isolation as a stepping stone towards inevitable complete speciation, theoretical and observational work suggests that it can sometimes be a stable state in its own right (Servedio & Hermisson, 2020;

Sullivan, 1995). In addition, incompatibilities during early speciation are often polymorphic within species and even within populations, leading to different levels of isolation depending on the population and genotype examined (Cutter, 2012; Sweigart & Flagel, 2015; Zuellig & Sweigart, 2018a). Partial isolating barriers, both premating and postmating, can also be found between regions or populations that are otherwise considered ‘good’ compatible species (Corbett-Detig et al., 2013; N. H. Martin & Willis, 2010; Wright et al., 2013). As a community, we have only begun to understand the various genetic, ecological, demographic, and stochastic forces that lead to these diverse outcomes across organisms.

Hybridization and its prevalence

The counterpart to reproductive isolation is hybridization: when reproductive isolation is incomplete, hybridization between lineages can occur. It has become abundantly clear over the decades that reproductive isolation is often incomplete, even when species are clearly phenotypically, genetically, and ecologically distinct (Christie et al., 2022; Lotsy, 1931). As a result, hybridization among lineages is widespread throughout the tree of life, both between sister pairs and even at wider phylogenetic scales (Abbott et al., 2013; Goulet et al., 2017; Mallet et al., 2015). The impacts of hybridization vary widely, ranging from the complete collapse of species boundaries (Behm et al., 2010; Kleindorfer et al., 2014) to the reinforcement of reproductive barriers (Bank et al., 2012; Hopkins, 2013). Hybridization can increase genetic diversity and resilience (Brauer et al., 2023; P. R. Grant & Grant, 2019), or it can threaten extinction (Ayres et al., 2004; Brauer et al., 2023). If hybrids are viable and fertile, they can intermate and persist across generations as a hybrid swarm (Hasselman et al., 2014; Ruhsam et al., 2011). Alternatively, they may mate with parental species, leading to the movement of particular genomic regions from one species into the gene pool of the other; this process is

termed introgression (Askelson et al., 2023; Goodman et al., 1999; Randi & Lucchini, 2002).

Hybridization is often described in a cline or tension zone between two overlapping geographic ranges (Carneiro et al., 2013; Macholán et al., 2007; Streisfeld & Kohn, 2005). Hybrids can also form in a mosaic across the landscape when parental species ranges are largely overlapping, especially when parents are separated by adaptation to different microhabitats (Bacilieri et al., 1996; Ross & Harrison, 2002; Simon et al., 2021). Though hybrid outcomes are varied both among and within species pairs, the ecological and genetic drivers of these outcomes are only partially understood, particularly when reproductive isolation varies across space and time (Brice et al., 2021; Hamlin et al., 2020; Klein et al., 2017).

Consequences of hybridization in the genome

The effects of hybridization are not uniform across the genome. When introgression occurs, recombination breaks up blocks of ancestry over time, some of which survive in the recipient population and others which do not. The distribution of these blocks is determined in part by neutral factors such as the time since hybridization, the frequency of hybrids relative to the total population, and the stochastic effects of genetic drift (Sedghifar et al., 2016). These blocks are also shaped strongly by selection: either positive selection on adaptive alleles, which might be preferentially maintained in the population, or negative selection on incompatible or maladaptive alleles, which are quickly removed (Racimo et al., 2015; Sachdeva & Barton, 2018). Evidence from a variety of model systems shows that introgression is more likely to be retained in regions of the genome with higher recombination rates (Burri et al., 2015; Juric et al., 2016; S. H. Martin et al., 2019; Schumer et al., 2018). This is taken as evidence of predominantly negative selection against introgression, which acts to remove long blocks of introgression containing detrimental alleles in low-recombination areas, but will maintain smaller blocks of

introgression if they are able to quickly recombine away from any nearby incompatibility loci (Juric et al., 2016; Schumer et al., 2018). Introgression will also affect particular loci involved in reproductive isolation or adaptive sweeps, and can be used as a tool to identify important ‘speciation genes’ (Pardo-Diaz et al., 2012; Riquet et al., 2019; Runemark et al., 2018; Taylor et al., 2014; Wessinger et al., 2023), though intrapopulation demographic factors can sometimes be mistaken for these introgression patterns (Cruickshank & Hahn, 2014). Understanding the stability, predictability, and repeatability of these patterns is an ongoing project in evolutionary genetics (Chaturvedi et al., 2020; Langdon et al., 2024; Nouhaud et al., 2022).

Mimulus monkeyflowers as a model for speciation

The monkeyflowers of the genus *Mimulus* (Phrymaceae) have been used since Darwin (Darwin et al., 1876) as a model system for questions involving pollinator ecology (Schemske & Bradshaw, 1999; Streisfeld & Kohn, 2005), population genetics (Kelly, 2022; Mojica et al., 2012), speciation genetics (Stathos & Fishman, 2014; Zuellig & Sweigart, 2018a), hybridization and introgression (Kenney & Sweigart, 2016; Nelson et al., 2021), responses to climate change (McDonald et al., 2023), floral development (Yuan et al., 2013, 2016), and more (Twyford et al., 2015; Wu et al., 2008; Yuan, 2019). *Mimulus* consists of well over 100 wildflower species at various stages of reproductive isolation and divergence (Baldwin et al., 2012; A. L. Grant, 1924; Vickery, 1978), including widespread generalists (Twyford et al., 2020) and rare habitat specialists (Ferris et al., 2014). The group is notable for its variety of floral morphologies and pollinator strategies (Cooley & Willis, 2009; DeMarche et al., 2015; Grossenbacher & Stanton, 2014; Karron et al., 2004; Schemske & Bradshaw, 1999; Streisfeld & Rausher, 2009; Wu et al., 2013), its adaptation onto a variety of soil conditions (Ferris et al., 2014; Ferris & Willis, 2018; Gardner & Macnair, 2000; Lowry & Willis, 2010; Wright et al., 2013), and its repeated

chromosomal restructuring (Lowry & Willis, 2010; Stathos & Fishman, 2014; Vickery, 1995). In particular, the *Mimulus guttatus* species complex is an ideal system for studying speciation in action: multiple phylogenetically intertwined groups are ecologically differentiated and partially reproductively isolated by a variety of different extrinsic and intrinsic barriers (Brandvain et al., 2014; Coughlan et al., 2020; Fishman et al., 2002, 2015; Fishman & Willis, 2006; Gardner & Macnair, 2000; Ivey et al., 2023; Kenney & Sweigart, 2016; Mantel & Sweigart, 2019; Sweigart et al., 2007; Tataru et al., 2023; Whitener et al., 2024; Wright et al., 2013; Zuellig & Sweigart, 2018a). *Mimulus guttatus* has one of the highest measured intra-population nucleotide diversity values of any species (Puzey et al., 2017), greater than the divergence between humans and gorillas (Scally et al., 2012). As an example of the genomic complexity of the group, northern and southern clades of *M. guttatus*, though largely compatible, are as divergent at the nucleotide level as either is to the derived selfing species *Mimulus nasutus* (Brandvain et al., 2014). In this dissertation, I take advantage of the immense diversity both across the *Mimulus* genus and within the well-studied speciation crucible of the *Mimulus guttatus* species complex, in order to answer globally relevant questions about the nature of forces shaping the speciation process.

A note on taxonomy and nomenclature: molecular evidence suggests that the genus *Mimulus* as broadly conscribed is paraphyletic (Beardsley et al., 2004), and recent treatments have proposed to split the group into multiple monophyletic genera, resulting in nomenclatural changes for the majority of taxa in the group (Barker et al., 2012). Evolutionary biologists have largely continued to use *Mimulus* in the broad sense, retaining names that have a long history in the evolutionary literature (Lowry et al., 2019; but see response to these concerns by Nesom et al., 2019). Likewise, I use *Mimulus* in the broad sense throughout this dissertation; if you prefer the alternative system, the species in section *Eunanus* discussed in Chapter II can be referred to

by the genus name *Diplacus*, while *Mimulus guttatus* and *Mimulus nasutus*, the focal species in Chapters III and IV, can be referred to as *Erythranthe guttata* and *Erythranthe nasuta*.

Three studies on variation in reproductive isolation within monkeyflowers

Throughout this dissertation, I use careful characterization of reproductive isolation, and its consequences for hybridization and introgression, to draw inferences about the process of speciation. I focus on how reproductive isolation varies among species, among populations, among individuals, and among genetic loci. By examining patterns of variation in reproductive isolation, I tackle big-picture questions about what intrinsic and extrinsic forces shape this variation across the tree of life.

In Chapter II, I characterize patterns of postmating reproductive isolation and genetic divergence within a group of charismatic but understudied monkeyflowers in *Mimulus* section *Eunanus*. I find that, in contrast to some other model *Mimulus* groups, intrinsic reproductive isolation is almost complete. Despite strong contemporary divergence, genomic patterns suggest a complex history of past gene flow, underscoring the dynamic nature of reproductive isolation across evolutionary time. Hybrid seed inviability is a key source of reproductive isolation in this group, adding to a growing body of work suggesting it is a widespread and early-evolving barrier across *Mimulus*. However, I argue that seed inviability in this system may be a byproduct of extrinsic environmental selection on seed size, a factor not implicated in other cases of seed inviability in *Mimulus*.

In Chapter III, I investigate the dynamics of hybridization and reproductive isolation within a group of admixed populations of *Mimulus guttatus* and *Mimulus nasutus*. Using a decade of sampling with fine spatial resolution, I ask how hybrid ancestry is structured within and between populations, how that structure changes over time, and how premating reproductive

barriers influence these dynamics. I find that admixed populations are highly structured by hybrid ancestry, with multiple distinct cohorts of ancestry at both fine (~20-50m) and coarser (~4km) resolution across the landscape. These cohorts are stable across time and partially isolated by a combination of mating system, phenology, and microhabitat. However, the degree of isolation, especially for phenology, is highly variable across years, likely in response to changes in water availability. I argue that this heterogeneity across years may help to maintain a dynamic equilibrium of distinct but interconnected ancestry cohorts, favoring complex hybridization outcomes over simple population homogenization. These findings have implications for understanding population responses to climate change, which is predicted to increase year-to-year variability in environmental conditions.

In Chapter IV, I compare genomic patterns of ancestry across hybridizing populations of *Mimulus guttatus* and *Mimulus nasutus* in two geographic areas separated by ~1000km. I tackle two major questions in the hybridization field: how repeatable are genomic patterns of hybridization, and how well do laboratory studies of reproductive isolation predict these genomic patterns. I find that genomic ancestry patterns from admixed and sympatric populations are predictive of nearby less-admixed or allopatric populations, providing a window into the process of adaptive introgression. Ancestry patterns are also similar across geographic space, though less so at larger distances, indicating shared genomic features and similar selective pressures operating in parallel. This is true even though overall admixture levels are quite different between study areas. I find a surprisingly low correspondence between laboratory-identified reproductive isolation loci and wild introgression patterns. I argue that this is a byproduct of high genetic diversity, large populations, and recent time since divergence, which has made the genetic architecture of reproductive isolation highly polygenetic and geographically localized.

This provides a contrast with some other study systems that have clear genomic islands of divergence or adaptive sweeps of introgression.

References

- Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C. A., Buggs, R., Butlin, R. K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S. H., Hermansen, J. S., Hewitt, G., Hudson, A. G., Jiggins, C., ... Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, 26(2), 229–246. <https://doi.org/10.1111/j.1420-9101.2012.02599.x>
- Arbuthnott, D. (2009). The genetic architecture of insect courtship behavior and premating isolation. *Heredity*, 103(1), 15–22. <https://doi.org/10.1038/hdy.2009.22>
- Askelson, K. K., Spellman, G. M., & Irwin, D. (2023). Genomic divergence and introgression between cryptic species of a widespread North American songbird. *Molecular Ecology*, 32(24), 6839–6853. <https://doi.org/10.1111/mec.17169>
- Ayres, D. R., Zaremba, K., & Strong, D. R. (2004). Extinction of a Common Native Species by Hybridization with an Invasive Congener. *Weed Technology*, 18(1), 1288–1291. [https://doi.org/10.1614/0890-037X\(2004\)018\[1288:EOACNS\]2.0.CO;2](https://doi.org/10.1614/0890-037X(2004)018[1288:EOACNS]2.0.CO;2)
- Bacilieri, R., Ducousso, A., Petit, R. J., & Kremer, A. (1996). Mating system and asymmetric hybridization in a mixed stand of european oaks. *Evolution*, 50(2), 900–908. <https://doi.org/10.1111/j.1558-5646.1996.tb03898.x>
- Baldwin, B. G., Goldman, D., Keil, D. J., Patterson, R., Rosatti, T. J., & Wilken, D. (Eds.). (2012). *The Jepson Manual: Vascular Plants of California, Thoroughly Revised and Expanded* (2nd ed.).
- Bank, C., Hermisson, J., & Kirkpatrick, M. (2012). Can reinforcement complete speciation? *Evolution*, 66(1), 229–239. <https://doi.org/10.1111/j.1558-5646.2011.01423.x>
- Barker, W. R. (Bill), Nesom, G. L., Beardsley, P. M., & Fraga, N. S. (2012). A taxonomic conspectus of phrymaceae: A narrowed circumscription for. *Phytoneuron*, May, 1–60.
- Barnard, A. A., Fincke, O. M., McPeck, M. A., & Masly, J. P. (2017). Mechanical and tactile incompatibilities cause reproductive isolation between two young damselfly species: Mating structures and reproductive isolation. *Evolution*, 71(10), 2410–2427. <https://doi.org/10.1111/evo.13315>
- Barracough, T. G., & Vogler, A. P. (2000). Detecting the Geographical Pattern of Speciation from Species-Level Phylogenies. *The American Naturalist*, 155(4), 419–434. <https://doi.org/10.1086/303332>

- Barton, N. H. (2020). On the completion of speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1806), 20190530. <https://doi.org/10.1098/rstb.2019.0530>
- Bateson, W. (1909). Heredity and Variation in Modern Lights. In A. C. Seward (Ed.), *Darwin and Modern Science* (pp. 85–101). Cambridge University Press.
- Beardsley, P. M., Schoenig, S. E., Whittall, J. B., & Olmstead, R. G. (2004). Patterns of evolution in western North American *Mimulus* (Phrymaceae). *American Journal of Botany*, 91(3), 474–489. <https://doi.org/10.3732/ajb.91.3.474>
- Behm, J. E., Ives, A. R., & Boughman, J. W. (2010). Breakdown in postmating isolation and the collapse of a species pair through hybridization. *The American Naturalist*, 175(1), 11–26. <https://doi.org/10.1086/648559>
- Brandvain, Y., & Haig, D. (2005). Divergent Mating Systems and Parental Conflict as a Barrier to Hybridization in Flowering Plants. *The American Naturalist*, 166(3), 330–338. <https://doi.org/10.1086/432036>
- Brandvain, Y., Kenney, A. M., Flagel, L., Coop, G., & Sweigart, A. L. (2014). Speciation and Introgression between *Mimulus nasutus* and *Mimulus guttatus*. *PLoS Genetics*, 10(6), e1004410. <https://doi.org/10.1371/journal.pgen.1004410>
- Brauer, C. J., Sandoval-Castillo, J., Gates, K., Hammer, M. P., Unmack, P. J., Bernatchez, L., & Beheregaray, L. B. (2023). Natural hybridization reduces vulnerability to climate change. *Nature Climate Change*, 13(3), 282–289. <https://doi.org/10.1038/s41558-022-01585-1>
- Brice, C., Zhang, Z., Bendixsen, D., & Stelkens, R. (2021). Hybridization Outcomes Have Strong Genomic and Environmental Contingencies. *The American Naturalist*, 198(3), E53–E67. <https://doi.org/10.1086/715356>
- Burri, R., Nater, A., Kawakami, T., Mugal, C. F., Olason, P. I., Smeds, L., Suh, A., Dutoit, L., Bureš, S., Garamszegi, L. Z., Hogner, S., Moreno, J., Qvarnström, A., Ružić, M., Sæther, S.-A., Sætre, G.-P., Török, J., & Ellegren, H. (2015). Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of *Ficedula* flycatchers. *Genome Research*, 25(11), 1656–1665. <https://doi.org/10.1101/gr.196485.115>
- Carneiro, M., Baird, S. J. E., Afonso, S., Ramirez, E., Tarroso, P., Teotónio, H., Villafuerte, R., Nachman, M. W., & Ferrand, N. (2013). Steep clines within a highly permeable genome across a hybrid zone between two subspecies of the European rabbit. *Molecular Ecology*, 22(9), 2511–2525. <https://doi.org/10.1111/mec.12272>
- Chaturvedi, S., Lucas, L. K., Buerkle, C. A., Fordyce, J. A., Forister, M. L., Nice, C. C., & Gompert, Z. (2020). Recent hybrids recapitulate ancient hybrid outcomes. *Nature Communications*, 11(1), Article 1. <https://doi.org/10.1038/s41467-020-15641-x>

- Christie, K., Fraser, L. S., & Lowry, D. B. (2022). The strength of reproductive isolating barriers in seed plants: Insights from studies quantifying premating and postmating reproductive barriers over the past 15 years. *Evolution*, evo.14565. <https://doi.org/10.1111/evo.14565>
- Christie, K., & Strauss, S. Y. (2018). Along the speciation continuum: Quantifying intrinsic and extrinsic isolating barriers across five million years of evolutionary divergence in California jewelfloweers. *Evolution*, 72(5), 1063–1079. <https://doi.org/10.1111/evo.13477>
- Cooley, A. M., & Willis, J. H. (2009). Genetic Divergence Causes Parallel Evolution of Flower Color in Chilean Mimulus. *The New Phytologist*, 183(3), 729–739.
- Corbett-Detig, R. B., Zhou, J., Clark, A. G., Hartl, D. L., & Ayroles, J. F. (2013). Genetic incompatibilities are widespread within species. *Nature*, 504(7478), 135–137. <https://doi.org/10.1038/nature12678>
- Coughlan, J. M., & Matute, D. R. (2020). The importance of intrinsic postzygotic barriers throughout the speciation process: Intrinsic barriers throughout speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1806). <https://doi.org/10.1098/rstb.2019.0533>
- Coughlan, J. M., Wilson Brown, M., & Willis, J. H. (2020). Patterns of Hybrid Seed Inviability in the Mimulus guttatus sp. Complex Reveal a Potential Role of Parental Conflict in Reproductive Isolation. *Current Biology*, 30(1), 83-93.e5. <https://doi.org/10.1016/j.cub.2019.11.023>
- Coyne, J. A., & Orr, H. A. (1997). Patterns of Speciation in Drosophila Revisited. *Evolution*, 51(1), 295. <https://doi.org/10.2307/2410984>
- Crespi, B., & Nosil, P. (2013). Conflictual speciation: Species formation via genomic conflict. *Trends in Ecology & Evolution*, 28(1), 48–57. <https://doi.org/10.1016/j.tree.2012.08.015>
- Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, 23(13), 3133–3157. <https://doi.org/10.1111/mec.12796>
- Cutter, A. D. (2012). The polymorphic prelude to Bateson-Dobzhansky-Muller incompatibilities. *Trends in Ecology and Evolution*, 27(4), 209–218. <https://doi.org/10.1016/j.tree.2011.11.004>
- Darwin, C. (1859). *On the origin of species by means of natural selection*. John Murray. http://darwin-online.org.uk/converted/pdf/1861_OriginNY_F382.pdf
- Darwin, C., Darwin, C., & Edinburgh, R. C. of P. of. (1876). *The effects of cross and self fertilisation in the vegetable kingdom*. John Murray. <https://doi.org/10.5962/bhl.title.110800>

- DeMarche, M. L., Miller, T. J., & Kay, K. M. (2015). An ultraviolet floral polymorphism associated with life history drives pollinator discrimination in *Mimulus guttatus*. *American Journal of Botany*, 102(3), 396–406. <https://doi.org/10.3732/ajb.1400415>
- Deutsch, J. C. (1997). Colour diversification in Malawi cichlids: Evidence for adaptation, reinforcement or sexual selection? *Biological Journal of the Linnean Society*, 62(1), 1–14. <https://doi.org/10.1006/BIJL.1997.0135>
- Dobzhansky, T. (1937). *Genetics and the Origin of Species*. Columbia University Press.
- Feder, J. L., Opp, S. B., Wlazlo, B., Reynolds, K., Go, W., Spisak, S., Gavrilovic, V., Filchak, K. E., Rull, J., & Aluja, M. (1994). Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. *Proceedings of the National Academy of Sciences of the United States of America*, 91(17), 7990–7994. <https://doi.org/10.1073/pnas.91.17.7990>
- Feller, A. F., Burgin, G., Lewis, N. F., Prabhu, R., & Hopkins, R. (2024). Mismatch between pollen and pistil size causes asymmetric mechanical reproductive isolation across *Phlox* species. *Evolution*, qpae128. <https://doi.org/10.1093/evolut/qpae128>
- Ferris, K. G., Sexton, J. P., & Willis, J. H. (2014). Speciation on a local geographic scale: The evolution of a rare rock outcrop specialist in *Mimulus*. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1648), 20140001. <https://doi.org/10.1098/rstb.2014.0001>
- Ferris, K. G., & Willis, J. H. (2018). Differential adaptation to a harsh granite outcrop habitat between sympatric *Mimulus* species. *Evolution*, 72(6), 1225–1241. <https://doi.org/10.1111/evo.13476>
- Firman, R. C., Gasparini, C., Manier, M. K., & Pizzari, T. (2017). Postmating Female Control: 20 Years of Cryptic Female Choice. *Trends in Ecology & Evolution*, 32(5), 368–382. <https://doi.org/10.1016/j.tree.2017.02.010>
- Fishman, L., Beardsley, P. M., Stathos, A., Williams, C. F., & Hill, J. P. (2015). The genetic architecture of traits associated with the evolution of self-pollination in *Mimulus*. *New Phytologist*, 205(2), 907–917. <https://doi.org/10.1111/nph.13091>
- Fishman, L., Kelly, A. J., & Willis, J. H. (2002). Minor quantitative trait loci underlie floral traits associated with mating system divergence in *Mimulus*. *Evolution*, 56(11), 2138–2155. <https://doi.org/10.1111/j.0014-3820.2002.tb00139.x>
- Fishman, L., & Willis, J. H. (2006). A cytonuclear incompatibility causes anther sterility in *Mimulus* hybrids. *Evolution*, 60(7), 1372–1381. <https://doi.org/10.1111/j.0014-3820.2006.tb01216.x>
- Freeman, B. G., Rolland, J., Montgomery, G. A., & Schluter, D. (2022). Faster evolution of a premating reproductive barrier is not associated with faster speciation rates in New World

- passerine birds. *Proceedings of the Royal Society B: Biological Sciences*, 289(1966), 20211514. <https://doi.org/10.1098/rspb.2021.1514>
- Gardner, M., & Macnair, M. (2000). Factors affecting the co-existence of the serpentine endemic *Mimulus nudatus* Curran and its presumed progenitor, *Mimulus guttatus* Fischer ex DC. *Biological Journal of the Linnean Society*, 69(4), 443–459. <https://doi.org/10.1006/bijl.1999.0387>
- Gavrilets, S., & Boake, C. R. B. (1998). On the Evolution of Premating Isolation after a Founder Event. *The American Naturalist*, 152(5), 706–716. <https://doi.org/10.1086/286201>
- Gilman, R. T., & Behm, J. E. (2011). Hybridization, Species Collapse, and Species Reemergence After Disturbance to Premating Mechanisms of Reproductive Isolation. *Evolution*, 65(9), 2592–2605. <https://doi.org/10.1111/j.1558-5646.2011.01320.x>
- Gobron, N., Waszczak, C., Simon, M., Hiard, S., Boivin, S., Charif, D., Ducamp, A., Wenes, E., & Budar, F. (2013). A Cryptic Cytoplasmic Male Sterility Unveils a Possible Gynodioecious Past for *Arabidopsis thaliana*. *PLoS ONE*, 8(4). <https://doi.org/10.1371/journal.pone.0062450>
- Goodman, S. J., Barton, N. H., Swanson, G., Abernethy, K., & Pemberton, J. M. (1999). Introgression Through Rare Hybridization: A Genetic Study of a Hybrid Zone Between Red and Sika Deer (Genus *Cervus*) in Argyll, Scotland. *Genetics*, 152(1), 355–371. <https://doi.org/10.1093/genetics/152.1.355>
- Goulet, B. E., Roda, F., & Hopkins, R. (2017). Hybridization in plants: Old ideas, new techniques. *Plant Physiology*, 173(1), 65–78. <https://doi.org/10.1104/pp.16.01340>
- Grant, A. L. (1924). A Monograph of the Genus *Mimulus*. *Annals of the Missouri Botanical Garden*, 11(2/3), 99–388. <https://doi.org/10.2307/2394024>
- Grant, P. R., & Grant, B. R. (2019). Hybridization increases population variation during adaptive radiation. *Proceedings of the National Academy of Sciences*, 116(46), 23216–23224. <https://doi.org/10.1073/pnas.1913534116>
- Grossenbacher, D. L., & Stanton, M. L. (2014). Pollinator-mediated competition influences selection for flower-color displacement in sympatric monkeyflowers. *American Journal of Botany*, 101(11), 1915–1924. <https://doi.org/10.3732/ajb.1400204>
- Guerrero, R. F., Muir, C. D., Josway, S., & Moyle, L. C. (2017). Pervasive antagonistic interactions among hybrid incompatibility loci. *PLOS Genetics*, 13(6), e1006817. <https://doi.org/10.1371/journal.pgen.1006817>
- Haldane, J. B. S. (1922). Sex ratio and unisexual sterility in hybrid animals. *Journal of Genetics*, 12(2), 101–109. <https://doi.org/10.1007/BF02983075>

- Hamlin, J. A. P., Hibbins, M. S., & Moyle, L. C. (2020). Assessing biological factors affecting postspeciation introgression. *Evolution Letters*, 4(2), 137–154. <https://doi.org/10.1002/evl3.159>
- Hasselman, D. J., Argo, E. E., McBride, M. C., Bentzen, P., Schultz, T. F., Perez-Umphrey, A. A., & Palkovacs, E. P. (2014). Human disturbance causes the formation of a hybrid swarm between two naturally sympatric fish species. *Molecular Ecology*, 23(5), 1137–1152. <https://doi.org/10.1111/mec.12674>
- Hood, G. R., Zhang, L., Hu, E. G., Ott, J. R., & Egan, S. P. (2019). Cascading reproductive isolation: Plant phenology drives temporal isolation among populations of a host-specific herbivore. *Evolution*, 73(3), 554–568. <https://doi.org/10.1111/evo.13683>
- Hopkins, R. (2013). Reinforcement in plants. *New Phytologist*, 197(4), 1095–1103. <https://doi.org/10.1111/nph.12119>
- Hopkins, R., & Rausher, M. D. (2011). Identification of two genes causing reinforcement in the Texas wildflower *Phlox drummondii*. *Nature*, 469(7330), 411–414. <https://doi.org/10.1038/nature09641>
- Hopkins, R., & Rausher, M. D. (2012). Pollinator-mediated selection on flower color allele drives reinforcement. *Science*, 335(6072), 1090–1092. <https://doi.org/10.1126/science.1215198>
- Irion, U., & Nüsslein-Volhard, C. (2019). The identification of genes involved in the evolution of color patterns in fish. *Current Opinion in Genetics & Development*, 57, 31–38. <https://doi.org/10.1016/J.GDE.2019.07.002>
- Irwin, D. E. (2020). Assortative Mating in Hybrid Zones Is Remarkably Ineffective in Promoting Speciation. *The American Naturalist*, 195(6), E150–E167. <https://doi.org/10.1086/708529>
- Ivey, C. T., Habecker, N. M., Bergmann, J. P., Ewald, J., Frayer, M. E., & Coughlan, J. M. (2023). Weak reproductive isolation and extensive gene flow between *Mimulus glaucescens* and *M. guttatus* in northern California. *Evolution*, 77(5), 1245–1261. <https://doi.org/10.1093/evolut/qpad044>
- Juric, I., Aeschbacher, S., & Coop, G. (2016). The Strength of Selection against Neanderthal Introgression. *PLOS Genetics*, 12(11), e1006340. <https://doi.org/10.1371/journal.pgen.1006340>
- Karron, J. D., Mitchell, R. J., Holmquist, K. G., Bell, J. M., & Funk, B. (2004). The influence of floral display size on selfing rates in *Mimulus ringens*. *Heredity*, 92(3), 242–248. <https://doi.org/10.1038/sj.hdy.6800402>
- Kelly, J. K. (2022). The genomic scale of fluctuating selection in a natural plant population. *Evolution Letters*, 6(6), 506–521. <https://doi.org/10.1002/evl3.308>

- Kenney, A. M., & Sweigart, A. L. (2016). Reproductive isolation and introgression between sympatric *Mimulus* species. *Molecular Ecology*, 25(11), 2499–2517. <https://doi.org/10.1111/mec.13630>
- Klein, E. K., Lagache-Navarro, L., & Petit, R. J. (2017). Demographic and spatial determinants of hybridization rate. *Journal of Ecology*, 105(1), 29–38. <https://doi.org/10.1111/1365-2745.12674>
- Kleindorfer, S., O'Connor, J. A., Dudaniec, R. Y., Myers, S. A., Robertson, J., & Sulloway, F. J. (2014). Species collapse via hybridization in Darwin's tree finches. *The American Naturalist*, 183(3), 325–341. <https://doi.org/10.1086/674899>
- Langdon, Q. K., Groh, J. S., Aguillon, S. M., Powell, D. L., Gunn, T., Payne, C., Baczenas, J. J., Donny, A., Dodge, T. O., Du, K., Schartl, M., Ríos-Cárdenas, O., Gutiérrez-Rodríguez, C., Morris, M., & Schumer, M. (2024). Swordtail fish hybrids reveal that genome evolution is surprisingly predictable after initial hybridization. *PLOS Biology*, 22(8), e3002742. <https://doi.org/10.1371/journal.pbio.3002742>
- Lee, H. Y., Chou, J. Y., Cheong, L., Chang, N. H., Yang, S. Y., & Leu, J. Y. (2008). Incompatibility of Nuclear and Mitochondrial Genomes Causes Hybrid Sterility between Two Yeast Species. *Cell*, 135(6), 1065–1073. <https://doi.org/10.1016/j.cell.2008.10.047>
- Li, Z., Pinson, S. R. M., Paterson, A. H., Park, W. D., & Stansel, J. W. (1997). Genetics of Hybrid Sterility and Hybrid Breakdown in an Intersubspecific Rice (*Oryza sativa* L.) Population. *Genetics*, 145(4), 1139–1148. <https://doi.org/10.1093/genetics/145.4.1139>
- Lotsy, J. P. (1931). On the species of the taxonomist in its relation to evolution. *Genetica*, 13(1), 1–16. <https://doi.org/10.1007/BF01725037>
- Lowry, D. B., Modliszewski, J. L., Wright, K. M., Wu, C. A., & Willis, J. H. (2008). The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1506), 3009–3021. <https://doi.org/10.1098/rstb.2008.0064>
- Lowry, D. B., Sobel, J. M., Angert, A. L., Ashman, T.-L., Baker, R. L., Blackman, B. K., Brandvain, Y., Byers, K. J. R. P., Cooley, A. M., Coughlan, J. M., Dudash, M. R., Fenster, C. B., Ferris, K. G., Fishman, L., Friedman, J., Grossenbacher, D. L., Holeski, L. M., Ivey, C. T., Kay, K. M., ... Yuan, Y.-W. (2019). The case for the continued use of the genus name *Mimulus* for all monkeyflowers. *TAXON*, 68(4), 617–623. <https://doi.org/10.1002/tax.12122>
- Lowry, D. B., & Willis, J. H. (2010). A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biology*, 8(9). <https://doi.org/10.1371/journal.pbio.1000500>
- Macholán, M., Munclinger, P., Šugerková, M., Dufková, P., Bímová, B., Božíková, E., Zima, J., & Piálek, J. (2007). Genetic analysis of autosomal and x-linked markers across a mouse

- hybrid zone. *Evolution*, 61(4), 746–771. <https://doi.org/10.1111/j.1558-5646.2007.00065.x>
- Mallet, J., Besansky, N., & Hahn, M. W. (2015). How reticulated are species? *BioEssays*, 38(2), 140–149. <https://doi.org/10.1002/bies.201500149>
- Malone, J. H., & Fontenot, B. E. (2008). Patterns of Reproductive Isolation in Toads. *PLoS ONE*, 3(12), e3900. <https://doi.org/10.1371/journal.pone.0003900>
- Mantel, S. J., & Sweigart, A. L. (2019). Divergence in drought-response traits between sympatric species of *Mimulus*. *Ecology and Evolution*. <https://doi.org/10.1002/ece3.5549>
- Martin, N. H., & Willis, J. H. (2007). Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric *Mimulus* species. *Evolution*, 61(1), 68–82. <https://doi.org/10.1111/j.1558-5646.2007.00006.x>
- Martin, N. H., & Willis, J. H. (2010). Geographical variation in postzygotic isolation and its genetic basis within and between two *Mimulus* species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1552), 2469–2478. <https://doi.org/10.1098/rstb.2010.0030>
- Martin, S. H., Davey, J. W., Salazar, C., & Jiggins, C. D. (2019). Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLoS Biology*, 17(2), 1–28. <https://doi.org/10.1371/journal.pbio.2006288>
- Masly, J. P., Jones, C. D., Noor, M. A. F., Locke, J., & Orr, H. A. (2006). Gene transposition as a cause of hybrid sterility in *Drosophila*. *Science*, 313(5792), 1448–1450. <https://doi.org/10.1126/science.1128721>
- Matute, D. R., Butler, I. A., Turissini, D. A., & Coyne, J. A. (2010). A test of the snowball theory for the rate of evolution of hybrid incompatibilities. *Science (New York, N.Y.)*, 329(5998), 1518–1521. <https://doi.org/10.1126/science.1193440>
- Matute, D. R., & Cooper, B. S. (2021). Comparative studies on speciation: 30 years since Coyne and Orr. *Evolution*, 1989, 1–15. <https://doi.org/10.1111/evo.14181>
- Mayr, E. (1942). *Systematics and the Origin of Species*. Columbia University Press.
- McDonald, L. M., Scharnagl, A., Turcu, A. K., Patterson, C. M., & Kooyers, N. J. (2023). Demographic consequences of an extreme heat wave are mitigated by spatial heterogeneity in an annual monkeyflower. *Ecology and Evolution*, 13(8), e10397. <https://doi.org/10.1002/ece3.10397>
- Melo, M. C., Greal, A., Brittain, B., Walter, G. M., & Ortiz-Barrientos, D. (2014). Strong extrinsic reproductive isolation between parapatric populations of an Australian groundsel. *New Phytologist*, 203(1), 323–334. <https://doi.org/10.1111/nph.12779>

- Mendelson, T. C. (2003). Sexual isolation evolves faster than hybrid inviability in a diverse and sexually dimorphic genus of fish (Percidae: Etheostoma). *Evolution*, 57(2), 317–327. <https://doi.org/10.1111/j.0014-3820.2003.tb00266.x>
- Mojica, J. P., Lee, Y. W., Willis, J. H., & Kelly, J. K. (2012). Spatially and temporally varying selection on intrapopulation quantitative trait loci for a life history trade-off in *Mimulus guttatus*. *Molecular Ecology*, 21(15), 3718–3728. <https://doi.org/10.1111/j.1365-294X.2012.05662.x>
- Moran, B. M., Payne, C. Y., Powell, D. L., Iverson, E. N. K., Donny, A. E., Banerjee, S. M., Langdon, Q. K., Gunn, T. R., Rodriguez-Soto, R. A., Madero, A., Baczenas, J. J., Kleczko, K. M., Liu, F., Matney, R., Singhal, K., Leib, R. D., Hernandez-Perez, O., Corbett-Detig, R., Frydman, J., ... Schumer, M. (2024). A lethal mitonuclear incompatibility in complex I of natural hybrids. *Nature*, 626(7997), 119–127. <https://doi.org/10.1038/s41586-023-06895-8>
- Moyle, L. C., & Nakazato, T. (2010). Hybrid incompatibility “snowballs” between *Solanum* species. *Science (New York, N.Y.)*, 329(5998), 1521–1523. <https://doi.org/10.1126/science.1193063>
- Moyle, L. C., Olson, M. S., & Tiffin, P. (2004). Patterns of reproductive isolation in three angiosperm genera. *Evolution*, 58(6), 1195–1208. <https://doi.org/10.1111/j.0014-3820.2004.tb01700.x>
- Muller, H. J. (1942). Isolating mechanisms, evolution, and temperature. *Biological Symposia*, 6, 71–125.
- Murlas Cosmides, L., & Tooby, J. (1981). Cytoplasmic inheritance and intragenomic conflict. *Journal of Theoretical Biology*, 89(1), 83–129. [https://doi.org/10.1016/0022-5193\(81\)90181-8](https://doi.org/10.1016/0022-5193(81)90181-8)
- Nelson, T. C., Stathos, A. M., Vanderpool, D. D., Finseth, F. R., Yuan, Y., & Fishman, L. (2021). Ancient and recent introgression shape the evolutionary history of pollinator adaptation and speciation in a model monkeyflower radiation (*Mimulus* section *Erythranthe*). *PLoS Genetics*, 17(2), e1009095. <https://doi.org/10.1371/journal.pgen.1009095>
- Nesom, G. L., Fraga, N. S., Barker, W. R., Beardsley, P. M., Tank, D. C., Baldwin, B. G., & Olmstead, R. G. (2019). Response to “The case for the continued use of the genus name *Mimulus* for all monkeyflowers.” *Taxon*, 68(4), 624–627.
- Nouhaud, P., Martin, S. H., Portinha, B., Sousa, V. C., & Kulmuni, J. (2022). Rapid and predictable genome evolution across three hybrid ant populations. *PLOS Biology*, 20(12), e3001914. <https://doi.org/10.1371/journal.pbio.3001914>
- Oneal, E., Willis, J. H., & Franks, R. G. (2016). Disruption of endosperm development is a major cause of hybrid seed inviability between *Mimulus guttatus* and *Mimulus nudatus*. *New Phytologist*, 210(3), 1107–1120. <https://doi.org/10.1111/nph.13842>

- Orr, H. A., & Turelli, M. (2001). The evolution of postzygotic isolation: Accumulating Dobzhansky-Muller incompatibilities. *Evolution*, 55(6), 1085–1094. <https://doi.org/10.1111/j.0014-3820.2001.tb00628.x>
- Osmolovsky, I., Shifrin, M., Gamliel, I., Belmaker, J., & Sapir, Y. (2022). Eco-Geography and Phenology Are the Major Drivers of Reproductive Isolation in the Royal Irises, a Species Complex in the Course of Speciation. *Plants*, 11(23), Article 23. <https://doi.org/10.3390/plants11233306>
- Pardo-Díaz, C., Salazar, C., Baxter, S. W., Merot, C., & Figueiredo-Ready, W. (2012). Adaptive Introgression across Species Boundaries in Heliconius Butterflies. *PLoS Genet*, 8(6), 1002752. <https://doi.org/10.1371/journal.pgen.1002752>
- Presgraves, D. C., Balagopalan, L., Abmayr, S. M., & Orr, H. A. (2003). Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. *Nature*, 423(6941), 715–719. <https://doi.org/10.1038/nature01679>
- Puzey, J. R., Willis, J. H., & Kelly, J. K. (2017). Population structure and local selection yield high genomic variation in *Mimulus guttatus*. *Molecular Ecology*, 26(2), 519–535. <https://doi.org/10.1111/mec.13922>
- Racimo, F., Sankararaman, S., Nielsen, R., & Huerta-Sánchez, E. (2015). Evidence for archaic adaptive introgression in humans. *Nature Reviews Genetics*, 16(6), 359–371. <https://doi.org/10.1038/nrg3936>
- Randi, E., & Lucchini, V. (2002). Detecting rare introgression of domestic dog genes into wild wolf (*Canis lupus*) populations by Bayesian admixture analyses of microsatellite variation. *Conservation Genetics*, 3(1), 29–43. <https://doi.org/10.1023/A:1014229610646>
- Richards, T. J., & Ortiz-Barrientos, D. (2016). Immigrant inviability produces a strong barrier to gene flow between parapatric ecotypes of *Senecio luteus*. *Evolution*, 70(6), 1239–1248. <https://doi.org/10.1111/evo.12936>
- Richmond, J. Q., Jockusch, E. L., & Latimer, A. M. (2011). Mechanical Reproductive Isolation Facilitates Parallel Speciation in Western North American Scincid Lizards. *The American Naturalist*, 178(3), 320–332. <https://doi.org/10.1086/661240>
- Riquet, F., Liautard-Haag, C., Woodall, L., Bouza, C., Louisy, P., Hamer, B., Otero-Ferrer, F., Aublanc, P., Béduneau, V., Briard, O., El Ayari, T., Hochscheid, S., Belkhir, K., Arnaud-Haond, S., Gagnaire, P., & Bierne, N. (2019). Parallel pattern of differentiation at a genomic island shared between clinal and mosaic hybrid zones in a complex of cryptic seahorse lineages. *Evolution*, 73(4), 817–835. <https://doi.org/10.1111/evo.13696>
- Ross, C. L., & Harrison, R. G. (2002). A Fine-Scale Spatial Analysis of the Mosaic Hybrid Zone Between *Gryllus firmus* and *Gryllus pennsylvanicus*. *Evolution*, 56(11), 2296–2312. <https://doi.org/10.1111/j.0014-3820.2002.tb00153.x>

- Ruhsam, M., Hollingsworth, P. M., & Ennos, R. A. (2011). Early evolution in a hybrid swarm between outcrossing and selfing lineages in Geum. *Heredity*, 107(3), 246–255. <https://doi.org/10.1038/hdy.2011.9>
- Runemark, A., Trier, C. N., Eroukhmanoff, F., Hermansen, J. S., Matschiner, M., Ravinet, M., Elgvin, T. O., & Sætre, G.-P. (2018). Variation and constraints in hybrid genome formation. *Nature Ecology & Evolution*, 2(3), Article 3. <https://doi.org/10.1038/s41559-017-0437-7>
- Sachdeva, H., & Barton, N. H. (2018). Introgression of a Block of Genome Under Infinitesimal Selection. *Genetics*, 209(4), 1279–1303. <https://doi.org/10.1534/genetics.118.301018>
- Sandstedt, G. D., & Sweigart, A. L. (2022). Developmental evidence for parental conflict in driving Mimulus species barriers. *New Phytologist*, 236(4), 1545–1557. <https://doi.org/10.1111/nph.18438>
- Scally, A., Dutheil, J. Y., Hillier, L. W., Jordan, G. E., Goodhead, I., Herrero, J., Hobolth, A., Lappalainen, T., Mailund, T., Marques-Bonet, T., McCarthy, S., Montgomery, S. H., Schwalie, P. C., Tang, Y. A., Ward, M. C., Xue, Y., Yngvadottir, B., Alkan, C., Andersen, L. N., ... Durbin, R. (2012). Insights into hominid evolution from the gorilla genome sequence. *Nature*, 483(7388), Article 7388. <https://doi.org/10.1038/nature10842>
- Schemske, D. W., & Bradshaw, H. D. (1999). Pollinator preference and the evolution of floral traits in monkeyflowers (Mimulus). *Proceedings of the National Academy of Sciences of the United States of America*, 96(21), 11910–11915. <https://doi.org/10.1073/pnas.96.21.11910>
- Schilthuizen, M., Giesbers, M. C. W. G., & Beukeboom, L. W. (2011). Haldane’s rule in the 21st century. *Heredity*, 107(2), Article 2. <https://doi.org/10.1038/hdy.2010.170>
- Schumer, M., Xu, C., Powell, D. L., Durvasula, A., Skov, L., Holland, C., Blazier, J. C., Sankararaman, S., Andolfatto, P., Gil, †, Rosenthal, G., & Przeworski, M. (2018). Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science*, 360(6389), 656–660.
- Scopece, G., Widmer, A., & Cozzolino, S. (2008). Evolution of Postzygotic Reproductive Isolation in a Guild of Deceptive Orchids. *The American Naturalist*, 171(3), 315–326. <https://doi.org/10.1086/527501>
- Sedghifar, A., Brandvain, Y., & Ralph, P. (2016). Beyond clines: Lineages and haplotype blocks in hybrid zones. *Molecular Ecology*, 25(11), 2559–2576. <https://doi.org/10.1111/mec.13677>
- Selz, O. M., Pierotti, M. E. R., Maan, M. E., Schmid, C., & Seehausen, O. (2014). Female preference for male color is necessary and sufficient for assortative mating in 2 cichlid sister species. *Behavioral Ecology*, 25(3), 612–626. <https://doi.org/10.1093/beheco/aru024>

- Servedio, M. R., & Hermisson, J. (2020). The evolution of partial reproductive isolation as an adaptive optimum. *Evolution*, 74(1), 4–14. <https://doi.org/10.1111/evo.13880>
- Sianta, S. A., & Kay, K. M. (2021). Parallel evolution of phenological isolation across the speciation continuum in serpentine-adapted annual wildflowers. *Proceedings. Biological Sciences*, 288(1948), 20203076. <https://doi.org/10.1098/rspb.2020.3076>
- Sianta, S. A., Moeller, D. A., & Brandvain, Y. (2024). The extent of introgression between incipient *Clarkia* species is determined by temporal environmental variation and mating system. *Proceedings of the National Academy of Sciences*, 121(12), e2316008121. <https://doi.org/10.1073/pnas.2316008121>
- Simon, A., Fraïsse, C., El Ayari, T., Liautard-Haag, C., Strelkov, P., Welch, J. J., & Bierne, N. (2021). How do species barriers decay? Concordance and local introgression in mosaic hybrid zones of mussels. *Journal of Evolutionary Biology*, 34(1), 208–223. <https://doi.org/10.1111/jeb.13709>
- Sobel, J. M. (2014). Ecogeographic Isolation and Speciation in the Genus *Mimulus*. *The American Naturalist*, 184(5), 565–579. <https://doi.org/10.1086/678235>
- Soliman, H. K., & Coughlan, J. M. (2024). United by conflict: Convergent signatures of parental conflict in angiosperms and placental mammals. *Journal of Heredity*, esae009. <https://doi.org/10.1093/jhered/esae009>
- Stathos, A., & Fishman, L. (2014). Chromosomal rearrangements directly cause underdominant F1 pollen sterility in *Mimulus lewisii*–*Mimulus cardinalis* hybrids. *Evolution*, 68(11), 3109–3119. <https://doi.org/10.1111/evo.12503>
- Stebbins, G. L. (1950). *Variation and Evolution in Plants*. Columbia University Press.
- Stelkens, R. B., Schmid, C., & Seehausen, O. (2015). Hybrid Breakdown in Cichlid Fish. *PLOS ONE*, 10(5), e0127207. <https://doi.org/10.1371/journal.pone.0127207>
- Streisfeld, M. A., & Kohn, J. R. (2005). Contrasting patterns of floral and molecular variation across a cline in *Mimulus aurantiacus*. *Evolution*, 59(12), 2548–2559. <https://doi.org/10.1111/j.0014-3820.2005.tb00968.x>
- Streisfeld, M. A., & Rausher, M. D. (2009). Altered trans-Regulatory Control of Gene Expression in Multiple Anthocyanin Genes Contributes to Adaptive Flower Color Evolution in *Mimulus aurantiacus*. *Molecular Biology and Evolution*, 26(2), 433–444. <https://doi.org/10.1093/molbev/msn268>
- Sullivan, B. K. (1995). Temporal stability in hybridization between *Bufo microscaphus* and *Bufo woodhousii* (Anura: Bufonidae): Behavior and morphology. *Journal of Evolutionary Biology*, 8(2), 233–247. <https://doi.org/10.1046/j.1420-9101.1995.8020233.x>

- Svensson, E. I., Eroukhmanoff, F., & Friberg, M. (2006). Effects of natural and sexual selection on adaptive population divergence and premating isolation in a damselfly. *Evolution*, 60(6), 1242–1253. <https://doi.org/10.1111/j.0014-3820.2006.tb01202.x>
- Sweigart, A. L., Fishman, L., & Willis, J. H. (2006). A simple genetic incompatibility causes hybrid male sterility in *mimulus*. *Genetics*, 172(4), 2465–2479. <https://doi.org/10.1534/genetics.105.053686>
- Sweigart, A. L., & Flagel, L. E. (2015). Evidence of natural selection acting on a polymorphic hybrid incompatibility locus in *mimulus*. *Genetics*, 199(2), 543–554. <https://doi.org/10.1534/genetics.114.171819>
- Sweigart, A. L., Mason, A. R., & Willis, J. H. (2007). Natural variation for a hybrid incompatibility between two species of *Mimulus*. *Evolution*, 61(1), 141–151. <https://doi.org/10.1111/j.1558-5646.2007.00011.x>
- Tataru, D., Wheeler, E. C., & Ferris, K. G. (2023). Spatially and temporally varying selection influence species boundaries in two sympatric *Mimulus*. *Proceedings of the Royal Society B: Biological Sciences*, 290(1992), 20222279. <https://doi.org/10.1098/rspb.2022.2279>
- Taylor, S. A., Curry, R. L., White, T. A., Ferretti, V., & Lovette, I. (2014). Spatiotemporally consistent genomic signatures of reproductive isolation in a moving hybrid zone. *Evolution*, 68(11), 3066–3081. <https://doi.org/10.1111/evo.12510>
- Trier, C. N., Hermansen, J. S., Sætre, G.-P., & Bailey, R. I. (2014). Evidence for Mito-Nuclear and Sex-Linked Reproductive Barriers between the Hybrid Italian Sparrow and Its Parent Species. *PLOS Genetics*, 10(1), e1004075. <https://doi.org/10.1371/journal.pgen.1004075>
- Twyford, A. D., Streisfeld, M. A., Lowry, D. B., & Friedman, J. (2015). Genomic studies on the nature of species: Adaptation and speciation in *Mimulus*. *Molecular Ecology*, 24(11), 2601–2609. <https://doi.org/10.1111/mec.13190>
- Twyford, A. D., Wong, E. L. Y., & Friedman, J. (2020). Multi-level patterns of genetic structure and isolation by distance in the widespread plant *Mimulus guttatus*. *Heredity*, 125(4), 227–239. <https://doi.org/10.1038/s41437-020-0335-7>
- Vickery, R. K. (1978). Case Studies in the Evolution of Species Complexes in *Mimulus*. In *Evolutionary Biology* (pp. 405–507). Springer US. https://doi.org/10.1007/978-1-4615-6956-5_7
- Vickery, R. K. (1995). Speciation by Aneuploidy and Polyploidy in *Mimulus* (scrophulariaceae). *The Great Basin Naturalist*, 55(2), 174–176.
- Wessinger, C. A., Katzer, A. M., Hime, P. M., Rausher, M. D., Kelly, J. K., & Hileman, L. C. (2023). A few essential genetic loci distinguish *Penstemon* species with flowers adapted to pollination by bees or hummingbirds. *PLOS Biology*, 21(9), e3002294. <https://doi.org/10.1371/journal.pbio.3002294>

- Whitener, M. R., Mangelson, H., & Sweigart, A. L. (2024). Patterns of genomic variation reveal a single evolutionary origin of the wild allotetraploid *Mimulus sookensis*. *Evolution*, qpae079. <https://doi.org/10.1093/evolut/qpae079>
- Wright, K. M., Lloyd, D., Lowry, D. B., Macnair, M. R., & Willis, J. H. (2013). Indirect Evolution of Hybrid Lethality Due to Linkage with Selected Locus in *Mimulus guttatus*. *PLoS Biology*, 11(2), e1001497. <https://doi.org/10.1371/journal.pbio.1001497>
- Wu, C. A., Lowry, D. B., Cooley, A. M., Wright, K. M., Lee, Y. W., & Willis, J. H. (2008). *Mimulus* is an emerging model system for the integration of ecological and genomic studies. *Heredity*, 100(2), 220–230. <https://doi.org/10.1038/sj.hdy.6801018>
- Wu, C. A., Streisfeld, M. A., Nutter, L. I., & Cross, K. A. (2013). The Genetic Basis of a Rare Flower Color Polymorphism in *Mimulus lewisii* Provides Insight into the Repeatability of Evolution. *PLoS ONE*, 8(12), e81173. <https://doi.org/10.1371/journal.pone.0081173>
- Xiong, T., & Mallet, J. (2022). On the impermanence of species: The collapse of genetic incompatibilities in hybridizing populations. *Evolution*, 76(11), 2498–2512. <https://doi.org/10.1111/evo.14626>
- Yuan, Y.-W. (2019). Monkeyflowers (*Mimulus*): New model for plant developmental genetics and evo-devo. *New Phytologist*, 222(2), 694–700. <https://doi.org/10.1111/nph.15560>
- Yuan, Y.-W., Rebocho, A. B., Sagawa, J. M., Stanley, L. E., & Bradshaw, H. D. (2016). Competition between anthocyanin and flavonol biosynthesis produces spatial pattern variation of floral pigments between *Mimulus* species. *Proceedings of the National Academy of Sciences*, 113(9), 2448–2453. <https://doi.org/10.1073/pnas.1515294113>
- Yuan, Y.-W., Sagawa, J. M., Young, R. C., Christensen, B. J., & Bradshaw, H. D., Jr. (2013). Genetic Dissection of a Major Anthocyanin QTL Contributing to Pollinator-Mediated Reproductive Isolation Between Sister Species of *Mimulus*. *Genetics*, 194(1), 255–263. <https://doi.org/10.1534/genetics.112.146852>
- Zuellig, M. P., & Sweigart, A. L. (2018a). A two-locus hybrid incompatibility is widespread, polymorphic, and active in natural populations of *Mimulus*. *Evolution*, 72(11), 2394–2405. <https://doi.org/10.1111/evo.13596>
- Zuellig, M. P., & Sweigart, A. L. (2018b). Gene duplicates cause hybrid lethality between sympatric species of *Mimulus*. *PLoS Genetics*, 14(4), e1007130. <https://doi.org/10.1101/201392>

CHAPTER II

STRONG POSTMATING REPRODUCTIVE ISOLATION IN *MIMULUS* SECTION
*EUNANUS*¹

¹Farnitano M.C. and A.L. Sweigart. 2023. Strong postmating reproductive isolation in *Mimulus* section *Eunanus*. *Journal of Evolutionary Biology*. 36(10):1393-1410. Reprinted here with permission of the publisher.

Abstract

Postmating reproductive isolation can help maintain species boundaries when premating barriers to reproduction are incomplete. The strength and identity of postmating reproductive barriers are highly variable among diverging species, leading to questions about their genetic basis and evolutionary drivers. These questions have been tackled in model systems but are less often addressed with broader phylogenetic resolution. In this study we analyze patterns of genetic divergence alongside direct measures of postmating reproductive barriers in an overlooked group of sympatric species within the model monkeyflower genus, *Mimulus*. Within this *Mimulus brevipes* species group, we find substantial divergence among species, including a cryptic genetic lineage. However, rampant gene discordance and ancient signals of introgression suggest a complex history of divergence. In addition, we find multiple strong postmating barriers, including postmating prezygotic isolation, hybrid seed inviability, and hybrid male sterility. *M. brevipes* and *M. fremontii* have substantial but incomplete postmating isolation. For all other tested species pairs, we find essentially complete postmating isolation. Hybrid seed inviability appears linked to differences in seed size, providing a window into possible developmental mechanisms underlying this reproductive barrier. While geographic proximity and incomplete mating isolation may have allowed gene flow within this group in the distant past, strong postmating reproductive barriers today have likely played a key role in preventing ongoing introgression. By producing foundational information about reproductive isolation and genomic divergence in this understudied group, we add new diversity and phylogenetic resolution to our understanding of the mechanisms of plant speciation.

Introduction

Evolutionary biologists have long been interested in understanding what patterns and processes lead to reproductive isolation among species (Coyne & Orr, 1997; Darwin, 1859; V. Grant, 1981). Reproductive isolation is often broken down into components that act at sequential stages of the life cycle – premating barriers such as ecogeographic isolation and pollinator-mediated isolation that prevent mating, postmating prezygotic barriers that reduce fertilization success after mating has occurred, and postmating postzygotic barriers that reduce the fitness of hybrid offspring relative to pure species (Ramsey et al., 2003; Sobel & Chen, 2014). In plants, premating barriers are typically thought to evolve more quickly than postmating barriers and to have a greater contribution to overall isolation (Christie et al., 2022). However, this general pattern masks a great deal of heterogeneity across species pairs: many systems are completely isolated by premating barriers, while others rely exclusively on strong postmating isolation; many have a mix of both. Postmating barriers are often the result of intrinsic genetic incompatibilities, making them potentially more stable over the long term than ecologically mediated premating barriers (Coughlan & Matute, 2020). Postmating barriers may also be a source of cryptic variation in groups without obvious morphological or ecological differences. Species complexes, consisting of many groups at various stages of divergence and gene flow, have been useful tools for speciation research. By placing reproductive isolation and divergence in a phylogenomic context, species complexes allow us to make informed comparisons across groups, distinguishing patterns and trends while uncovering crucial differences. For example, the degree of reproductive isolation between two taxa tends to increase with genetic distance (Christie & Strauss, 2018; Coyne & Orr, 1989; Malone & Fontenot, 2008; Scopece et al., 2007), but careful work in irises and other species complexes has demonstrated substantial

heterogeneity in this relationship (Moyle et al., 2004; Osmolovsky et al., 2022). In addition, the emergence speed of different categories of isolation (Christie & Strauss, 2018) or of specific reproductive barriers (e.g., sterility vs. inviability) (Coyne & Orr, 1989; Le Gac et al., 2007; Malone & Fontenot, 2008; Presgraves, 2002) may vary considerably.

The reasons behind these differences are a subject of ongoing investigation, but may include differences in genetic architecture (Guerrero et al., 2017; Moyle & Payseur, 2009) or the forms of selection acting on each barrier (Baack et al., 2015). For example, conflicting selection pressures between maternal and paternal genomes in the mammalian placenta or the flowering plant endosperm are thought to give rise to reproductive barriers as these conflicts are resolved differently in independent lineages (Crespi & Nosil, 2013; Haig & Westoby, 1991; Lafon-Placette & Köhler, 2016); the relative strength of this parental conflict could influence the rate of barrier evolution (Coughlan et al., 2020; Raunsgard et al., 2018). Different forms of isolation may also be genetically linked in colocalized QTL or chromosomal rearrangements, causing them to evolve non-independently (Charron et al., 2014; Ferris et al., 2017; Noor et al., 2001). The importance of this genomic colocalization in speciation, and the role of chance vs. selection in its appearance, are subjects of ongoing debate (Cruickshank & Hahn, 2014; Duranton et al., 2018; Fuller et al., 2018; Kirkpatrick & Barrett, 2015; Renaut et al., 2013).

An excellent system for building a phylogenetically informed understanding of reproductive isolation is the western North American radiation of monkeyflowers in the genus *Mimulus* (Lamiales: Phrymaceae). Multiple species complexes within *Mimulus* have already been studied extensively in a speciation context, including the *Mimulus guttatus* (e.g., Brandvain et al., 2014; Ferris et al., 2014; Lowry & Willis, 2010), *M. aurantiacus* (e.g., Stankowski et al., 2019; Streisfeld & Kohn, 2005), and *M. lewisii* (e.g., Nelson et al., 2021; Ramsey et al., 2003)

complexes. A number of pre- and postmating reproductive barriers have been mapped to single genes or QTL (e.g., Fishman et al., 2014; Fishman & Willis, 2006; Streisfeld & Rausher, 2009; Sweigart et al., 2006; Yuan et al., 2013; Zuellig & Sweigart, 2018). Recently, hybrid seed inviability in particular has been identified as a strong barrier in multiple independent species pairs of the *Mimulus guttatus* species complex, with parental conflict repeatedly implicated as a common mechanism (Coughlan et al., 2020; Oneal et al., 2016; Sandstedt et al., 2020; Sandstedt & Sweigart, 2022). An expanding set of genetic and developmental resources across the genus allow for a holistic approach to understanding reproductive isolation from genic to phylogenetic scales (Yuan, 2019).

However, some *Mimulus* subgroups have received less attention, in part because they are less experimentally tractable or more difficult to access. Studies in these other groups would provide an important complement to existing knowledge, facilitating comparative analysis at the genus level while highlighting overlooked diversity in the genetic and ecological mechanisms of speciation. One such neglected group is *Mimulus* section *Eunanus* [synonym *Diplacus* section *Eunanus*, see (Barker et al., 2012; Lowry et al., 2019) for a discussion of nomenclatural issues]. Section *Eunanus* is a clade of ~23 species (Grant, 1924; Nesom, 2013) which prior to this publication has had no genome-scale sequencing and few assessments of reproductive isolation. The group includes both widespread and narrowly endemic species occupying a range of elevations and habitats in western North America. Here, we focus primarily on three species from this group: *Mimulus brevipes* Benth., *Mimulus fremontii* (Benth.) Gray, and *Mimulus johnstonii* Grant (Baldwin et al., 2012; A. L. Grant, 1924) (Figure 2.1A). These three morphologically defined species make up a well-supported clade according to a phylogenetic study of *Mimulus* based on three genetic markers (Beardsley et al., 2004). Their ranges are

overlapping; *M. johnstonii* is restricted to higher elevations in the Transverse Ranges of southern California, while *M. fremontii* and *M. brevipes* are more widespread throughout the coastal ranges from Monterey into Baja California (Figure 2.1B).

Studies of reproductive isolation in these three species are limited. Species distribution modeling between *Mimulus johnstonii* and *M. brevipes* identified strong but incomplete ecogeographical isolation (RI=0.65 and 0.82 reciprocally) (Sobel, 2014), and they can be found within hundreds of feet of each other (pers. obs.). No distribution models have been made with *M. fremontii*, but *M. brevipes* and *M. fremontii* are found at similar elevations, flower at similar times, and can be found growing within inches of each other (pers. obs.). Major differences in flower color, size, and shape between *M. brevipes* and its relatives might suggest some degree of pollinator-mediated isolation, though little is known about the identity of pollinators or levels of outcrossing; *M. fremontii* and *M. johnstonii* have similar floral characteristics and seem unlikely to have substantial pollinator isolation. Postmating barriers have not been studied, with the exception of one study showing low germination rates of F1 hybrid seeds between *M. brevipes* and *M. johnstonii*, potentially suggesting hybrid seed inviability (Sobel, 2010).

In this study, we present the first genome-scale analysis of divergence relationships in *Mimulus* section *Eunanus*, incorporating our three focal species as well as two more distant species, *M. constrictus* and *M. nanus*. In addition, we use controlled crosses to examine multiple postmating reproductive barriers between the focal species. We find clear genetic divergence and multiple strong postmating reproductive barriers between all tested species, but signals of historical introgression persist. Our results add to a growing body of evidence that intrinsic postmating barriers, especially hybrid seed inviability, are a key component in the maintenance of plant species.

Materials and Methods

Sample collections and growth conditions

Fruits were collected from wild plants of three focal species (*M. brevipes*, *M. fremontii*, and *M. johnstonii*) at nine locations in southern California in summer 2019 (Figure 2.1B, Table S2.1). A population of *M. constrictus* from California in 2019 and a population of *M. nanus* from Oregon in 2021 were collected as additional comparisons; these species are phylogenetically within section *Eunanus* but outside of the immediate focal group. To germinate, we treated seeds overnight with 1 mM gibberellic acid, then rinsed three times with distilled water and placed them individually in 2x2cm soil plugs. Soil was a 1:1:1 mix of sand, perlite, and peat with a thin layer of finely sifted soil mix on top. Germination trays were set in standing bottom water and kept moist with a squirt bottle as necessary. When germination rates were low, liquid smoke was added to standing water to simulate post-fire conditions and stimulate germination. After germinants developed one or two pairs of true leaves, we transplanted soil plugs into plastic conical pots filled with the same soil mix. The conical pots were maintained in standing bottom water in a growth chamber set to 23C days/18C nights, with low relative humidity and 16 hours of daylight. Germination and establishment was low overall and highly variable across species, populations, and trials, so germination and establishment rates were not used as a proxy for seed viability or hybrid fitness.

One population initially collected as *M. johnstonii*, S25 from Sespe Creek, showed morphological differences suggesting it may be a unique lineage, and is treated separately throughout this paper. Compared to *M. johnstonii*, Sespe Creek has inserted stigmas, a lighter purple corolla with a slight white wash at the mouth, and longer leaf trichomes.

Genomic sequencing

We collected leaf tissue from 33 individuals for genomic sequencing. All sequenced individuals were generated from outbred field-collected seeds and germinated in growth chambers. We flash-froze leaf tissue in liquid nitrogen and used a modified CTAB-based protocol to extract genomic DNA (Fishman, 2020). Briefly, tissue was ground to a powder and incubated in CTAB buffer with β -mercaptoethanol. Supernatant was extracted with phenol-chloroform-isoamyl alcohol, then again with chloroform-isoamyl alcohol, followed by precipitation of carbohydrates with sodium chloride and polyethylene glycol. DNA was then precipitated with isopropanol, washed with ethanol, and resuspended in distilled water.

We used the Illumina Nextera XT DNA library prep kit (FC-131-1024), which uses a tagmentation enzyme to simultaneously fragment DNA and ligate adapters, to prepare genomic DNA for sequencing. The 33 individually barcoded samples were then sequenced with a read length of 2x150bp (paired end) in four batches at the Duke University Sequencing and Genomic Technologies Shared Resource with either an Illumina NextSeq 500 or an Illumina NovaSeq 6000 sequencer, in combined lanes with other barcoded projects (Table S2.2).

Sequence alignments and SNP calling

We combined our dataset of 33 samples with five previously published samples (Stankowski et al., 2019) from the *Mimulus aurantiacus* species complex as an outgroup for analysis. *M. aurantiacus* is outside of section *Eunanus* but is the closest group with previous genome-scale data, as well as the closest group for which a reference genome assembly is currently available. For all 38 samples, we trimmed raw sequencing reads using Trimmomatic v0.39 (Bolger et al., 2014) to remove adapters and low-quality ends. We mapped reads to the *Mimulus aurantiacus* reference genome (Stankowski et al., 2019) using bwa mem v0.7.17 (Li & Durbin, 2009), then sorted and indexed with samtools v1.10 (Danecek et al., 2021). We removed

PCR duplicates using the MarkDuplicates tool in Picard v2.21.6 (Broad Institute, 2019), and removed reads with unmapped mate pairs or mapping quality <29 using samtools view (Danecek et al., 2021). Average read coverage per sample was 2.2-15.7 (mean 6.8) across 33 novel samples (details in Table S2.2).

We called SNPs in each sample using the GATK v4.1.6.0 haplotype caller, then used GATK GenotypeGVCFs to jointly call variant and invariant sites for the entire dataset in ‘all-sites’ mode (Van der Auwera & O’Connor, 2020). Only sites mapping to the ten major linkage groups of the reference genome, corresponding to ten nuclear chromosomes, were included. We split the resulting VCF into SNPs and invariant sites, and separately filtered each before recombining them into a joint dataset, following recommendations by Dmitri Kryvokhyza (Kryvokhyzha, 2022). We filtered both SNPs and invariant sites to remove sites with combined DP <83 (~1/3 of mean combined depth) or >1376 (two standard deviations above mean combined depth), QD<2, SOR>3, or MQ<40. We additionally filtered SNPs to keep only biallelic sites across the entire dataset, and to remove sites with QUAL<40, FS>60, MQRankSum<-12.5, ReadPosRankSum>12.5, or ReadPosRankSum<-12.5.

Starting with all sites that pass the above filters (14,814,282 biallelic SNPs and 96,871,941 invariant sites), we generated four datasets for analysis. For genome-wide phylogenetic and introgression analyses, we used a ‘complete’ dataset, which retained only sites where at least 31 of 38 (>80%) of samples had a called genotype, resulting in 7,728,322 biallelic SNPs and 32,658,171 invariant sites. When calculating genome-wide introgression metrics, uneven sampling of species or uneven coverage across individuals could produce a bias whereby variants in one taxon are called more frequently than in another. To reduce the effect of this bias, we created four independent iterations of ‘downsampled’ data with equalized sampling and

coverage. For three iterations, we randomly chose two individuals from each species; for the fourth iteration, we chose the two most divergent individuals from each species (except for *M. constrictus* for which we only had one sample). We then randomly sampled 12 million raw read pairs from each chosen individual and used the same pipeline as above to generate SNPs and invariant sites. Again, we retained only sites where 11 of 13 samples (>80%) had called genotypes, resulting in between 3,253,505 and 3,778,609 SNPs.

For the calculation of diversity and divergence metrics, we made a ‘synonymous sites’ dataset. Starting with the ‘complete’ dataset, we used the *M. aurantiacus* gene annotation to select only sites in the 3rd codon position of coding sequences that were four-fold degenerate, i.e., any nucleotide change to that site in the reference background would not change the protein sequence, using a custom script by Tim Sackton (Sackton, 2014/2022). The ‘synonymous sites’ dataset included 967,470 SNPs and 1,654,798 invariant sites. We note that multiple mutation hits within a single codon are possible given the large proportion of variant sites across the dataset, which could change the synonymous nature of these sites, but we expect these events to be rare enough to not substantially affect our genome-wide results on average.

For gene-by-gene analyses, we used a ‘genic’ dataset that included all sites passing the initial filters that fell within each of the 22,421 genes in the *M. aurantiacus* genome annotation.

Phylogenetic tree building

To determine phylogenetic relationships among our samples, we built a maximum likelihood phylogenetic tree using the GTR+Gamma model in RAxML v8.2.12 (Stamatakis, 2014) with SNPs from the ‘complete’ dataset. SNPs were extracted into a concatenated fasta alignment using the ‘consensus’ function from bcftools v1.13 (Danecek et al., 2021) and invariant sites were excluded. Heterozygous sites were initially coded as IUPAC ambiguity

codes, then randomly converted to a single allele with the ‘randbase’ function in seqtk v1.3 (Li, 2018). An ascertainment bias correction was included using the ‘Felsenstein’ method implemented in RAxML to account for the exclusion of invariant sites. We ran 4 separate iterations of RAxML with unique random seeds, then chose the iteration with the largest log-likelihood. We used the RAxML -b option to run 1000 rapid bootstrap trees. We also generated a neighbor-joining tree with 1000 bootstrap replicates from the same data, using the dist.ml, nj, and bootstrap.phyDat functions in the R package phangorn (Schliep, 2011).

We generated gene trees to examine patterns of gene tree discordance on the phylogeny. We used SNPs from the ‘genic’ dataset in the *M. aurantiacus* genome annotation, following the same approach as above to generate a maximum-likelihood phylogenetic tree in RAxML for each gene. Genes with insufficient data to resolve a tree (any individual with no called sites, or any two individuals with identical sequence) were excluded, resulting in trees for 17,573 genes. We then input these trees into ASTRAL v5.6.1 (Rabiee et al., 2019) which uses a multispecies coalescent approach to calculate a consensus species tree and quartet support scores for each node.

Quantifying homozygosity, diversity, and divergence

Pairwise sequence differences were calculated from the ‘synonymous sites’ dataset using the ‘dxy’ and ‘pi’ functions in the program pixy v1.2.3 (Korunes & Samuk, 2021). For each comparison, sequence differences were calculated in 1Mb windows across the genome, then added together and divided by the total number of informative sites. We calculated pairwise diversity (π_s) for each pair of individuals within a species, pairwise divergence (d_s) for each pair of individuals across species, and heterozygosity (h_s , pairwise diversity within an individual) for

each individual. Diversity, divergence, and heterozygosity values were averaged across individuals within each species, or across individual pairs within each species pair.

To estimate the divergence times between species, we used the following molecular clock formula: (divergence time in generations) = (synonymous divergence – synonymous ancestral diversity)/(2*mutation rate). We assumed ancestral diversity for any two species to be the average of within-species diversity for the two species, and took the mutation rate to be 1.5×10^{-8} (Koch et al., 2000). These estimates are approximate, as they do not take into account the effect of changing effective population sizes, such as demographic bottlenecks, or the possibility that ancestral diversity was very different from current diversity.

Quantifying introgression

To explore historical introgression between species, we used three complementary approaches: ABBA-BABA tests and the related f_4 -statistics; gene tree discordance bias using TWISST; and the model-based TreeMix program.

We used Dsuite v0.4 (Malinsky et al., 2021) to calculate Patterson’s D, a measure of the relative frequency of ABBA and BABA sites in the genome, and the related f_4 statistic, which uses allele frequencies to estimate the proportion of the genome resulting from introgression. We calculated D and f_4 for all possible trios of species in our dataset, using the ‘complete’ dataset as well as four ‘downsampled’ datasets, with the *M. aurantiacus* complex as the outgroup. Z-scores and associated p-values were calculated for each trio using a block-jackknife approach with 100 blocks in Dsuite; to account for multiple tests, a Bonferroni correction was used to obtain a corrected p-value (p_{corr}). Because there are multiple trios available to estimate introgression for the same pair of species, we summarized results across trios using the F_{branch} statistic in Dsuite for each species pair (Malinsky *et al.* 2018).

We used TWISST (github version [d56cefb](#)) (Martin & Van Belleghem, 2017) to obtain the proportion of gene trees supporting particular tree topologies. Gene trees produced for the ASTRAL analysis above, based on genes in the *M. aurantiacus* annotation, were used as input for TWISST, again excluding trees where any sample has 0 called sites or where any two samples have identical sequence. For each trio of species, we compared the two possible alternate (different from the consensus species tree) tree topologies using a binomial test with expected ratio 1:1 to look for an excess of one topology over another. The *M. aurantiacus* complex was always used as the outgroup. This test is analogous to the SNP-based ABBA-BABA test, but uses gene trees rather than SNPs.

We used TreeMix (Pickrell & Pritchard, 2012) to create phylogenetic network models with or without migration edges as a further test of introgression. We used the ‘synonymous sites’ dataset, including only variant SNPs with no missing data, and grouped individuals by species. The dataset was pruned to select SNPs not in linkage disequilibrium using a custom script by Joana Meier (<https://github.com/joanam/scripts/raw/master/LdPruning.sh>). We ran models with zero to eight migration edges, choosing the best of 10 replicate runs for each model type, then used likelihood ratio tests to determine whether each additional edge significantly improved the model fit. The program OptM (Fitak, 2021) was used as an alternative measure of the best fitting TreeMix migration model.

Identifying candidate windows with historical introgression

To identify candidate genomic regions with signatures of historical introgression, we used Dsuite (Malinsky et al., 2021) to calculate the window-based *df* statistic (Pfeifer and Kapan, 2019) for windows of 100 informative SNPs, each overlapping the previous by 50 SNPs. Low diversity is a potential confounding factor when using D- and f-statistics, particularly for small

genomic windows, so we used pixy as before to calculate ‘ π ’ for each species in 10kb regions, using the ‘complete’ dataset of SNPs and invariant sites (the ‘synonymous sites’ dataset did not provide good coverage of some areas at 10kb resolution). We assigned a π value to each 100-SNP df window based on the 10kb region containing the midpoint of that 100-SNP window. We calculated df across the genome for two focal trios with genome-wide support for introgression, using the *M. aurantiacus* complex as the outgroup. We then excluded any df window whose midpoint lies within a 10kb region with $\pi=0$ for any of the three focal species for that trio. To test the relationship between π and df , we binned all df windows by their associated π value for each species into 10 deciles of increasing π , and calculated mean df for each decile. Finally, to identify the most extreme outliers for df across the genome, we selected the 100 windows (out of 4732 or 4141, depending on the quartet) with the most extreme (positive or negative) values as potential introgression outliers, with positive values set to match the direction of the genome-wide F_{branch} signal.

Measuring reproductive barriers

We assessed three different barriers to reproduction: postmating prezygotic isolation, F1 hybrid seed inviability, and F1 hybrid sterility. We conducted hand pollinations of greenhouse-grown plants within and among species by first removing anthers from the maternal flower to prevent self-pollination, then using forceps to place pollen from the paternal flower on the maternal stigma surface. Stigma lobes in these species close quickly in response to touch, so hand-pollination is not always successful. In a few cases, supplemental pollen from the same paternal plant was added the following day once the stigma had re-opened. We collected mature, browned fruits, categorized according to cross type. Cross type throughout this paper refers to the combination of maternal species and paternal species, with maternal species always listed

first; the 16 cross types are JxJ, JxB, BxJ, JxF, FxJ, JxS, SxJ, BxB, BxF, FxB, BxS, SxB, FxF, FxS, SxF, SxS where J=*M. johnstonii*, B=*M. brevipes*, F=*M. fremontii*, S=Sespe creek population. For the cross type BxF, we also germinated and grew F1 hybrids from four independent crosses in the same conditions as parental plants (Table S2.3). These F1 hybrids were crossed to themselves and reciprocally with *M. fremontii* and *M. brevipes* as above (cross types HxB, BxH, HxH, HxF, FxH, where H=F1 hybrid). F1 hybrids between other species pairs were not grown because of low crossing success or low seed viability.

To measure postmating prezygotic isolation, we quantified crossing success (probability that a pollination produces at least one seed) and seed production (number of seeds per fruit) after hand-pollination; both viable and inviable seeds were included but unfertilized ovules were not. To test for an association between cross type and crossing success, we ran a penalized Firth regression with a binomial family function using the R package ‘brglm’. A Firth regression was chosen to account for complete separation (all successes or all failures) of some groups. To test the dominance relationships of this barrier, we compared the crossing success of *M. brevipes* x *M. fremontii* F1 hybrids to their parental species. We ran a generalized linear mixed-effects model (GLMM) with a binomial family and logit link function using the R package ‘lme4’ on cross types BxB, BxH, BxF, HxB, HxH, HxF, FxB, FxH, and FxF, where H indicates F1 hybrid offspring from a BxF cross. For this model, maternal population was included as a random effect, treating each independent F1 cross as its own population.

To assess seed viability in each cross, we used a combination of visual inspection and chemical treatment. All seeds were inspected under a dissecting scope and scored as inviable or viable. Viable seeds were plump and generally ovate, although they were allowed to be slightly misshapen as long as they appeared to be full. Inviabile seeds included any seed that was severely

shriveled, concave, empty, much smaller than the typical within-species seed, or with a shape very different from the typical ovate shape. To confirm that seeds visually scored as inviable were in fact inviable, we treated a subset of seeds with tetrazolium chloride, which stains viable seeds dark red. For this treatment, seeds were scarified using 1:5 bleach:distilled water with 0.83 uL/mL Triton-X for 15 minutes, washed twice with distilled water, and placed in 1% (w/v) tetrazolium chloride, then incubated in the dark at room temperature for ~48 hours. To test for an association between cross type and the relative counts of viable vs. inviable seeds, we ran a penalized Firth regression using the R package ‘brglm’ with a binomial family and logit link function. A Firth regression was chosen to account for complete separation (all viable or all inviable) of some groups.

Parental conflict is a potential driver of hybrid seed inviability, and is associated with asymmetries in hybrid seed traits (Haig & Westoby, 1991). To investigate seed size and shape in parental vs. hybrid seeds, we imaged a random subset of seeds from 49 representative fruits under a dissecting scope and used imageJ to manually measure seed length and width. To avoid conflating viability with seed measurements within a cross, we measured only viable seeds for majority-viable cross types, and only inviable seeds for majority-inviable cross types. We ran a linear mixed-effects model using the R package ‘lme4’ to test for differences in seed length between the four intraspecific cross types (JxJ, BxB, FxF, and SxS), using fruit as a random effect. We also ran six independent LMMs, one for each pair of species, to directly compare intra- and interspecific seeds. For example, we tested the *M. johnstonii* vs. *M. brevipes* pair by comparing cross types JxJ, JxB, BxJ, BxB. For each species-pair, we ran one LMM to test for an association between cross type and seed length, and a second to test for an association between cross type and seed length/width ratio.

To assess hybrid male sterility of *M. brevipes*-*M. fremontii* F1 plants compared with their parental species, we scored pollen number and pollen viability using an aniline blue stain. For each individual, all anthers from a single flower were collected in 50 μ L 0.25% aniline blue in lactophenol solution. Blue stained pollen grains were scored as viable, while unstained clear pollen grains were counted as inviable. Pollen was counted in full 1mm² (0.1 μ L) squares of a hemocytometer until 9 squares or at least 100 total pollen grains were counted. For each individual, we estimated pollen viability by determining the proportion of viable pollen grains out of the total counted (for some individuals, pollen viability was measured from multiple flowers and the average was used). We used an LMM to test for an association between species ID and the total count of pollen grains scored per hemocytometer square, with population as a random effect. Similarly, we used a GLM with a binomial family and logit link function to test for an association between species ID (*M. brevipes*, *M. fremontii*, or F1 hybrid) and the relative counts of viable vs. inviable pollen grains, adding population as a random effect (each hybrid family was treated as its own population).

To assess *M. brevipes*-*M. fremontii* F1 female fertility, we hand-pollinated F1 hybrids with pollen from either parent and measured seed production. To determine whether F1 hybrids produce fewer seeds than within-species crosses, we used a linear model to test for an association between cross type and the number of seeds produced per fruit, comparing cross types BxB, BxH, HxF, and FxF and only counting fruits that produced at least one seed. For every statistical model above, we ran Tukey post-hoc tests implemented in the package ‘multcomp’ to test for pairwise differences between each cross type or group.

We followed the methods of Sobel and Chen (2014) to calculate standardized measures of reproductive isolation for each of three reproductive barriers, as well as total postmating reproductive isolation, between each species. All measures used the equation

$$RI = 1 - 2 * (H / (H + C))$$

where H=heterospecific fitness and C=conspecific fitness. In this framework, RI ranges from -1 to 1 where -1 is complete heterosis, 0 is random assortment, and 1 is complete isolation. We calculated total measured reproductive isolation sequentially with the equation

$$RI_{TOTAL} = RI_{previous} + (1 - RI_{previous}) * RI_{new}$$

where $RI_{previous}$ is the total RI from all previous barriers and RI_{new} is the next barrier. Barriers that could not be measured due to complete or near-complete previous RI were left blank.

Results

Genome-wide variation in the M. brevipes group defines clear species but also reveals a complex history of introgression

A maximum-likelihood phylogeny generated from whole-genome SNP data resolves each of the three focal species (*M. brevipes*, *M. johnstonii*, and *M. fremontii*) as monophyletic with 100% bootstrap support (Figure 2.2A). Within these species, each population is also recovered as monophyletic, with the exception of one sample from a *M. brevipes* population (Figure 2.2A: one individual from population B11 clusters with B19). *M. johnstonii* is supported as sister to *M. brevipes*, in contrast to (Beardsley et al., 2004) which found *M. brevipes* and *M. fremontii* as sister based on just three loci. Genome-wide estimates of pairwise nucleotide divergence at synonymous sites also support strong divergence among the three focal species, with pairwise divergence values (d_s , 4.8-6.0%) consistently well outside the range of nucleotide

diversity within species (π_s , range 1.6-3.7%) (Figure 2.2B, Table S2.4). Heterozygosity is similar to pairwise diversity (h_s , range 1.1-3.1%), supporting a primarily outcrossing strategy in these species (Figure 2.2B, Table S2.4). Speciation times were estimated at >1 million generations for all species pairs (Figure 2.2A).

The Sespe Creek population (S25) is clearly a distinct lineage well outside of the *M. brevipes* trio (Figure 2.2A). Although S25 is most closely related to *M. constrictus*, high genetic divergence between these lineages ($d_s = 0.06$) suggest that S25 is a distinct species (Figure 2.2A-B, Table S2.4). Other species in this clade are not sampled and could be closer relatives to Sespe Creek, but none are known to occur in that geographic area. *M. nanus* is recovered as an outgroup to the other *Eunanus* section species as expected, with the *M. aurantiacus* complex set as the ultimate outgroup (Figure 2.2B). Species-level relationships from the maximum-likelihood tree were corroborated by a neighbor-joining tree and ASTRAL consensus tree (Figure S2.1).

Despite clear divergence between these *Mimulus* species, we also discovered a complex history of introgression. Using ABBA-BABA tests (Tables S2.5-S2.6) and associated F statistics (Figures 2.3A and S2.2), TWISST (Figure S2.3), and TreeMix (Figure 2.3B), we detected two signals of gene flow that were consistent across all methods, as well as additional signals that were more ambiguous (summarized in Figure S2.4). First, we found a strong signal of directional introgression from the *M. brevipes*-*M. fremontii*-*M. johnstonii* clade into the Sespe Creek population (Table S2.5: $D = 0.186$ - 0.199 ; Figure 2.3A: $F_{branch} = 0.072$ - 0.093 ; Figure S2.3: TWISST quartet scores 41.1/41.8/17.0, Figure 2.3B: TreeMix $m=0.193$). Second, we detected a signal of gene flow between *M. nanus* and the *M. constrictus* – Sespe Creek clade (Table S2.5: $D = 0.091$ - 0.131 , Figure 2.3A: $F_{branch} = 0.032$, Figure S2.3: TWISST quartet scores 73.5/15.2/11.3, Figure 2.3B: TreeMix $m=0.296$).

Introgression from *Sespe Creek* back into *M. fremontii* was supported by TWISST (Figure S2.3: quartet scores 70.3/15.7/14.0) and by ABBA-BABA tests for the 'complete' dataset (Table S2.5: $D = 0.080$; Figure 2.3A: $F_{branch} = 0.016$) but only some downsampled iterations of f -statistics (Table S2.6: median $D = 0.044$, Figure S2.2: median $F_{branch} = 0.007$) and not by TreeMix (Figure 2.3B). Introgression from *Sespe Creek* into *M. johnstonii* was supported by ABBA-BABA tests for the 'complete' dataset (Table S2.5: $D = 0.077$; Figure 2.3A: $F_{branch} = 0.023$) and by TreeMix (Figure 2.3B: $m=0.193$), but not by TWISST. Results from TWISST were in the opposite direction of D - and f -statistics for the trio (*M. fremontii*, *M. brevipes*, *M. johnstonii*) (Figure 2.3A), while downsampled datasets produced a wide range of F_{branch} values from 0 to 0.033 (Figure S2.2), indicating that this signal was unstable depending on which individuals were included in the dataset. We note that F_{branch} and TWISST results cannot explicitly test for the directionality of introgression, though they may show asymmetry depending on the direction and chosen taxa; we use our TreeMix results to infer the directionality of each introgression case when possible.

For the two strongest cases of species-level introgression, we chose the trio with the strongest signal and scanned the genome for signatures of introgression using the statistic df . For both comparisons, df windows were asymmetric, with both mean and median df greater than 0 and more positive outliers (signals of introgression in the same direction as the genome-wide F_{branch} pattern) than negative outliers of the same magnitude (Figure 2.4). Positive outliers were widely distributed throughout the genome, present on all ten linkage groups for both trios and with no large blocks of uninterrupted introgression, indicating that gene flow in this group is ancient rather than ongoing in these comparisons (Figure 2.4). Higher df values were generally associated with lower-diversity regions: top outliers were found in lower- π regions than the

genome-wide average, and df decreased across deciles of increasing pi (Table S2.7). However, df remained greater than 0 across all deciles of pi for all species and for both quartets, suggesting that elevated df is not driven only by low diversity (Table S2.7). Excluding df windows with $pi=0$ for any member of the trio had only a minor effect on the genome-wide average of df (Table S2.7).

*Multiple strong postmating reproductive barriers between species in the *M. brevipes* group*

Most crosses between four focal species (*M. johnstonii*, *M. brevipes*, *M. fremontii*, and the Sespe Creek population) showed little to no reduction in crossing success (probability of producing a seed, or seeds produced per fruit) compared to within-species crosses (Figures 2.5 and S5, Table S2.8). In one case (FxB), there is evidence of strong unidirectional postmating prezygotic isolation: only one in 58 crosses between maternal *M. fremontii* and paternal *M. brevipes* produced any seeds, a significantly lower proportion than the within-species or reciprocal cross types (Figure 2.4, Table S2.8). This pattern appears to be additive on both the maternal and paternal sides in *M. brevipes* x *M. fremontii* F1 hybrids, with intermediate crossing success in both HxB and FxH crosses (Figure S2.6).

A few other interspecific cross types also had low or zero seed production (e.g., reciprocal crosses of *M. fremontii* and Sespe Creek), which could indicate a postmating prezygotic barrier, but these crosses had low sample sizes and did not reach significance compared to within-species crosses (Figure 2.5). In addition, cross type BxJ had significantly lower fruit success than BxB, possibly indicating premating postzygotic isolation – however, JxJ crosses also had relatively low fruit success, making it difficult to rule out poor performance of *M. johnstonii* pollen overall. We note that these species are difficult to keep happy in growth chamber conditions, and some cases of fruit failure (such as low intraspecific cross success in *M.*

johnstonii) may be a byproduct of less-than-perfect growth conditions rather than the particulars of a given cross type.

For crosses that did produce seeds, we found strong F1 hybrid seed inviability in nine interspecific crosses: 0% viability for five cross types and <2% viability for four others (Figure 2.6A-B, Table S2.8). The exception was species pair *M. fremontii* and *M. brevipes*, which produced viable seeds in both directions (Figure 2.6A-B, Table S2.8). We could not score seed viability for cross type SxF because the only potential seeds produced were not distinguishable from unfertilized ovules. Seed viability results from the tetrazolium staining assay were qualitatively similar to results from scoring seeds by eye (Table S2.9).

M. johnstonii intraspecific seeds were significantly larger than *M. brevipes* or *M. fremontii* seeds, with Sespe Creek seeds intermediate in size (Figure 2.6C, Table S2.10). Inviabile seeds overall showed no evidence of overgrowth or undergrowth relative to parents, but instead tended to match the seed length of the maternal parent. However, F1 hybrids with Sespe Creek as maternal parent did seem to have slightly smaller seeds than either parent (this trend was only significant for SxF) (Figures 2.6A and S2.7A, Table S2.10). Crosses with *M. johnstonii*, the largest-seeded species, as the maternal parent produced inviable seeds that were collapsed and thin, as shown by a significant increase in hybrid seed length/width ratios (three cross types) compared to both parents and to the reciprocal cross type (Figures 2.6A and S2.7B, Table S2.10). In the reciprocal direction, hybrid seeds were typically shriveled, as were hybrid seeds in both directions of Sespe Creek crosses (Figures 2.6A and S2.7B).

Although in most cases, strong seed inviability prevented us from quantifying later-acting barriers, we were able to assess F1 hybrid sterility between *M. brevipes* and *M. fremontii*. Pollen number per flower was intermediate for hybrids compared to the two parental species (Figure

2.7A, Table S2.10). However, only ~20% of hybrid pollen grains were fertile, a sharp reduction from either parent, indicating strong F1 male sterility (Figure 2.7B, Table S2.11). We find no evidence of hybrid female sterility in *M. brevipes* x *M. fremontii* F1s; seed production from pollinated hybrids was intermediate compared to pure species (Figure 2.7C).

Total postmating reproductive isolation was near-complete ($RI \geq 0.95$) for all tested species pairs except for *M. brevipes* vs. *M. fremontii*, primarily driven by strong hybrid seed inviability (Table 2.1). The exception, *M. brevipes* vs. *M. fremontii*, had incomplete but still substantial total postmating RI from a combination of postmating prezygotic isolation and F1 male sterility (B \rightarrow F: $RI=0.91$; F \rightarrow B: $RI=0.66$).

Discussion

A major goal of speciation research is to understand similarities and differences in how reproductive isolation emerges as groups of species diverge. We find that *Mimulus brevipes*, *Mimulus johnstonii*, and *Mimulus fremontii* are clearly delineated species both genetically and by intrinsic reproductive barriers. We find strong hybrid seed inviability between multiple species pairs, as well as a postmating prezygotic barrier and severe hybrid male sterility between one species pair. While premating barriers have not yet been thoroughly quantified in this group, the combination of close geographic proximity (within centimeters in some cases) and similar floral morphology (excluding *M. brevipes*) makes it unlikely that premating barriers alone can eliminate hybridization. These strong postmating barriers are therefore likely to be important in preventing contemporary gene flow.

While many studies have shown that reproductive isolation increases with divergence on average, individual species pairs are often idiosyncratic (Coyne & Orr, 1989; Malone &

Fontenot, 2008; Matute & Cooper, 2021; Moyle et al., 2004). Our results highlight this heterogeneity – while reproductive isolation and divergence are both strong overall, the species pair with the least postmating reproductive isolation in this group (*M. brevipes* and *M. fremontii*) is not the most recently diverged pair (*M. brevipes* and *M. johnstonii*). Paying greater attention to the exceptions may help us better understand the factors that drive reproductive isolation.

Gene discordance and introgression

We find evidence of extensive phylogenetic discordance across the genome, including signatures of historical introgression between multiple species in this group. Our strongest signals of introgression appear between groups with essentially complete postmating barriers and without stark differences in floral morphology. These signals are likely ancient, and may have predated the postmating barriers we see today. Without postmating reproductive barriers, even small amounts of introgression can lead to species collapse in sympatry, especially if premating barriers are weak or are disrupted by environmental change (Behm et al., 2010; Kleindorfer et al., 2014; Xiong & Mallet, 2022). Alternatively, species can persist despite ongoing introgression, depending on the circumstances (Kay et al., 2018; Kenney & Sweigart, 2016; Servedio & Hermisson, 2020). Given the lack of obvious pre-zygotic barriers between these species, we argue that postmating reproductive barriers have likely played an important and ongoing role in reducing gene flow as divergence increases in section *Eunanus*, contradicting the hypothesis that postmating barriers are ‘after-the-fact’ byproducts with little influence on speciation trajectories. More work on pre-zygotic barriers in section *Eunanus* would help to confirm or refute this hypothesis.

Without further sampling of populations and species in this section, attributing these historical signals of introgression to precise events on the phylogeny is difficult. In particular,

gene flow involving unsampled lineages can result in incorrectly attributed ‘ghost’ introgression (Beerli, 2004; Slatkin, 2005; Tricou et al., 2022), and complex demographic histories may produce false positive signals of hybridization. Taken as a whole, however, these signals demonstrate a high degree of complexity in the divergence process, consistent with reticulate evolutionary histories of other groups within *Mimulus* (e.g., Brandvain et al., 2014; Nelson et al., 2021; Stankowski et al., 2019), in plants more broadly (e.g., Curtu et al., 2009; Goulet et al., 2017; Hamlin et al., 2020; Kay et al., 2018; Scascitelli et al., 2010), and across the tree of life (e.g., D’Angiolo et al., 2020; Green et al., 2010; Kleindorfer et al., 2014; Mallet et al., 2007; Schumer et al., 2018).

Cryptic diversity in Mimulus section Eunanus

Genetic data allow us to uncover cryptic variation in the form of distinct genetic lineages that are not readily distinguished by morphological features in the field (Bickford et al., 2007; Struck et al., 2018). Our population from Sespe Creek does not appear to match any of the likely species from the area, and we find that it is a genetically distinct lineage with a history of introgression. In addition, it appears to be reproductively isolated from other members of the group by strong hybrid seed inviability. We do not yet know the exact nature of its origin, or how widely distributed this lineage is; more sampling from the area and genetic data from more species in the group will be necessary to fully resolve these questions. Studies of speciation in plants often focus on ecologically mediated premating reproductive isolation, but cryptic diversity may be better explained by postmating barriers (Coughlan & Matute, 2020). *Mimulus* section *Eunanus* is home to many more understudied, small, pink-flowered taxa; if our study is any indication, strong postmating reproductive isolation may play an outsized role in generating diversity in this group.

Parallel evolution of hybrid seed inviability in Mimulus

Hybrid seed inviability has been found repeatedly in species pairs from the *Mimulus guttatus* species complex (section *Simiolus*) (Coughlan et al., 2020; Oneal et al., 2016; Sandstedt & Sweigart, 2022). By demonstrating hybrid seed inviability in a distant section of *Mimulus*, we show that this pattern is widespread and important, both across *Mimulus* and likely in plants more broadly. In the *M. guttatus* complex and other systems, hybrid seed inviability has been tied to parental conflict in resource allocation, a process mediated by the endosperm (Coughlan et al., 2020; Haig & Westoby, 1991; Lafon-Placette et al., 2017; Oneal et al., 2016; Rebernig et al., 2015; Roth et al., 2019; Sandstedt & Sweigart, 2022). We do not find clear evidence of parental conflict in our system -- hybrid seed sizes tended to track maternal seed sizes, without the telltale overgrowth or undergrowth phenotypes. However, hybrid seed size is not always a good predictor of parental conflict (Sandstedt & Sweigart, 2022), so future work to characterize differences in seed development would be required to rule out a conflict hypothesis.

Alternatively, differences in parental seed sizes between our species could result in Dobzhansky-Muller-like incompatibilities during hybrid seed development, without invoking parental conflict as a driver. Seed size is an important life history trait tied to ecological strategies; larger seeds tend to have more success establishing under a variety of hazardous conditions, from nutrient deprivation to low soil moisture to deep burial (Leishman et al., 2000). Larger, heavier seeds may also prevent long-range dispersal and keep offspring in maternal habitats for which they are locally adapted, as is the case in dune-adapted plants (Bowers, 1982; Schwarzbach et al., 2001). Both *M. johnstonii* and the Sespe Creek population are found on unstable scree slopes, which may pose a particular challenge to seed establishment or favor local dispersal to maintain local adaptation. This habitat preference is therefore a possible candidate for a selective force, which

could have acted in parallel in *M. johnstonii* and Sespe Creek to drive seed size differences leading to hybrid inviability.

Postmating prezygotic isolation

Crossing failure (postmating prezygotic isolation) can be caused by a failure of pollen tube germination, pollen tube growth, or fertilization (Wheeler et al., 2001); it may be a passive incompatibility between pollen and pistil, or an active rejection mechanism to prevent maladaptive hybridization (Hogenboom et al., 1975; Roda & Hopkins, 2018; Rushworth et al., 2022). Most studies on the mechanisms of postmating prezygotic isolation come from systems with self-incompatibility, where they are thought to be related to self-incompatibility mechanisms, e.g. *Solanum* (Bernacchi & Tanksley, 1997; Tovar-Méndez et al., 2014), *Nicotiana* (Kuboyama et al., 1994), and *Lilium* (Ascher & Drewlow, 1975), though a mechanism unrelated to self-incompatibility has been described in *Brassica* (Fujii et al., 2019). Since *Mimulus* lacks self-incompatibility mechanisms, this group provides an opportunity to investigate how interspecific incompatibility may arise on its own. In other systems with differences in style length, germinating pollen tubes can under- or over-shoot the ovules (Gore et al., 1990; Williams & Rouse, 1988). Style length could be important in our system, since *M. brevipes* styles are substantially longer than *M. fremontii* styles, though our case is unusual in that pollen from the long-styled parent fails while pollen from the short-styled parent is successful. Pollen competition driving differential coevolution of the pollen and pistil is another possible source of incompatibilities (Brandvain & Haig, 2005; Skogsmyr & Lankinen, 2002), a scenario which has been documented in the *M. guttatus* complex (Aagaard et al., 2013; Fishman et al., 2008) and *Arabidopsis* (Takeuchi & Higashiyama, 2012).

Hybrid male sterility

Hybrid male sterility has been characterized multiple times in *Mimulus* sections *Simiolus* and *Erythranthe*, with a variety of genetic causes, including a nuclear DMI interaction (Sweigart et al., 2006), a cytoplasmic-nuclear interaction (Fishman & Willis, 2006), or underdominant chromosomal rearrangements (Stathos & Fishman, 2014), but the relative frequency of these mechanisms across plants is unknown. *M. brevipes* and *M. fremontii* are more genetically divergent than most pairs for which male sterility has been studied, in *Mimulus* or elsewhere. Since they still produce viable hybrids and sterility is not 100% complete, genetic mapping could be used in the future to determine the mechanisms underlying sterility in this system, and the degree of parallelism with other cases in *Mimulus*.

Future directions

As the first in-depth genomic and phenotypic investigation of speciation within *Mimulus* section *Eunanus*, this study fills an important phylogenetic gap within the speciation literature, helping to position the genus *Mimulus* as a leading model for broad phylogenetically informed comparisons of how species form and diverge. Our study system highlights both universal and idiosyncratic patterns in the speciation process; we add to a growing body of evidence that hybrid seed inviability and hybrid male sterility are important barriers to gene flow in *Mimulus* and across plant taxa; and we demonstrate the complexity of gene discordance and historical introgression among species despite substantial overall divergence. In addition, we lay the groundwork for fruitful avenues of future mechanistic study: seed size evolution and hybrid seed inviability; style length and postmating prezygotic isolation; and highly divergent hybrid male sterility are all worthy of future exploration in this group. Our study also provides the starting point for a more complete phylogenetic sampling of section *Eunanus*, which would enhance our

understanding of cryptic divergence and the interplay between ecological, geographic, and genetic forces driving the diversification of species.

References

- Aagaard, J. E., George, R. D., Fishman, L., MacCoss, M. J., & Swanson, W. J. (2013). Selection on Plant Male Function Genes Identifies Candidates for Reproductive Isolation of Yellow Monkeyflowers. *PLoS Genetics*, 9(12). <https://doi.org/10.1371/journal.pgen.1003965>
- Ascher, P. D., & Drewlow, L. W. (1975). The effect of prepollination injection with stigmatic exudate on interspecific pollen tube growth in *Lilium longiflorum* thunb. Styles. *Plant Science Letters*, 4(6), 401–405. [https://doi.org/10.1016/0304-4211\(75\)90269-2](https://doi.org/10.1016/0304-4211(75)90269-2)
- Baack, E., Melo, M. C., Rieseberg, L. H., & Ortiz-Barrientos, D. (2015). The origins of reproductive isolation in plants. *New Phytologist*, 207(4), 968–984. <https://doi.org/10.1111/nph.13424>
- Baldwin, B. G., Goldman, D., Keil, D. J., Patterson, R., Rosatti, T. J., & Wilken, D. (Eds.). (2012). *The Jepson Manual: Vascular Plants of California, Thoroughly Revised and Expanded* (2nd ed.).
- Barker, W. R. (Bill), Nesom, G. L., Beardsley, P. M., & Fraga, N. S. (2012). A taxonomic conspectus of phrymaceae: A narrowed circumscription for. *Phytoneuron*, May, 1–60.
- Beardsley, P. M., Schoenig, S. E., Whittall, J. B., & Olmstead, R. G. (2004). Patterns of evolution in western North American *Mimulus* (Phrymaceae). *American Journal of Botany*, 91(3), 474–489. <https://doi.org/10.3732/ajb.91.3.474>
- Beerli, P. (2004). Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. *Molecular Ecology*, 13(4), 827–836. <https://doi.org/10.1111/j.1365-294X.2004.02101.x>
- Behm, J. E., Ives, A. R., & Boughman, J. W. (2010). Breakdown in postmating isolation and the collapse of a species pair through hybridization. *The American Naturalist*, 175(1), 11–26. <https://doi.org/10.1086/648559>
- Bernacchi, D., & Tanksley, S. D. (1997). An Interspecific Backcross of *Lycopersicon esculentum* × *L. hirsutum*: Linkage Analysis and a QTL Study of Sexual Compatibility Factors and Floral Traits. *Genetics*, 147(2), 861–877. <https://doi.org/10.1093/genetics/147.2.861>
- Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., Ingram, K. K., & Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, 22(3), 148–155. <https://doi.org/10.1016/j.tree.2006.11.004>

- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bowers, J. E. (1982). The plant ecology of inland dunes in western North America. *Journal of Arid Environments*, 5(3), 199–220. [https://doi.org/10.1016/S0140-1963\(18\)31444-7](https://doi.org/10.1016/S0140-1963(18)31444-7)
- Brandvain, Y., & Haig, D. (2005). Divergent Mating Systems and Parental Conflict as a Barrier to Hybridization in Flowering Plants. *The American Naturalist*, 166(3), 330–338. <https://doi.org/10.1086/432036>
- Brandvain, Y., Kenney, A. M., Flagel, L., Coop, G., & Sweigart, A. L. (2014). Speciation and Introgression between *Mimulus nasutus* and *Mimulus guttatus*. *PLoS Genetics*, 10(6), e1004410. <https://doi.org/10.1371/journal.pgen.1004410>
- Broad Institute. (2019). *Picard Toolkit*, *GitHub repository* [Computer software]. <https://broadinstitute.github.io/picard/>
- Charron, G., Leducq, J. B., & Landry, C. R. (2014). Chromosomal variation segregates within incipient species and correlates with reproductive isolation. *Molecular Ecology*, 23(17), 4362–4372. <https://doi.org/10.1111/mec.12864>
- Christie, K., Fraser, L. S., & Lowry, D. B. (2022). The strength of reproductive isolating barriers in seed plants: Insights from studies quantifying premating and postmating reproductive barriers over the past 15 years. *Evolution*, evo.14565. <https://doi.org/10.1111/evo.14565>
- Christie, K., & Strauss, S. Y. (2018). Along the speciation continuum: Quantifying intrinsic and extrinsic isolating barriers across five million years of evolutionary divergence in California jewelflowers. *Evolution*, 72(5), 1063–1079. <https://doi.org/10.1111/evo.13477>
- Coughlan, J. M., & Matute, D. R. (2020). The importance of intrinsic postzygotic barriers throughout the speciation process: Intrinsic barriers throughout speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1806). <https://doi.org/10.1098/rstb.2019.0533>
- Coughlan, J. M., Wilson Brown, M., & Willis, J. H. (2020). Patterns of Hybrid Seed Inviability in the *Mimulus guttatus* sp. Complex Reveal a Potential Role of Parental Conflict in Reproductive Isolation. *Current Biology*, 30(1), 83–93.e5. <https://doi.org/10.1016/j.cub.2019.11.023>
- Coyne, J. A., & Orr, H. A. (1989). Patterns of Speciation in *Drosophila*. *Evolution*, 43(2), 362–381.
- Coyne, J. A., & Orr, H. A. (1997). Patterns of Speciation in *Drosophila* Revisited. *Evolution*, 51(1), 295. <https://doi.org/10.2307/2410984>
- Crespi, B., & Nosil, P. (2013). Conflictual speciation: Species formation via genomic conflict. *Trends in Ecology & Evolution*, 28(1), 48–57. <https://doi.org/10.1016/j.tree.2012.08.015>

- Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, 23(13), 3133–3157. <https://doi.org/10.1111/mec.12796>
- Curtu, A. L., Gailing, O., & Finkeldey, R. (2009). Patterns of contemporary hybridization inferred from paternity analysis in a four-oak-species forest. *BMC Evolutionary Biology*, 9(1), 284. <https://doi.org/10.1186/1471-2148-9-284>
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, 10(2), giab008. <https://doi.org/10.1093/gigascience/giab008>
- D’Angiolo, M., De Chiara, M., Yue, J.-X., Irizar, A., Stenberg, S., Persson, K., Llored, A., Barré, B., Schacherer, J., Marangoni, R., Gilson, E., Warringer, J., & Liti, G. (2020). A yeast living ancestor reveals the origin of genomic introgressions. *Nature*, 587(7834), Article 7834. <https://doi.org/10.1038/s41586-020-2889-1>
- Darwin, C. (1859). *On the origin of species by means of natural selection*. John Murray. http://darwin-online.org.uk/converted/pdf/1861_OriginNY_F382.pdf
- Duranton, M., Allal, F., Fraïsse, C., Bierne, N., Bonhomme, F., & Gagnaire, P. A. (2018). The origin and remolding of genomic islands of differentiation in the European sea bass. *Nature Communications*, 9(1), 1–11. <https://doi.org/10.1038/s41467-018-04963-6>
- Ferris, K. G., Barnett, L. L., Blackman, B. K., & Willis, J. H. (2017). The genetic architecture of local adaptation and reproductive isolation in sympatry within the *Mimulus guttatus* species complex. *Molecular Ecology*, 26(1), 208–224. <https://doi.org/10.1111/mec.13763>
- Ferris, K. G., Sexton, J. P., & Willis, J. H. (2014). Speciation on a local geographic scale: The evolution of a rare rock outcrop specialist in *Mimulus*. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1648), 20140001. <https://doi.org/10.1098/rstb.2014.0001>
- Fishman, L. (2020). 96-well CTAB-chloroform DNA extraction. protocols.io. <https://dx.doi.org/10.17504/protocols.io.bgv6jw9e>
- Fishman, L., Aagaard, J., & Tuthill, J. C. (2008). Toward the Evolutionary Genomics of Gametophytic Divergence: Patterns of Transmission Ratio Distortion in Monkeyflower (*mimulus*) Hybrids Reveal a Complex Genetic Basis for Conspecific Pollen Precedence. *Evolution*, 62(12), 2958–2970. <https://doi.org/10.1111/j.1558-5646.2008.00475.x>
- Fishman, L., Sweigart, A. L., Kenney, A. M., & Campbell, S. (2014). Major quantitative trait loci control divergence in critical photoperiod for flowering between selfing and outcrossing species of monkeyflower (*Mimulus*). *New Phytologist*, 201(4), 1498–1507. <https://doi.org/10.1111/nph.12618>

- Fishman, L., & Willis, J. H. (2006). A cytonuclear incompatibility causes anther sterility in *Mimulus* hybrids. *Evolution*, 60(7), 1372–1381. <https://doi.org/10.1111/j.0014-3820.2006.tb01216.x>
- Fitak, R. R. (2021). OptM: Estimating the optimal number of migration edges on population trees using Treemix. *Biology Methods and Protocols*, 6(1), bpab017. <https://doi.org/10.1093/biomethods/bpab017>
- Fujii, S., Tsuchimatsu, T., Kimura, Y., Ishida, S., Tangpranomkorn, S., Shimosato-Asano, H., Iwano, M., Furukawa, S., Itoyama, W., Wada, Y., Shimizu, K. K., & Takayama, S. (2019). A stigmatic gene confers interspecies incompatibility in the Brassicaceae. *Nature Plants*, 5(7), Article 7. <https://doi.org/10.1038/s41477-019-0444-6>
- Fuller, Z. L., Leonard, C. J., Young, R. E., Schaeffer, S. W., & Phadnis, N. (2018). Ancestral polymorphisms explain the role of chromosomal inversions in speciation. *PLOS Genetics*, 14(7), e1007526. <https://doi.org/10.1371/journal.pgen.1007526>
- Gore, P., Potts, B., Volker, P., & Megalos, J. (1990). Unilateral Cross-Incompatibility in *Eucalyptus*: The Case of Hybridisation Between *E. globulus* and *E. nitens*. *Australian Journal of Botany*, 38(4), 383. <https://doi.org/10.1071/BT9900383>
- Goulet, B. E., Roda, F., & Hopkins, R. (2017). Hybridization in plants: Old ideas, new techniques. *Plant Physiology*, 173(1), 65–78. <https://doi.org/10.1104/pp.16.01340>
- Grant, A. L. (1924). A Monograph of the Genus *Mimulus*. *Annals of the Missouri Botanical Garden*, 11(2/3), 99–388. <https://doi.org/10.2307/2394024>
- Grant, V. (1981). *Plant Speciation*. Columbia University Press.
- Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M. H.-Y., Hansen, N. F., Durand, E. Y., Malaspinas, A.-S., Jensen, J. D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H. A., ... Pääbo, S. (2010). A draft sequence of the Neandertal genome. *Science (New York, N.Y.)*, 328(5979), 710–722. <https://doi.org/10.1126/science.1188021>
- Guerrero, R. F., Muir, C. D., Josway, S., & Moyle, L. C. (2017). Pervasive antagonistic interactions among hybrid incompatibility loci. *PLOS Genetics*, 13(6), e1006817. <https://doi.org/10.1371/journal.pgen.1006817>
- Haig, D., & Westoby, M. (1991). Genomic imprinting in endosperm: Its effect on seed development in crosses between species, and between different ploidies of the same species, and its implications for the evolution of apomixis. *Philosophical Transactions - Royal Society of London, B*, 333, 1–13.
- Hamlin, J. A. P., Hibbins, M. S., & Moyle, L. C. (2020). Assessing biological factors affecting postspeciation introgression. *Evolution Letters*, 4(2), 137–154. <https://doi.org/10.1002/evl3.159>

- Hogenboom, N. G., Mather, K., Heslop-Harrison, J., & Lewis, D. (1975). Incompatibility and incongruity: Two different mechanisms for the non-functioning of intimate partner relationships. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 188(1092), 361–375. <https://doi.org/10.1098/rspb.1975.0025>
- Kay, K. M., Woolhouse, S., Smith, B. A., Pope, N. S., & Rajakaruna, N. (2018). Sympatric serpentine endemic *Monardella* (Lamiaceae) species maintain habitat differences despite hybridization. *Molecular Ecology*, 27(9), 2302–2316. <https://doi.org/10.1111/mec.14582>
- Kenney, A. M., & Sweigart, A. L. (2016). Reproductive isolation and introgression between sympatric *Mimulus* species. *Molecular Ecology*, 25(11), 2499–2517. <https://doi.org/10.1111/mec.13630>
- Kirkpatrick, M., & Barrett, B. (2015). Chromosome inversions, adaptive cassettes and the evolution of species' ranges. *Molecular Ecology*, 24(9), 2046–2055. <https://doi.org/10.1111/mec.13074>
- Kleindorfer, S., O'Connor, J. A., Dudaniec, R. Y., Myers, S. A., Robertson, J., & Sulloway, F. J. (2014). Species collapse via hybridization in Darwin's tree finches. *The American Naturalist*, 183(3), 325–341. <https://doi.org/10.1086/674899>
- Koch, M. A., Haubold, B., & Mitchell-Olds, T. (2000). Comparative Evolutionary Analysis of Chalcone Synthase and Alcohol Dehydrogenase Loci in *Arabidopsis*, *Arabis*, and Related Genera (Brassicaceae). *Molecular Biology and Evolution*, 17(10), 1483–1498. <https://doi.org/10.1093/oxfordjournals.molbev.a026248>
- Korunes, K. L., & Samuk, K. (2021). pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Molecular Ecology Resources*, 21(4), 1359–1368. <https://doi.org/10.1111/1755-0998.13326>
- Kryvokhyzha, D. (n.d.). *GATK: The best practice for genotype calling in a non-model organism*. Dmytro Kryvokhyzha - Bioinformatics & Genomics Scientist. Retrieved August 8, 2022, from <https://evodify.com/gatk-in-non-model-organism/>
- Kuboyama, T., Chung, C. S., & Takeda, G. (1994). The diversity of interspecific pollen-pistil incongruity in *Nicotiana*. *Sexual Plant Reproduction*, 7(4), 250–258. <https://doi.org/10.1007/BF00232744>
- Lafon-Placette, C., Johannessen, I. M., Hornslien, K. S., Ali, M. F., Bjerkan, K. N., Bramsiepe, J., Glöckle, B. M., Rebernig, C. A., Brysting, A. K., Grini, P. E., & Köhler, C. (2017). Endosperm-based hybridization barriers explain the pattern of gene flow between *Arabidopsis lyrata* and *Arabidopsis arenosa* in Central Europe. *Proceedings of the National Academy of Sciences*, 114(6), E1027–E1035. <https://doi.org/10.1073/pnas.1615123114>
- Lafon-Placette, C., & Köhler, C. (2016). Endosperm-based postzygotic hybridization barriers: Developmental mechanisms and evolutionary drivers. *Molecular Ecology*, 25(11), 2620–2629. <https://doi.org/10.1111/mec.13552>

- Le Gac, M., Hood, M. E., & Giraud, T. (2007). Evolution of Reproductive Isolation Within a Parasitic Fungal Species Complex. *Evolution*, 61(7), 1781–1787. <https://doi.org/10.1111/j.1558-5646.2007.00144.x>
- Leishman, M., Wright, I., Moles, A., & Westoby, M. (2000). The Evolutionary Ecology of Seed Size. *Seeds: The Ecology of Regeneration in Plant Communities*, 2. <https://doi.org/10.1079/9780851994321.0031>
- Li, H. (2018). *Seqtk* (Version 1.3) [Computer software]. <https://github.com/lh3/seqtk>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Lowry, D. B., Sobel, J. M., Angert, A. L., Ashman, T.-L., Baker, R. L., Blackman, B. K., Brandvain, Y., Byers, K. J. R. P., Cooley, A. M., Coughlan, J. M., Dudash, M. R., Fenster, C. B., Ferris, K. G., Fishman, L., Friedman, J., Grossenbacher, D. L., Holeski, L. M., Ivey, C. T., Kay, K. M., ... Yuan, Y.-W. (2019). The case for the continued use of the genus name *Mimulus* for all monkeyflowers. *TAXON*, 68(4), 617–623. <https://doi.org/10.1002/tax.12122>
- Lowry, D. B., & Willis, J. H. (2010). A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biology*, 8(9). <https://doi.org/10.1371/journal.pbio.1000500>
- Malinsky, M., Matschiner, M., & Svardal, H. (2021). Dsuite—Fast D-statistics and related admixture evidence from VCF files. *Molecular Ecology Resources*, 21(2), 584–595. <https://doi.org/10.1111/1755-0998.13265>
- Mallet, J., Beltrán, M., Neukirchen, W., & Linares, M. (2007). Natural hybridization in heliconiine butterflies: The species boundary as a continuum. *BMC Evolutionary Biology*, 7(1), 28. <https://doi.org/10.1186/1471-2148-7-28>
- Malone, J. H., & Fontenot, B. E. (2008). Patterns of Reproductive Isolation in Toads. *PLoS ONE*, 3(12), e3900. <https://doi.org/10.1371/journal.pone.0003900>
- Martin, S. H., & Van Belleghem, S. M. (2017). Exploring Evolutionary Relationships Across the Genome Using Topology Weighting. *Genetics*, 206(1), 429–438. <https://doi.org/10.1534/genetics.116.194720>
- Matute, D. R., & Cooper, B. S. (2021). Comparative studies on speciation: 30 years since Coyne and Orr. *Evolution*, 1989, 1–15. <https://doi.org/10.1111/evo.14181>
- Moyle, L. C., Olson, M. S., & Tiffin, P. (2004). Patterns of reproductive isolation in three angiosperm genera. *Evolution*, 58(6), 1195–1208. <https://doi.org/10.1111/j.0014-3820.2004.tb01700.x>

- Moyle, L. C., & Payseur, B. A. (2009). Reproductive isolation grows on trees. *Trends in Ecology & Evolution*, 24(11), 591–598. <https://doi.org/10.1016/j.tree.2009.05.010>
- Nelson, T. C., Stathos, A. M., Vanderpool, D. D., Finseth, F. R., Yuan, Y., & Fishman, L. (2021). Ancient and recent introgression shape the evolutionary history of pollinator adaptation and speciation in a model monkeyflower radiation (*Mimulus* section *Erythranthe*). *PLoS Genetics*, 17(2), e1009095. <https://doi.org/10.1371/journal.pgen.1009095>
- Nesom, G. L. (2013). Taxonomic notes on *Diplacus* (Phrymaceae). *Phytoneuron*, 8.
- Noor, M. A. F., Grams, K. L., Bertucci, L. A., & Reiland, J. (2001). Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences*, 98(21), 12084–12088. <https://doi.org/10.1073/pnas.221274498>
- Oneal, E., Willis, J. H., & Franks, R. G. (2016). Disruption of endosperm development is a major cause of hybrid seed inviability between *Mimulus guttatus* and *Mimulus nudatus*. *New Phytologist*, 210(3), 1107–1120. <https://doi.org/10.1111/nph.13842>
- Osmolovsky, I., Shifrin, M., Gamliel, I., Belmaker, J., & Sapir, Y. (2022). Eco-Geography and Phenology Are the Major Drivers of Reproductive Isolation in the Royal Irises, a Species Complex in the Course of Speciation. *Plants*, 11(23), Article 23. <https://doi.org/10.3390/plants11233306>
- Pickrell, J. K., & Pritchard, J. K. (2012). Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. *PLoS Genetics*, 8(11). <https://doi.org/10.1371/journal.pgen.1002967>
- Presgraves, D. C. (2002). Patterns of Postzygotic Isolation in Lepidoptera. *Evolution*, 56(6), 1168–1183. <https://doi.org/10.1111/j.0014-3820.2002.tb01430.x>
- Rabiee, M., Sayyari, E., & Mirarab, S. (2019). Multi-allele species reconstruction using ASTRAL. *Molecular Phylogenetics and Evolution*, 130, 286–296. <https://doi.org/10.1016/j.ympev.2018.10.033>
- Ramsey, J., Bradshaw, H. D., & Schemske, D. W. (2003). Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae). *Evolution*, 57(7), 1520–1534. <https://doi.org/10.1111/j.0014-3820.2003.tb00360.x>
- Raunsgard, A., Opedal, Ø. H., Ekrem, R. K., Wright, J., Bolstad, G. H., Armbruster, W. S., & Pélabon, C. (2018). Intersexual conflict over seed size is stronger in more outcrossed populations of a mixed-mating plant. *Proceedings of the National Academy of Sciences*, 115(45), 11561–11566. <https://doi.org/10.1073/pnas.1810979115>
- Rebernick, C. A., Lafon-Placette, C., Hatorangan, M. R., Slotte, T., & Köhler, C. (2015). Non-reciprocal Interspecies Hybridization Barriers in the *Capsella* Genus Are Established in the Endosperm. *PLOS Genetics*, 11(6), e1005295. <https://doi.org/10.1371/journal.pgen.1005295>

- Renaut, S., Grassa, C. J., Yeaman, S., Moyers, B. T., Lai, Z., Kane, N. C., Bowers, J. E., Burke, J. M., & Rieseberg, L. H. (2013). Genomic islands of divergence are not affected by geography of speciation in sunflowers. *Nature Communications*, 4(1), 1827. <https://doi.org/10.1038/ncomms2833>
- Roda, F., & Hopkins, R. (2018). Correlated evolution of self and interspecific incompatibility across the range of a Texas wildflower. *New Phytologist*. <https://doi.org/10.1111/nph.15340>
- Roth, M., Florez-Rueda, A. M., & Städler, T. (2019). Differences in Effective Ploidy Drive Genome-Wide Endosperm Expression Polarization and Seed Failure in Wild Tomato Hybrids. *Genetics*, 212(1), 141–152. <https://doi.org/10.1534/genetics.119.302056>
- Rushworth, C. A., Wardlaw, A. M., Ross-Ibarra, J., & Brandvain, Y. (2022). Conflict over fertilization underlies the transient evolution of reinforcement. *PLOS Biology*, 20(10), e3001814. <https://doi.org/10.1371/journal.pbio.3001814>
- Sackton, T. (2022). *Natural selection constrains neutral diversity across a wide range of species* [R]. https://github.com/tsackton/linked-selection/blob/da62043544dce9c9342f0a393f41c63bec292fcf/misc_scripts/Identify_4D_Sites.pl (Original work published 2014)
- Sandstedt, G. D., & Sweigart, A. L. (2022). Developmental evidence for parental conflict in driving *Mimulus* species barriers. *New Phytologist*, 236(4), 1545–1557. <https://doi.org/10.1111/nph.18438>
- Sandstedt, G. D., Wu, C. A., & Sweigart, A. L. (2020). Evolution of multiple postzygotic barriers between species of the *Mimulus tilingii* complex. *Evolution*, 75, 600–613. <https://doi.org/10.1111/evo.14105>
- Scascitelli, M., Whitney, K. D., Randell, R. A., King, M., Buerkle, C. A., & Rieseberg, L. H. (2010). Genome scan of hybridizing sunflowers from Texas (*Helianthus annuus* and *H. debilis*) reveals asymmetric patterns of introgression and small islands of genomic differentiation. *Molecular Ecology*, 19(3), 521–541. <https://doi.org/10.1111/j.1365-294X.2009.04504.x>
- Schliep, K. P. (2011). phangorn: Phylogenetic analysis in R. *Bioinformatics*, 27(4), 592–593. <https://doi.org/10.1093/bioinformatics/btq706>
- Schumer, M., Xu, C., Powell, D. L., Durvasula, A., Skov, L., Holland, C., Blazier, J. C., Sankararaman, S., Andolfatto, P., Gil, †, Rosenthal, G., & Przeworski, M. (2018). Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science*, 360(6389), 656–660.
- Schwarzbach, A. E., Donovan, L. A., & Rieseberg, L. H. (2001). Transgressive character expression in a hybrid sunflower species. *American Journal of Botany*, 88(2), 270–277. <https://doi.org/10.2307/2657018>

- Scopece, G., Musacchio, A., Widmer, A., & Cozzolino, S. (2007). Patterns of reproductive isolation in Mediterranean deceptive orchids. *Evolution*, 61(11), 2623–2642. <https://doi.org/10.1111/j.1558-5646.2007.00231.x>
- Servedio, M. R., & Hermisson, J. (2020). The evolution of partial reproductive isolation as an adaptive optimum. *Evolution*, 74(1), 4–14. <https://doi.org/10.1111/evo.13880>
- Skogsmyr, I., & Lankinen, Å. (2002). Sexual selection: An evolutionary force in plants? *Biological Reviews*, 77(4), 537–562. <https://doi.org/10.1017/S1464793102005973>
- Slatkin, M. (2005). Seeing ghosts: The effect of unsampled populations on migration rates estimated for sampled populations. *Molecular Ecology*, 14(1), 67–73. <https://doi.org/10.1111/j.1365-294X.2004.02393.x>
- Sobel, J. M. (2010). *Speciation in the western North American wildflower genus Mimulus* [Ph.D., Michigan State University]. <https://www.proquest.com/docview/873378456/abstract/950327CDE46D4690PQ/1>
- Sobel, J. M. (2014). Ecogeographic Isolation and Speciation in the Genus *Mimulus*. *The American Naturalist*, 184(5), 565–579. <https://doi.org/10.1086/678235>
- Sobel, J. M., & Chen, G. F. (2014). Unification of methods for estimating the strength of reproductive isolation. *Evolution*, 68(5), 1511–1522. <https://doi.org/10.1111/evo.12362>
- Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
- Stankowski, S., Chase, M. A., Fuiten, A. M., Rodrigues, M. F., Ralph, P. L., & Streisfeld, M. A. (2019). Widespread selection and gene flow shape the genomic landscape during a radiation of monkeyflowers. *PLOS Biology*, 17(7), e3000391. <https://doi.org/10.1371/journal.pbio.3000391>
- Stathos, A., & Fishman, L. (2014). Chromosomal rearrangements directly cause underdominant F1 pollen sterility in *Mimulus lewisii*–*Mimulus cardinalis* hybrids. *Evolution*, 68(11), 3109–3119. <https://doi.org/10.1111/evo.12503>
- Streisfeld, M. A., & Kohn, J. R. (2005). Contrasting patterns of floral and molecular variation across a cline in *Mimulus aurantiacus*. *Evolution*, 59(12), 2548–2559. <https://doi.org/10.1111/j.0014-3820.2005.tb00968.x>
- Streisfeld, M. A., & Rausher, M. D. (2009). Altered trans-Regulatory Control of Gene Expression in Multiple Anthocyanin Genes Contributes to Adaptive Flower Color Evolution in *Mimulus aurantiacus*. *Molecular Biology and Evolution*, 26(2), 433–444. <https://doi.org/10.1093/molbev/msn268>
- Struck, T. H., Feder, J. L., Bendiksby, M., Birkeland, S., Cerca, J., Gusarov, V. I., Kistenich, S., Larsson, K.-H., Liow, L. H., Nowak, M. D., Stedje, B., Bachmann, L., & Dimitrov, D.

- (2018). Finding Evolutionary Processes Hidden in Cryptic Species. *Trends in Ecology & Evolution*, 33(3), 153–163. <https://doi.org/10.1016/j.tree.2017.11.007>
- Sweigart, A. L., Fishman, L., & Willis, J. H. (2006). A simple genetic incompatibility causes hybrid male sterility in *mimulus*. *Genetics*, 172(4), 2465–2479. <https://doi.org/10.1534/genetics.105.053686>
- Takeuchi, H., & Higashiyama, T. (2012). A Species-Specific Cluster of Defensin-Like Genes Encodes Diffusible Pollen Tube Attractants in *Arabidopsis*. *PLOS Biology*, 10(12), e1001449. <https://doi.org/10.1371/journal.pbio.1001449>
- Tovar-Méndez, A., Kumar, A., Kondo, K., Ashford, A., Baek, Y. S., Welch, L., Bedinger, P. A., & McClure, B. A. (2014). Restoring pistil-side self-incompatibility factors recapitulates an interspecific reproductive barrier between tomato species. *Plant Journal*, 77(5), 727–736. <https://doi.org/10.1111/tpj.12424>
- Tricou, T., Tannier, E., & de Vienne, D. M. (2022). Ghost Lineages Highly Influence the Interpretation of Introgression Tests. *Systematic Biology*, 71(5), 1147–1158. <https://doi.org/10.1093/sysbio/syac011>
- Van der Auwera, G., & O'Connor, B. (2020). *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra (1st Edition)*. O'Reilly Media.
- Wheeler, M. J., Franklin-Tong, V. E., & Franklin, F. C. H. (2001). The molecular and genetic basis of pollen–pistil interactions. *New Phytologist*, 151(3), 565–584. <https://doi.org/10.1046/j.0028-646x.2001.00229.x>
- Williams, E., & Rouse, J. (1988). Disparate Style Lengths Contribute to Isolation of Species in *Rhododendron*. *Australian Journal of Botany*, 36(2), 183. <https://doi.org/10.1071/BT9880183>
- Xiong, T., & Mallet, J. (2022). On the impermanence of species: The collapse of genetic incompatibilities in hybridizing populations. *Evolution*, 76(11), 2498–2512. <https://doi.org/10.1111/evo.14626>
- Yuan, Y.-W. (2019). Monkeyflowers (*Mimulus*): New model for plant developmental genetics and evo-devo. *New Phytologist*, 222(2), 694–700. <https://doi.org/10.1111/nph.15560>
- Yuan, Y.-W., Sagawa, J. M., Young, R. C., Christensen, B. J., & Bradshaw, H. D., Jr. (2013). Genetic Dissection of a Major Anthocyanin QTL Contributing to Pollinator-Mediated Reproductive Isolation Between Sister Species of *Mimulus*. *Genetics*, 194(1), 255–263. <https://doi.org/10.1534/genetics.112.146852>
- Zuellig, M. P., & Sweigart, A. L. (2018). A two-locus hybrid incompatibility is widespread, polymorphic, and active in natural populations of *Mimulus*. *Evolution*, 72(11), 2394–2405. <https://doi.org/10.1111/evo.13596>

Table 2.1: Summary of postmating reproductive isolation (RI).

Pair a, b	Crossing success RI		Seed viability RI		F1 viability RI		Total measured RI	
	a→b	b→a	a→b	b→a	a→b	b→a	a→b	b→a
J, B	0.30	0.10	0.99	1.00	--	--	0.99	1.00
J, F	-0.13	-0.05	0.97	1.00	--	--	0.97	1.00
B, F	0.91	0.13	-0.03	0.09	--	0.57	0.91	0.66
J, S	0.35	-0.07	1.00	0.97	--	--	1.00	0.96
B, S	0.05	0.16	1.00	0.99	--	--	1.00	0.99
F, S	1.00	0.68	--	1.00	--	--	1.00	1.00

a→b indicates reproductive isolation preventing gene flow from species a into species b, calculated by comparing the fitness of b x a crosses relative to b x b crosses, with b→a indicating the reciprocal direction using the fitness of a x b crosses relative to a x a crosses, following (Sobel & Chen, 2014). Values range from -1 (complete heterosis) to 0 (free mating) to 1 (complete reproductive isolation).

Bolded values include contributions from barriers with significant model results.

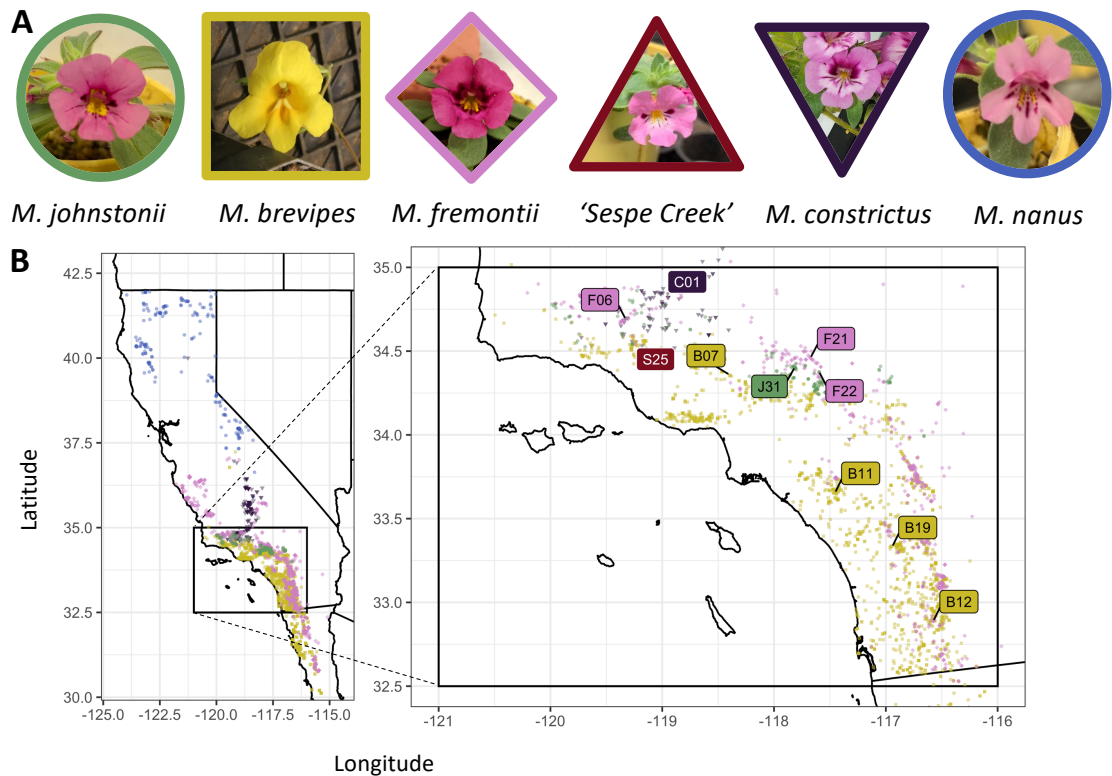


Figure 2.1. Species and populations sampled in this study. (A) Representative floral images for sampled species grown in growth chambers. (B) Map of occurrence records and sampled populations from the GBIF database for five species within California, USA and Baja California, Mexico. Smaller icons indicate GBIF records (GBIF.org, 2022) for each species, which include iNaturalist research grade observations and herbarium collections. Inset shows focal region of southern California where species ranges overlap. Larger named icons indicate sampled populations used in this study. The range of *M. nanus* extends into Nevada, Oregon, and Washington, but only occurrences for California are shown; population N01 (*M. nanus*) was collected from Washington state and is not shown. Population S25, ‘Sespe Creek’, is treated as a separate lineage despite being initially collected as *M. johnstonii*.

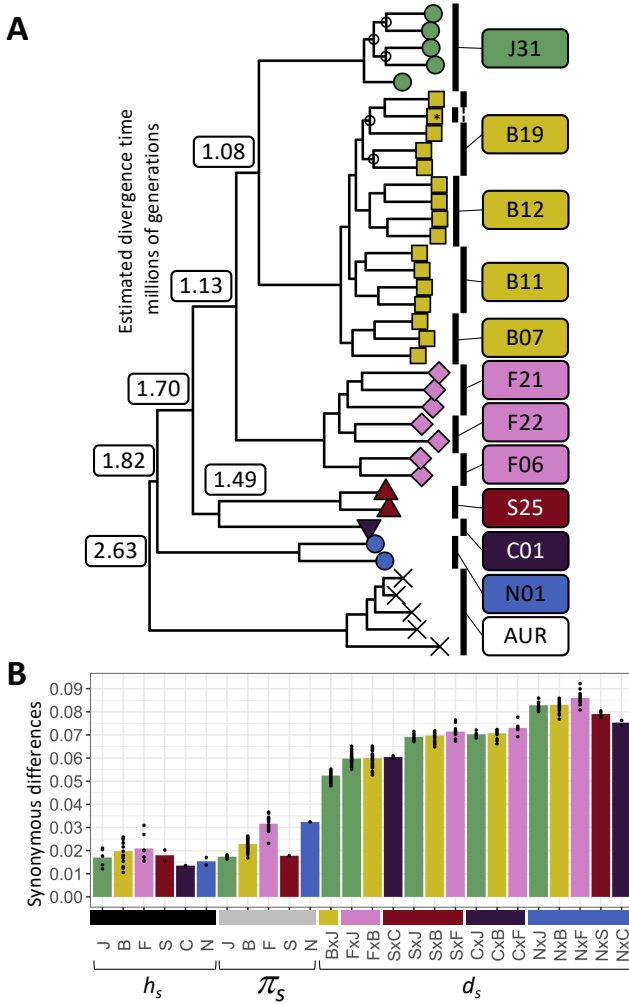


Figure 2.2. Phylogenetic relationships and divergence between sampled species in *Mimulus* section *Eunanus*. (A) Maximum likelihood phylogeny inferred using RAxML from the ‘complete’ dataset of genome-wide SNPs, rooted by the *M. aurantiacus* complex. All nodes except those marked with open circles have 100% bootstrap support. Species divergence times were estimated separately using synonymous divergence relative to synonymous diversity in Pixy. All sampled populations are monophyletic with strong support, with the exception of one sample from B11 that clusters with B19 (marked with an asterisk). (B) Nucleotide heterozygosity (h_s), pairwise diversity (π_s), and divergence (d_s) at genome-wide synonymous sites, indicating substantially higher between-species divergence than within-species diversity. Bars represent species or comparison means, while points represent individual pairwise sample comparisons. Diversity includes both within- and between-population comparison where possible. Sespe creek is substantially diverged from its closest sampled relative, *M. constrictus* (SxC $d_s=6.0\%$). J = *M. johnstonii*, B = *M. brevipes*, F = *M. fremontii*, S = Sespe creek population, C = *M. constrictus*, N = *M. nanus*, AUR = *M. aurantiacus* complex.

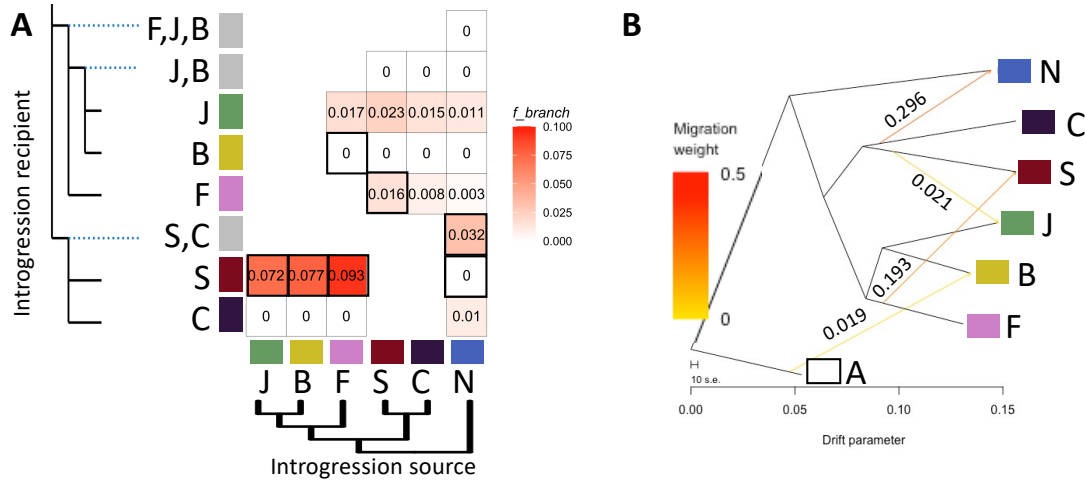


Figure 2.3. Signatures of historical introgression between species in *Mimulus* section *Eunanus*. (A) Genome-wide F_{branch} statistics, an estimate of the proportion of the genome affected by introgression based on multiple f_4 -ratio statistics, calculated from the ‘complete’ SNP dataset. Seven bold-outlined boxes indicate comparisons with significant gene discordance bias supporting introgression from TWISST gene tree summaries ($p_{corr} < 0.01$). Full TWISST summaries are provided in Figure S2.3. (B) TreeMix network analysis results for four migration edges. Migration weights are provided next to each edge. The four-edge model was a significant improvement over the three-edge or smaller models (likelihood ratio test $p = 0.048$, $llik = 197.423$); a five-edge model did not significantly improve the model. The optimizer program OptM selected a single-edge model as the model that maximized the change in log-likelihood; this model had an edge of weight 0.249 from *M. constrictus* into *M. nanus*.

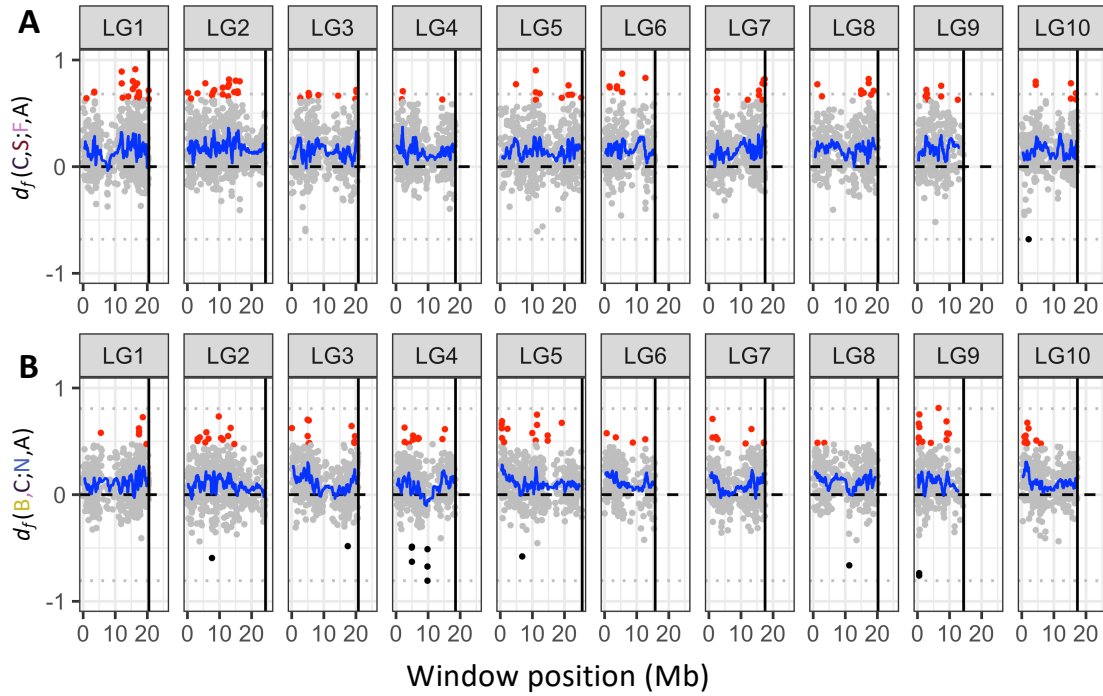


Figure 2.4. Introgession outlier signals are distributed throughout the genome. (A) Position along ten linkage groups of extreme d_f values in 100-SNP windows for the quartet (*M. constrictus*, Sespe Creek; *M. fremontii*, *M. aurantiacus* group). The top 100 windows with the largest absolute value of d_f are highlighted: positive outliers (red) indicate allele sharing consistent with introgression between Sespe Creek and *M. fremontii*, while negative outliers (black) indicate excess allele sharing between *M. constrictus* and *M. fremontii*. The blue line shows a rolling mean of 51 windows, indicating that the pattern of elevated d_f is distributed across the genome rather than focused in a few blocks. (B) Position along ten linkage groups of extreme d_f values in 100-SNP windows for the quartet (*M. brevipes*, *M. constrictus*; *M. nanus*, *M. aurantiacus* group). The top 100 windows with the largest absolute value of d_f are highlighted: positive outliers (red) indicate allele sharing consistent with introgression between *M. constrictus* and *M. nanus*, while negative outliers (black) indicate excess allele sharing between *M. brevipes* and *M. nanus*. The blue line shows a rolling mean of 21 windows. Dotted grey lines indicate the most extreme negative outlier and its equivalent positive value, and dashed black lines indicate $d_f=0$ (no excess allele sharing). Total tested windows after removal of values in regions where $\pi=0$ for any species in the tested trio: (A) 4732, (B) 4141.

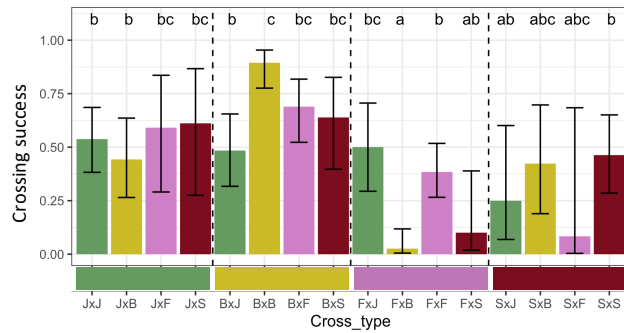


Figure 2.5. Postmating prezygotic isolation between at least one species pair. Crossing success (probability of producing at least one seed from a cross) for intra- and interspecific cross types. Shown are least-square means and asymptotic confidence intervals from a Firth binomial model. Letters indicate significance in post-hoc Tukey tests; cross types sharing a letter are not significantly different. One interspecific cross type, FxB, was significantly reduced compared to both intraspecific parental crosses, indicating an asymmetric barrier to fruit set. One cross, BxJ, was significantly reduced compared to the maternal (BxB) parent but not the paternal (JxJ) parent. Two other crosses, FxS and SxF, had low fruit set possibly indicating a reproductive barrier, but were not significant in the model due to low sample sizes; e.g., 0 out of 4 SxF crosses resulted in seeds. J = *M. johnstonii*, B = *M. brevipes*, F = *M. fremontii*, S = Sespe creek population. Cross types are listed as (maternal x paternal).

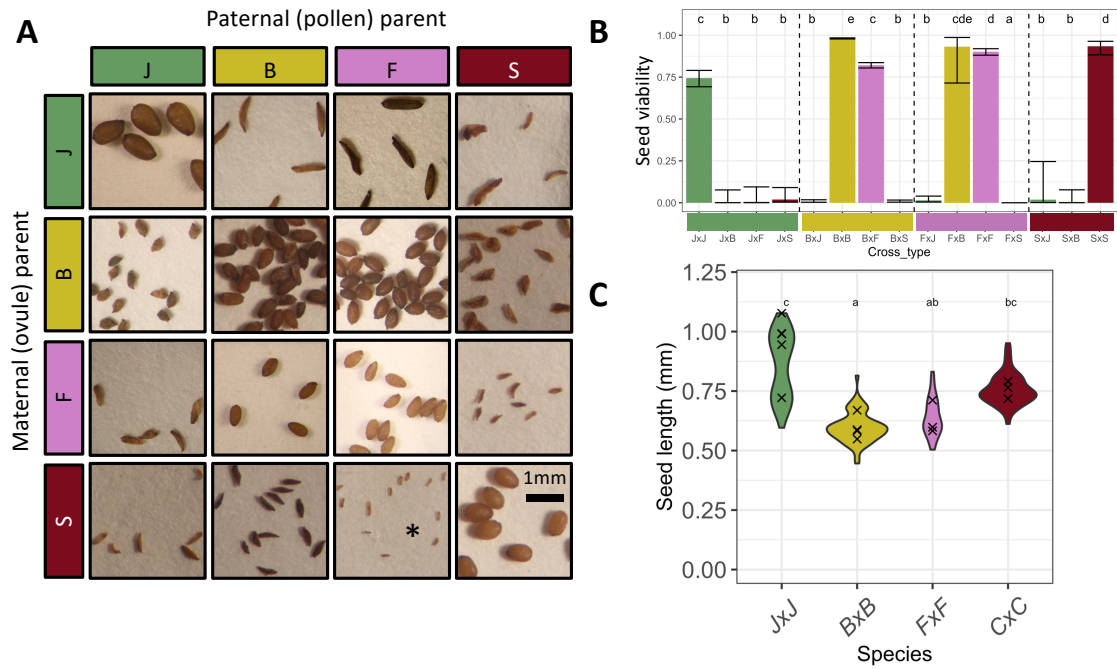


Figure 2.6. Strong hybrid seed inviability between multiple species pairs. (A) Representative seed images from intra- and interspecific laboratory crosses between four species. All images are adjusted to the same scale. For the SxF cross, marked with an asterisk, the only ‘seeds’ produced were not distinguishable from unfertilized ovules and were not counted as seeds for crossing success or viability scoring, but it is possible they represent early seed abortion. (B) Proportion of viable seeds produced by intra- and interspecific laboratory crosses. Shown are least-square means and asymptotic confidence intervals from a Firth binomial model, with letters indicating significance in post-hoc Tukey tests. Intraspecific, BxF, and FxB crosses are mostly viable, but all other interspecific cross types have almost complete inviability. Cross types are listed as (maternal x paternal). (C) Seed lengths for intraspecific crosses. Violin plots show distribution of data for individual seeds, with x’s marking the means of individual fruits. Letters indicate significance from post-hoc Tukey tests of an LMM.

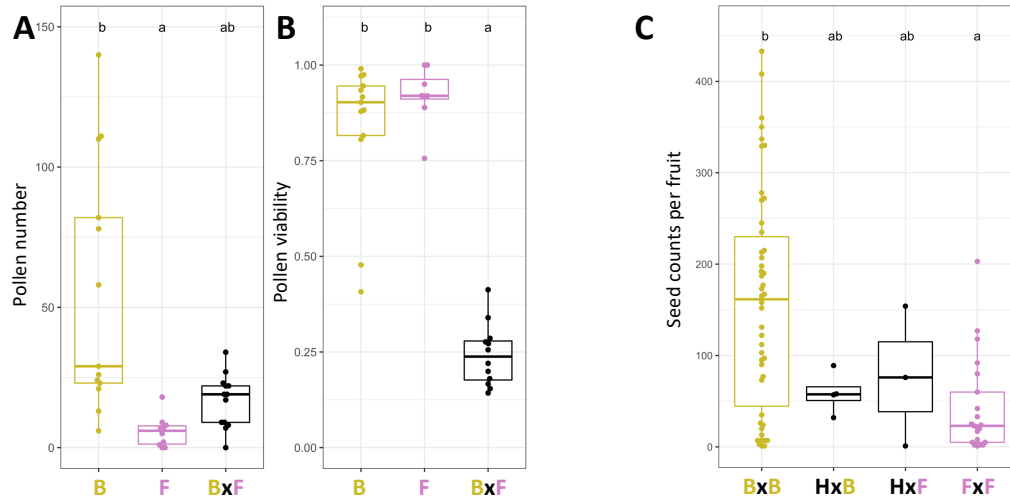


Figure 2.7. Strong hybrid male sterility in hybrids between *M. brevipes* and *M. fremontii*. (A) Pollen number (counts per mm² on a hemocytometer) and (B) pollen viability (proportion of pollen grains scored as viable in aniline blue stain). Letters indicate significance in post-hoc Tukey tests from a poisson (abundance) or binomial (viability) GLMM with population as a random variable. Each point represents the average across 1-3 flowers from a single individual. For each panel, groups sharing a letter are not significantly different. (C) Seed production from hybrid vs. parental fruits as a test of hybrid female sterility, including only fruits that produced at least one seed. Letters indicate significance in post-hoc Tukey tests from a linear model. B=*M. brevipes*, F=*M. fremontii*, H=*M. brevipes* x *M. fremontii* F1 hybrid. Cross types are listed as (maternal x paternal).

CHAPTER III

FLUCTUATING REPRODUCTIVE ISOLATION AND STABLE ANCESTRY STRUCTURE
IN A FINE-SCALED MOSAIC OF HYBRIDIZING *MIMULUS* MONKEYFLOWERS¹

¹Farnitano M.C., K. Karoly, and A.L. Sweigart. Fluctuating reproductive isolation and stable ancestry structure in a fine-scaled mosaic of hybridizing *Mimulus* monkeyflowers.
Submitted to *PLOS Genetics*, 9/18/2024.

Abstract

Hybridization among taxa impacts a variety of evolutionary processes from adaptation to extinction. We seek to understand both patterns of hybridization across taxa and the evolutionary and ecological forces driving those patterns. To this end, we use whole-genome low-coverage sequencing of 459 wild-grown and 1565 offspring individuals to characterize the structure, stability, and mating dynamics of admixed populations of *Mimulus guttatus* and *Mimulus nasutus* across a decade of sampling. In three streams, admixed genomes are common and a *M. nasutus* organellar haplotype is fixed in *M. guttatus*, but new hybridization events are rare. Admixture is strongly unidirectional, but each stream has a unique distribution of ancestry proportions. In one stream, three distinct cohorts of admixed ancestry are spatially structured at ~20-50m resolution and stable across years. Mating system provides almost complete isolation of *M. nasutus* from both *M. guttatus* and admixed cohorts, and is a partial barrier between admixed and *M. guttatus* cohorts. Isolation due to phenology is near-complete between *M. guttatus* and *M. nasutus*. Phenological isolation is a strong barrier in some years between admixed and *M. guttatus* cohorts, but a much weaker barrier in other years, providing a potential bridge for gene flow. These fluctuations are associated with differences in water availability across years, supporting a role for climate in mediating the strength of reproductive isolation. Together, mating system and phenology accurately predict fluctuations in assortative mating across years, which we estimate directly using paired maternal and offspring genotypes. Climate-driven fluctuations in reproductive isolation may promote the longer-term stability of a complex mosaic of hybrid ancestry, preventing either complete isolation or complete collapse of species barriers.

Introduction

Hybridization can have a wide range of impacts on species, depending on both the success of hybrid individuals and their degree of reproductive isolation from progenitors. At one extreme, unfit hybrids can be an evolutionary dead-end (Dobzhansky, 1934; Farnitano & Sweigart, 2023; Richards & Ortiz-Barrientos, 2016; Scopece et al., 2008; Stebbins, 1959; Wodsdalek, 1916). Conversely, hybrids may freely and successfully mate with both progenitors, eroding differences between species until a single undifferentiated population remains (Behm et al., 2010; Kleindorfer et al., 2014; Xiong & Mallet, 2022). Another possibility is that hybrids are successful but become strongly reproductively isolated from progenitors, forming a new, independent lineage (Nice et al., 2013; Rosser et al., 2024; Schumer et al., 2014; Ungerer et al., 1998). Many cases fall between these extremes, with partial but incomplete reproductive barriers (Christie et al., 2022; Coyne & Orr, 1997; Malone & Fontenot, 2008; Moyle et al., 2004).

Partial reproductive isolation allows for ongoing gene flow between species (introgression) without collapse into a single lineage. Plant biologists have long understood the importance of hybridization among members of a ‘syngameon’, a group of species that exchange genes but maintain distinctiveness (Lotsy, 1931; Suarez-Gonzalez et al., 2018). In the genomic era, introgression has been detected throughout the tree of life, suggesting that partial reproductive isolation is common (Goulet et al., 2017; Mallet et al., 2015). Introgression plays an important role in numerous evolutionary processes such as adaptation (Arnold & Kunte, 2017; Chhatre et al., 2018), niche expansion (Manzoor et al., 2020; Pfennig et al., 2016), evolutionary rescue (Baskett & Gomulkiewicz, 2011; Oziolor et al., 2019), extinction (Ayres et al., 2004; Rhymer & Simberloff, 1996), and invasion (Quilodrán et al., 2020; Wolfe et al., 2007). But the

dynamics of introgression are varied. Hybridization can be rare and transient, followed by backcrossing within a few generations to leave a signature of introgression at just a few genetic loci (Goodman et al., 1999; Green et al., 2010; Randi & Lucchini, 2002). In other cases, hybridization produces a persistent swarm of intermating hybrids, which then have the potential to interact more extensively with progenitor species (Harrison & Larson, 2016; Hasselman et al., 2014; Schumer et al., 2018; Yan et al., 2019). The coexistence of a partially isolated hybrid swarm alongside progenitor species provides an opportunity for significant adaptive introgression, since hybrids can act as a bridge for gene flow between progenitors (Larson et al., 2013; Martinsen et al., 2001). However, hybridization can also introduce maladaptive alleles or allele combinations (Dobzhansky, 1934; Fishman & Sweigart, 2018; Muller, 1942; Orr & Turelli, 2001) that an admixed population must contend with (Mantel & Sweigart, 2024; Schumer et al., 2018; Shuker et al., 2005; Zuellig & Sweigart, 2018). Theory suggests that intermediate levels of reproductive isolation may, under certain circumstances, be an evolutionarily stable state rather than simply a transition state on the route to complete speciation (Servedio & Hermisson, 2020). But hybridization outcomes appear to be highly contingent on environmental and genetic conditions (Borge et al., 2005; Brice et al., 2021; Harrison & Larson, 2016; Otis et al., 2017). More empirical work is needed to understand what conditions lead to persistent partial reproductive isolation in hybrid populations.

One factor that could lead to partial reproductive isolation is spatial heterogeneity. When multiple distinct microhabitats are available, divergent directional selection can maintain multiple ecotypes in parallel niches (Butlin et al., 2014; Lowry & Willis, 2010; Melo et al., 2014). Selection on locally favored alleles in these microhabitats can act as a premating reproductive barrier, reducing gene flow between ecotypes by eliminating migrants (Borzée et

al., 2016; Ebersole, 1985; Ling et al., 2022; Mantel & Sweigart, 2019). On the other hand, dispersal between these microhabitats could erode both local adaptation and reproductive isolation between ecotypes, leading to collapse into a single gene pool (Bank et al., 2012; Lenormand, 2002; Xiong & Mallet, 2022). Within this pool, multiple microhabitats may instead promote a diversity of alleles through balancing selection (Delph & Kelly, 2014; Levene, 1953). The relative importance of divergent selection and balancing selection will depend on the scale and frequency of dispersal relative to the scale of environmental heterogeneity (Kisel & Barraclough, 2010; Klein et al., 2017; Musker et al., 2021; Slatkin, 1987). In some cases, dispersal may be sufficient to prevent complete isolation but not strong enough to erode differences completely. In fact, gene flow is often detected in sympatry even when preexisting reproductive barriers are strong (Martin & Willis, 2007; Roda et al., 2017; Sambatti et al., 2012). Furthermore, hybrids may establish their own niche in intermediate or underutilized microhabitats, in which persistence is subject to the same balance of migration and selection (Abbott, 1992; Otis et al., 2017; Schwarzbach et al., 2001; Selz & Seehausen, 2019; Ungerer et al., 1998).

Environmental heterogeneity over time, both within and across years, might also contribute to the maintenance of partial reproductive isolation. Differences in the timing of mating throughout the year can isolate groups (Hendry et al., 1999; Hood et al., 2019; Osmolovsky et al., 2022). If reproductive phenology is dependent on environmental cues, then variation in those cues across years can modulate the strength of temporal isolation (Franks & Weis, 2009; Seehausen et al., 2008; Sianta et al., 2024). Also, variation in selective pressures across years can lead to fluctuating selection that maintains a diversity of genotypes (Abdul-Rahman et al., 2021; Bergland et al., 2014; Han et al., 2020; Troth et al., 2018). Within a

hybridizing population, this can result in selection against hybridization in some years but not in others (Grant & Grant, 1996; Tataru et al., 2023).

Understanding and predicting the effects of environmental variation on hybridization and reproductive isolation requires tracking hybrid populations across time. Multi-year studies have given us important insights into hybridization, demonstrating directional shifts in composition over time (Carney et al., 2000), shifts across space (Taylor et al., 2014), stability over time (Moore & Buchanan, 1985; Sullivan, 1995), and fluctuations in the strength of species barriers (Grant & Grant, 1996; Hendry et al., 2000; Kleindorfer et al., 2014). Still, we know little about the environmental and genetic circumstances driving these different outcomes.

With this in mind, we turn to a previously identified hybridizing population of *Mimulus guttatus* and *Mimulus nasutus* monkeyflowers, at Catherine Creek (CAC) just north of the Columbia River Gorge in Washington, USA (Figure 3.1A). *M. nasutus* diverged from an *M. guttatus* progenitor ~200KYA, expanding to share much of the *M. guttatus* range across the western United States, where the two hybridize in multiple locations of secondary contact (Brandvain et al., 2014). At Catherine Creek and the surrounding area, both species occupy a series of ephemeral seeps, where they co-occur at small spatial scales (Kenney & Sweigart, 2016). Previously, *M. guttatus* collections in 2012 from Catherine Creek showed levels of *M. nasutus* genomic ancestry ranging from near 0 to approximately 50%, indicating a history of hybridization followed by backcrossing to the *M. guttatus* parental population (Brandvain et al., 2014; Kenney & Sweigart, 2016).

Despite this evidence of hybridization, several prezygotic reproductive barriers have been documented between *M. guttatus* and *M. nasutus*. One major source of isolation is mating system: *M. guttatus* is primarily outcrossing (though self-compatible), with large showy flowers

visited by a variety of bee pollinators (Arathi & Kelly, 2004; Kelly & Willis, 2002; Macnair et al., 1989), while *M. nasutus* is predominantly selfing, with small and often cleistogamous flowers that self-pollinate prior to opening (Dole, 1992; Fishman et al., 2002). In addition, the species have phenological differences closely tied to water availability and drought escape: *M. nasutus* tends to be found on mossy rock outcrops that dry out more quickly, while the seepy microhabitats of *M. guttatus* stay wet later into the spring (Kenney & Sweigart, 2016; Kiang & Hamrick, 1978; Mantel & Sweigart, 2019). *M. nasutus* flowers earlier in the season, in part due to a shorter photoperiod requirement for flowering, which has been mapped to multiple large-effect genetic loci (Fishman et al., 2014). In addition, *M. nasutus* is better able to accelerate its life cycle to escape terminal drought (Mantel & Sweigart, 2019). The Catherine Creek area has highly variable precipitation (both in quantity and seasonal timing, Figure 3.1B), so we predict that heterogeneous water availability plays an important role in the persistence and isolation of these species and their hybrids.

Here, we sequence wild-growing individuals and their offspring from three additional years at Catherine Creek and from an additional nearby site, which, combined with previous data, span a decade of collections. This multi-year dataset allows us to ask whether the hybrid population at Catherine Creek is stable across years. By examining fine-scale spatial and temporal patterns of hybrid ancestry, we ask whether environmental heterogeneity might contribute to partial reproductive isolation and hybrid persistence. We address multiple possible contributors to isolation and persistence: spatial segregation of different ancestry cohorts; mating system variation; and phenological isolation across the flowering season. Then, using offspring genotypes, we document how these factors influence the actual mating dynamics of the admixed

population, and how these dynamics change across years in response to fluctuating environmental conditions.

Materials and Methods

Environmental data

We obtained daily precipitation data for the years 2010-2022 from the PRISM online database (PRISM Climate Group, 2024) for the 4km grid square centered on (Latitude=45.7113, Longitude=-121.3637), which includes the Catherine Creek site, and took monthly averages.

Field sampling and collections

We sampled tissue for DNA extraction and sequencing from wild-growing *Mimulus* individuals during the 2019, 2021, and 2022 growing seasons (April through June). In 2019, we sampled from a single stream at Catherine Creek, CAC_S1. In 2021 and 2022, we sampled from both CAC_S1 and an adjacent parallel stream ~120m to the west, CAC_S2 (Figure 3.1A). In 2021, we also sampled from a third more distant stream ~4km away (Little Maui, or LM). In 2021, because we arrived at the field sites in mid-May (due to the ongoing COVID19 pandemic), CAC samples were limited to late-flowering individuals (flowering at LM is shifted to later in the season and thus was unaffected).

Within each stream, we set out 0.5m x 0.5m plots in sites along the stream where *Mimulus* individuals were growing; all samples were collected from within these plots. Plot locations were not always exactly the same across years, but we assigned plots within 10 m of each other to the same plot ID, resulting in a total of 7 plots each for CAC_S1 and CAC_S2, plus 8 plots for LM (Figure 3.1C); some plot IDs were not represented in every year. Approximately once per week throughout the flowering season in both 2019 and 2022, we counted the number

of open *Mimulus* flowers within each plot. The total number of flowers across *M. guttatus*, *M. nasutus*, and potential hybrids was recorded for each plot at each time point; individual flowers were not field-identified to species as hybrids are difficult to distinguish in the field. Flowers are typically open for 1-3 days, and a single individual may have multiple flowers both simultaneously and in sequence (individual flower number ranges from 1 to >50).

During each visit to CAC in 2019 and 2022 and LM in 2021, we used acrylic paint to mark the calyx of three random open flowers per plot (if available). We used a different color of paint on each visit to indicate the date flowers were open. Later in the season, we attempted to relocate these same individuals and, if successful, collected fruits from the marked flowers, as well as leaf tissue into envelopes with silica for DNA. Seeds from these fruits were later germinated in the UGA Botany greenhouses; leaf/bud tissue was collected from the resulting offspring (N = 1-16 per wild-sampled maternal family, from 1-3 individual fruits per family; details in Table S3.1) and stored at -80°C for DNA extraction and sequencing.

DNA Extraction and Illumina sequencing

We extracted genomic DNA from both the dried wild-collected tissue samples and the greenhouse-grown offspring samples using a CTAB extraction protocol with phenol-chloroform extraction (Fishman, 2020). Dried tissue was flash-frozen in liquid nitrogen immediately before grinding; fresh tissue from greenhouse-grown offspring was kept at -80C until flash-freezing and grinding. DNA yield was quantified using a Quant-iT DNA quantification kit (Invitrogen P11496) and plate reader, then normalized to equal concentrations for library preparation.

To prepare libraries for Illumina sequencing, we used a Tagmentation approach (Adey et al., 2010; Picelli et al., 2014); our protocol is available at (Farnitano & Sweigart, 2024). Briefly, Tn5 enzyme was purified in bulk and pre-loaded with universal Illumina adapters following (Lu

et al., 2017). We then added the loaded Tn5 to approximately 1 ng of sample DNA and incubated to fragment DNA and add universal adapters. Next, we added OneTaq Hot Start polymerase (New England Biolabs M0488L) along with combinatorically barcoded forward and reverse primers and ran 18 cycles of PCR to amplify fragments. After PCR, samples were pooled into sets of 48 and cleaned using SPRI magnetic beads. These sets of 48 samples were quantified with a Qubit fluorometer and then merged in equimolar amounts for Illumina sequencing. Samples were sequenced at the Duke University Center for Genomic and Computational Biology using an Illumina NovaSeq 6000 machine to generate paired-end 150bp reads with a targeted depth of ~1X coverage per sample. Reads were demultiplexed based on their combinatorial barcodes into individual samples. Samples we sequenced across three separate runs. Across a total of 3055 samples, we sequenced ~5.5 billion read pairs, for an average of 1.18 million read pairs per sample (standard deviation 2.17 million read pairs). After filtering to remove samples with less than 25,000 called ancestry-informative sites (see *Assigning local ancestry*), our final dataset included 2708 samples with mean sequencing depth of 1.78 million read pairs per sample (standard deviation 1.56 million read pairs, range 45,098 read pairs to 15.72 million read pairs). This final dataset included 459 wild-collected maternal samples and 2248 greenhouse-grown offspring. Details about this final dataset are provided in Table S3.1.

Reference alignment, genotyping, and creation of SNP panels

For each Illumina sample, we used Trimmomatic v0.39 (Bolger et al., 2014) to remove adapter sequences and low-quality ends. We aligned reads with bwa v0.7.17 ‘mem’ (Li & Durbin, 2009) to the *Mimulus guttatus* IM62 v3 reference genome (<https://phytozome-next.jgi.doe.gov>), removed duplicates with picard v2.21.6 ‘MarkDuplicates’ (Broad Institute, 2019), and used samtools v1.10 (Danecek et al., 2021) to keep only properly paired reads with

map quality ≥ 29 . Coverage stats were obtained with qualimap v2.2.1 (Okonechnikov et al., 2016).

We created panels of high-quality, informative SNPs using 36 previously sequenced lines (Table S3.2), including one *M. nasutus* and 8 admixed *M. guttatus* lines from CAC, plus 3 additional *M. nasutus* and 24 allopatric *M. guttatus* from throughout the species' ranges. We followed the same steps above to align these Illumina panel lines to the *Mimulus guttatus* IM62 v3 reference genome. The panel was genotyped using GATK 4.4.0.0 HaplotypeCaller and GenotypeGVCFs in all-sites mode (Van der Auwera & O'Connor, 2020). Called sites were split into biallelic SNPs and invariant sites, with indels and multiallelic sites removed. SNPs were further filtered with GATK to remove sites with $QD < 2$, $QUAL < 40$, $SOR > 3$, $FS > 60$, $MQ < 40$, $MQRankSum < -12.5$, or $ReadPosRankSum > 12.5$ or < -12.5 . Invariant sites were filtered to remove sites with $QD < 2$, $SOR > 3$, or $MQ < 40$. Sites were further filtered at the individual genotype level using vcftools v0.1.16 (Danecek et al., 2011), setting genotypes to missing if $DP < 6$, $DP > 100$, or $GQ < 15$ for that sample. Heterozygous calls were retained. From the resulting genotype file, we created a list of 3,493,514 SNPs called in at least 31 of the 36 reference individuals.

For a more targeted panel, we created a test set of 100 representative wild CAC+LM samples, then subset our variant list to 19,633 sites with nonzero read coverage in at least 60% of test samples and a minor allele frequency of at least 20%. In addition, we created a panel of ancestry-informative sites distinguishing *M. guttatus* and *M. nasutus*, subsetting our full variant list to 208,560 SNPs with $\geq 80\%$ allele frequency difference between the 24 allopatric *M. guttatus* lines and the four high-coverage *M. nasutus* lines.

Population structure analyses with ANGSD

We used ANGSD v0.940 (Korneliussen et al., 2014) with the GATK likelihood model to obtain genotype likelihoods for all wild-collected samples at the targeted variant list of 19,633 SNPs. We then used these genotype likelihoods to run a genomic PCA analysis with PCAngsd (Meisner & Albrechtsen, 2018). The resulting covariance matrix was converted into principal components using the eigen function in R (R Core Team, 2023). As a complementary approach to our ancestry HMM (below), we used NGSadmix (Skotte et al., 2013) with K=2 groups to estimate admixture proportions using the ANGSD genotype likelihoods.

Assigning local ancestry

We used ancestry_HMM (Corbett-Detig & Nielsen, 2017) to assign *M. guttatus* vs. *M. nasutus* ancestry across the genome for each maternal and offspring individual, following the *ancestryinfer* pipeline from (Schumer et al., 2020) and using our 208,560 ancestry-informative sites (above). Reads for each maternal and offspring sample were aligned to both the *Mimulus guttatus* IM62 v3 reference and the *Mimulus nasutus* SF v2 reference (<https://phytozome-next.jgi.doe.gov>), keeping only reads that aligned exactly once to each genome. At each ancestry-informative site, we counted the number of reads supporting each allele. These read counts per allele were used as input for ancestry_HMM (Corbett-Detig & Nielsen, 2017), along with allele frequencies in the 24 allopatric *M. guttatus* and 4 *M. nasutus* reference lines. Recombination rates between each SNP position were approximated using the bp distance between adjacent sites multiplied by the global recombination rate estimate of 3.9e-8 Morgans/bp, calculated using a total genetic map length of 14.7 Morgans (Brandvain et al., 2014) across a genome size of 375 Mb, approximately the mean genome size between *M. guttatus* (~430 Mb) and *M. nasutus* (~320 Mb). Ancestry_HMM was run with the model parameters ‘-a 2 0.5 0.5 -p 0 -100 0.5 -p 1 -100 0.5 -e 0.02’ to estimate posterior probabilities of

M. guttatus, *M. nasutus*, or heterozygous ancestry at each site for each sample. Posterior probabilities of at least 0.9 for any genotype were converted to hard genotype calls, with lower values set to missing. Samples with fewer than 25,000 called genotypes (out of 208,560 sites) were excluded from all analyses, leaving a final dataset of 2685 samples from an initial set of 3055 samples (Table S3.1). Hybrid index (HI) equals the number of sites with homozygous *M. nasutus* calls plus half the number of heterozygous calls, divided by the total number of called sites. Ancestry heterozygosity (AH) equals the number of heterozygous calls divided by the total number of called sites.

To test the reliability of our local ancestry measurements, we subset the raw reads from each allopatric *M. guttatus*, *M. nasutus*, and high-coverage CAC line (36 lines total) to 2 million read pairs per sample, and followed the above local ancestry pipeline to call ancestry across the genomes in these lines. All 4 *M. nasutus* lines had hybrid index >0.99; 18 of 24 allopatric *M. guttatus* lines had HI<0.01 and the remaining six had HI<0.05. The 8 high-coverage CAC *M. guttatus* were more admixed, with hybrid index ranging from 0.11 to 0.34, consistent with previous ancestry estimates for these lines (Mantel & Sweigart, 2024).

Our data include 11 maternal samples from plots S2_1 and S2_2, all from 2022, that we determined to be the allopolyploid species *M. sookensis*. These individuals had hybrid indices near 0.5 and ancestry heterozygosity >0.85, which could indicate polyploidy or new F1 hybridization. Offspring of these individuals had very similar hybrid indices and similarly high ancestry heterozygosity, which is consistent with fixed heterozygosity in these highly selfing polyploids (Modliszewski & Willis, 2012; Sweigart et al., 2008; Whitener et al., 2024) but would not be true for the F2 progeny of F1 hybrids. In addition, these individuals formed a distinct cluster in PCA space: they diverged from the 1-1 correlation between hybrid index and PC1, and

they were separated cleanly by PC5. Finally, offspring grown in the greenhouse had a small flower size and shape that matched *M. sookensis* and *M. nasutus*; F1 hybrids typically have larger flowers more similar to *M. guttatus* (Fishman et al. 2002). We removed all *M. sookensis* individuals from the analyses of reproductive isolation, individual phenology, mating system, and offspring ancestry deviations.

We combined our data with previously obtained hybrid indices from 75 individuals collected in 2012 from CAC_S1 and CAC_S2 (Kenney & Sweigart, 2016). These samples were genotyped using MSG sequencing (Andolfatto et al., 2011), and hybrid indices were calculated from an HMM approach implemented in HapMix (Price et al., 2009); note that these are different data types and were processed differently than our 2019-2022 data. However, as both approaches use low-coverage whole-genome sequencing and HMM ancestry calling, we expect them to produce broadly comparable hybrid indices even if there are slight differences at individual loci.

Organellar haplotype networks

To investigate the haplotype structure of maternally transmitted organellar genomes in our population, we aligned each of our maternal samples, along with 8 high-coverage CAC lines (Table S3.2), to the *M. guttatus* IM767 v1.1 chloroplast assembly (<https://phytozome-next.jgi.doe.gov>) using BWA v0.7.17 (Li & Durbin, 2009), and called genotypes using GATK v4.4.0.0 HaplotypeCaller and GenotypeGVCFs functions (Van der Auwera & O'Connor, 2020). Read coverage aligned to the chloroplast genome was high compared to the nuclear genome, so we restricted each analysis to samples with mean coverage ≥ 40 , resulting in 328 of our samples and 6 of the previously sequenced CAC lines. We then subset each VCF to a set of SNPs from a recent analysis of organellar haplotypes across *M. guttatus*, *M. nasutus*, *M. decorus*, and *M.*

sookensis (Whitener et al., 2024), further removing any SNPs with missing data, heterozygous calls, or individual depth <4 for any sample using BCFtools v1.15.1 (Danecek et al., 2021). We merged this dataset with the genotype calls for the 41 samples from (Whitener et al., 2024), output SNP calls as a fasta sequence for each sample using BCFtools consensus, and converted to nexus format using EMBOSS v6.6.0 (Rice et al., 2000). We then used the Integer NJ Net function in popart (Leigh & Bryant, 2015) to generate a haplotype network from the resulting alignment. Our final chloroplast dataset included 102 segregating sites, of which 43 were parsimony-informative. We followed this same approach to generate a mitochondrial genome haplotype network, aligning sequences to an *M. guttatus* IM62 mitochondrial assembly (Mower et al., 2012) and excluding regions annotated as chloroplast-derived. After filtering for mean coverage ≥ 40 , we retained 150 of our samples and 8 previously sequenced CAC lines, which were merged with the 41 samples used in the recent analysis. Our final mitochondrial dataset included 120 segregating sites, of which 39 were parsimony-informative.

Estimating selfing rates

To estimate selfing and outcrossing rates, we first excluded maternal-offspring pairs whose genotypes were incompatible with a maternal-offspring relationship. For each maternal-offspring pair, we used ancestry calls from ancestry_HMM (at 208,560 sites) to calculate the number of informative sites where genotypes were incompatible with maternal-offspring relationships (i.e. maternal and offspring samples were homozygous for opposite ancestries), divided by the total number of homozygous maternal calls with a called genotype in the offspring. Allowing for some uncertainty in ancestry calling, we excluded any offspring for which this incompatibility ratio was $\geq 5\%$, as well as entire families if greater than half the

offspring were excluded. After these filters, our final family dataset contained 185 maternal plants with offspring, plus 1565 offspring.

We then randomly thinned our list of ancestry-informative sites to one per 10kb, leaving 17,514 sites, and calculated genotype likelihoods at each site for both maternal and offspring samples using ANGSD v0.940 with the GATK likelihood model (Korneliussen et al., 2014). We estimated probabilities of selfing vs. outcrossing for each maternal-offspring pair, and probabilities of half-sibling vs. full-sibling relationships for each outcrossed sibling pair, using the Bayesian model implemented in BORICE.genomic v3 (Colicchio et al., 2020). We tuned allele frequencies for 20 steps, then had 20 steps of burn-in and a total chain length of 100, with records kept every 2 steps. Offspring with a posterior probability ≥ 0.9 were called as selfed or outcrossed, and other offspring were excluded. We counted pairs as half-siblings if the posterior probability was $>50\%$, full-siblings if it was $<50\%$, and missing data if the posterior probability was exactly 50% .

Calculating the strength of reproductive isolation

We calculated standardized estimates (Sobel & Chen, 2014) of reproductive isolation (RI) between three ancestry cohorts within CAC_S1, defined by hybrid index: an *M. guttatus* (lower admixture levels <0.15) cohort, an admixed (higher admixture levels >0.15), and an *M. nasutus* cohort. We calculated two components of premating reproductive isolation, mating system isolation and phenological isolation, as well as their combined effects. Each measure ranges from -1 (complete disassortative mating) to 1 (complete assortative mating), with 0 indicating random mating.

For mating system isolation, we used selfing rates estimated from BORICE, dividing individuals into three cohorts by hybrid index. We assumed for this calculation that outcrossing

events contribute no isolation to the maternal cohort ($RI=0$), while selfing events contribute perfect isolation to their maternal cohort ($RI=1$). The degree of mating system isolation is then equal to the rate of selfing of maternal plants in each cohort. While selfing flowers may also contribute less to the pollen pool, causing increased mating system isolation from the paternal direction, this indirect effect is not included in our calculation.

To calculate phenological isolation, we chose one plot from within CAC_S1 to represent each ancestry cohort: plot S1_4 for *M. guttatus*, plot S1_2 for the admixed cohort, and plot S1_5A for *M. nasutus*. For each pair of plots, we calculated the probability of within-plot vs. between-plot matings based on the number of open flowers in each plot on each day: for each open flower, between-plot mating probability is equal to $B/(A+B)$ and within-plot mating probability is equal to $A/(A+B)$, where A and B are the number of open flowers on the same day as the focal flower from the same and opposing plots, respectively. This assumes that each open flower contributes equally to the pollen pool and ignores the effects of physical distance. These probabilities were then summed across all open flowers across the season to obtain an estimated total number of between-plot and within-plot pairings. Reproductive isolation is then equal to $RI=1-2P$, where P is the relative probability of these between-plot vs. within-plot pairings.

Following (Sobel & Chen, 2014), we calculated the combined effect of mating system (RI_M) and phenological (RI_P) isolation using the following equation:

$$RI_{Combined} = RI_M + ((1 - RI_M) * RI_P)$$

Note that the order of isolating barriers does not matter for this calculation.

Using offspring to assess assortative mating

If assortative mating by ancestry is strong, we expect offspring hybrid index to closely match maternal hybrid index; in contrast, deviations between offspring and maternal hybrid

index may indicate mating across different ancestry cohorts. To test for deviations from maternal hybrid index within the offspring, we first separated each fruit into selfed and outcrossed offspring based on BORICE. We then calculated the mean for each fruit of the difference between Offspring HI and Maternal HI (Offspring HI deviation), averaging selfed and outcrossed offspring separately. We expect a deviation near 0 for selfed offspring, or under strong assortative mating by ancestry; positive values indicate that the offspring have more *M. nasutus* ancestry than their maternal parent, while negative values indicate they have more *M. guttatus* ancestry than their maternal parent.

Statistical models

To test for an effect of maternal ancestry and year on selfing rates, we ran binomial GLMs using the R package stats (R Core Team, 2023). We ran these models for all families, then for all families except *M. nasutus* (HI>0.95) and for only families with maternal HI<0.5.

To test whether phenology was associated with hybrid index at the plot level, we paired flower censuses throughout the growing season from 2012, 2019, and 2022 with hybrid indices from our sequenced maternal plants from those plots. For each plot in each year, we calculated the date that the median censused flower was open in that plot. We then took the median hybrid index of all sequenced individuals from that plot. We ran a beta regression model using the R package ‘betareg’ (Cribari-Neto & Zeileis, 2010) to test for an effect of median flower date and year on median hybrid index, with year as a factor variable.

We also tested whether flower date was associated with hybrid index at the individual level. For sequenced individuals in 2019 and 2022, we have dates that marked flowers were open – note that these are not necessarily the first flowers on a given individual but were instead randomly chosen flowers from the plot on the day of marking. We ran a beta regression model

using the R package ‘betareg’ (Cribari-Neto & Zeileis, 2010) to test for an effect of individual marked flower date and year on maternal hybrid index, with year as a factor variable.

We ran linear models using the `lm` function in the R package `stats` (R Core Team, 2023) to test for an effect of Maternal HI and Year on Offspring HI Deviation (the difference between offspring hybrid index and maternal hybrid index, averaged per fruit). We ran separate linear models for the selfed offspring and for the outcrossed offspring, as well as a combined model with ‘Mating History’ (selfed or outcrossed) as an additional predictor. We also ran an ANOVA on each linear model using the ‘`anova`’ function in the R package ‘`stats`’ (R Core Team, 2023) to assess the relative effects of each predictor and interaction variable.

For all statistical models, we added each predictor sequentially, allowing for interactions, and used likelihood ratio tests to determine whether the additional predictor significantly improved the model fit.

Results

*Introgression from *M. nasutus* in replicated contact zones*

To explore patterns of genomic variation within and between sympatric populations, we collected and sequenced 459 wild *Mimulus* samples from three streams at the Catherine Creek and Little Maui field sites (Figure 3.1A,C) across the 2019, 2021, and 2022 growing seasons, and performed a PCA on these samples (Figure 3.2A). PC axis 1 (17.91% of variation) separates species by ancestry, showing a clearly differentiated *M. nasutus* group and a cloud of *M. guttatus*-like individuals with varying levels of hybrid ancestry; PC1 is highly correlated with hybrid index as determined by local ancestry inference ($r^2=0.978$, Figure 3.2B). A separate NGSadmixture analysis with $K=2$ is also highly correlated with hybrid index ($r^2=0.979$,

Figure S3.1). Additional PC axes reveal population structure associated with geography: PC2 (1.96% of variation) cleanly separates samples collected 4 km apart at the LM and CAC sites (including admixed individuals at PC1 \sim 0 and *M. nasutus* at PC1 \sim 0.1), while PC3 (1.74%) and PC4 (1.31%) identify three distinct clusters of admixed individuals within Catherine Creek at a scale of < 300 m (Figure 3.2A). These three CAC clusters appear to reflect microspatial structure at this site, with two of the three clusters containing only samples from one stream (CAC_S1) and the third containing samples from both CAC streams. Within CAC_S1, each study plot is largely confined to one of these three clusters (Table S3.3), suggesting genetic differentiation between samples separated by as few as 50 m. Meanwhile, samples from the second CAC stream (CAC_S2) are not differentiated from CAC_S1 samples within their shared cluster.

Consistent with previous population genomic analyses (Brandvain et al., 2014; Kenney & Sweigart, 2016), the distribution of hybrid index (HI, the proportion of the genome with *M. nasutus* ancestry) suggests mostly unidirectional backcrossing of hybrids with *M. guttatus* (i.e., a majority of admixed individuals have HI<0.5, Figure 3.2C). In fact, virtually all CAC and LM *M. guttatus*-like individuals have at least some detectable *M. nasutus* ancestry: out of 374 majority-*guttatus* samples, only 12 have HI<0.05, while only two of those have HI<0.01. The *M. nasutus* samples, in contrast, have little to no introgression: out of 74 majority-*nasutus* samples, 61 have HI>0.95 and 57 of those have HI>0.99. The remaining 13 have HI between 0.5 and 0.8, indicating that backcrossing with *M. nasutus* does occasionally happen. No individuals were sampled with HI between 0.8 and 0.95, suggesting that backcrossing to *M. nasutus* does not typically continue for multiple generations as it does with *M. guttatus*. Despite pervasive admixture, we detect no first-generation (F1) hybrids (Figure S3.2). Although we did discover a group of 11 individuals from CAC_S2 with HIs near 0.5, they were determined to be the

allopolyploid species *M. sookensis* (Figures 3.2B, S3.2, S3.3, Materials and Methods) and were removed from further analyses of admixture and reproductive isolation. Of the remaining 27 individuals with HIs between 0.4 and 0.6, all have ancestry heterozygosity values less than 0.62 (Figure S3.2), indicating they are second- or later-generation hybrids as opposed to new F1s (which should have heterozygosity near 1.0). Taken together, our finding that directional introgression composed of later-generation and backcrossed hybrids is replicated in all three sampled streams indicates admixture in secondary contact is not only common but follows consistent patterns across the landscape.

Given the admixture we detect in the nuclear genome, we asked whether organellar genomes also show signatures of introgression. Remarkably, we found that all 325 CAC and LM samples - *M. guttatus*, admixed, and *M. nasutus* - have a single chloroplast haplotype (Figure 3.3). A similar pattern is seen in the mitochondria (Figure S3.4). This same haplotype (or close derivatives) is carried by *all* sampled *M. nasutus* and *M. sookensis* lineages, which include collections spanning most of the two species' ranges (Whitener et al., 2024). In contrast, the haplotype is almost never found in *M. guttatus*, which has much higher organellar diversity across its geographic range (Figures 3.3, S3.4): of 18 previously sampled *M. guttatus* accessions, only two carry this CAC/LM haplotype. Surprisingly, one of these two accessions is the only other *M. guttatus* sample in our dataset from a known sympatric site: DPR, which is ~1000 km south of CAC/LM in the Sierras of California (Brandvain et al., 2014; Zuellig & Sweigart, 2018). We therefore infer that this haplotype is derived from *M. nasutus* and has been completely captured by sympatric *M. guttatus* through introgression at both CAC and LM sites (and, potentially, at other sympatric populations elsewhere in the range). This haplotype structure

shows that not a single *M. guttatus* individual in our sample is free from the effects of introgression.

Patterns of ancestry at small spatial scales are persistent across years.

While we find substantial admixture in all three streams, ancestry proportions vary across the landscape (Figure 3.2C). In each stream, we see multiple distinct, though sometimes overlapping, peaks of ancestry, which we refer to as ancestry cohorts. All three streams have a clear *M. nasutus* cohort (HI>0.95). LM and CAC_S2 each have a second, majority-*M. guttatus* cohort, but with different peak ancestries (HI=0.1-0.2 for LM and 0.25-0.3 for CAC_S2). Within CAC_S1, we see three cohorts: *M. nasutus*, a *M. guttatus*-like cohort centered around HI=0.05-0.1, and a more admixed cohort centered around HI=0.25-0.3. Across multiple years of sampling, the distribution of ancestry in each stream appears similar across years, including the presence of all three cohorts (*M. guttatus*, admixed, and *M. nasutus*) in CAC_S1 and two cohorts (admixed, *M. nasutus*) in CAC_S2. These similarities, despite minor differences in sampling from year to year, suggest that the ancestry structure of these populations is stable across time, even as it varies among streams.

We asked whether the multiple peaks of ancestry within CAC_S1 correspond to spatial structure at a smaller scale. The distribution of hybrid ancestry varies across plots, with clear spatial segregation between the three ancestry cohorts (Figure 3.4A). These differences are apparent on very fine scales: plots ~20m apart have distinct distributions of ancestry (i.e., plots S1_3, S1_4, and S1_5A; Figures 3.1C, 3.4A). This spatial pattern of ancestry is remarkably consistent across years wherever we have multiple years of sampling. This is true even comparing 2012 (which used a different sequencing methodology) to later years, with a few exceptions: plot S1_5B, for example, appears to have become more admixed between 2012 and

2019. But overall, differences in admixture proportion at small spatial scales are remarkably consistent across a decade of sampling.

Partial reproductive isolation between admixture cohorts due to mating system

We next sought to investigate potential reasons for the persistence of spatial ancestry structure at such fine scales. One possible cause is premating isolation driven by self-fertilization, so we used maternal and offspring genotype data from 150 maternal families within the two CAC streams to infer selfing rates for our samples. We found that, as expected, *M. nasutus* samples are highly selfing: only 1 of 45 *M. nasutus* offspring was inferred to be outcrossed. *M. guttatus* and admixed cohorts had a mix of selfing and outcrossing, with similar selfing rates between *M. guttatus* (HI<0.15, selfing rates=0.071-0.321, Table 3.1) and admixed (HI=0.15-0.8, selfing rates=0.218-0.356, Table 3.1) maternal plants. Among admixed individuals, selfing rate significantly increased with increasing *M. nasutus* ancestry, though this pattern is driven by a small number of individuals with hybrid index >0.5 (p=0.0004 with and p=0.9058 without these individuals, Table S3.4, Figure S3.6). Selfing rate also significantly varied across years, with 2022 having the highest selfing rates across ancestry cohorts (Table 3.1, Table S3.4, Figure S3.6). Within the *M. guttatus* and admixed cohorts, selfed offspring were well-distributed across fruits: out of 170 fruits with at least 2 offspring, 51% (86) had a mix of selfed and outcrossed offspring (Table S3.1). Within mixed fruits, most were majority outcrossing with a mean selfing proportion of 35% (Table S3.5). Multiple paternity within a fruit was also common: out of 115 fruits with at least two outcrossed offspring, 100 (87%) had more than one inferred pollen donor, with an average probability of 77.6% that two outcrossed offspring in the same fruit were half-siblings rather than full-siblings (Table S3.6).

We used selfing rates to estimate the strength of reproductive isolation due to mating system. For *M. nasutus*, mating system is an almost complete barrier to reproduction (RI=0.96-1.0, Table 3.2). Between the *M. guttatus* and admixed cohorts, mating system was a partial, but important, barrier (RI=0.209-0.405, depending on the year, Table 3.2). Note that these estimates do not consider potential secondary effects of selfing on reproductive isolation, such as lower contributions to the outcrossing pollen pool by self-fertilizing flowers.

Partial phenological isolation between ancestry cohorts

Another potential source of isolation between ancestry cohorts is flowering phenology, so we conducted a multi-year census of flowering phenology at CAC to address phenological isolation. For a given plot, peak flowering times were typically consistent across years (Figure 3.4B), despite differences in overall flower abundance (Figure 3.4C). As previously shown for the 2012 growing season (Kenney & Sweigart, 2016), we found that plot phenology at CAC consistently tracks hybrid ancestry (Figures 3.4A-B): median flower date of plots is highly correlated with their median hybrid index (pseudo- $r^2 = 0.607$, Figure 3.4D and Table S3.4). The flowering time of individual plants is also associated with ancestry (pseudo- $r^2 = 0.326$, Table S3.4): *M. nasutus* flowers were typically marked early in the season, admixed individuals mid-season, and late-season *M. guttatus* individuals later in the spring (Figure 3.4E).

Focusing on three representative plots in CAC_S1 with consistent sampling across years, we calculated the strength of phenological isolation between *M. nasutus* (plot S1_5A), admixed (plot S1_2), and *M. guttatus* (plot S1_4) cohorts. Isolation between *M. nasutus* and *M. guttatus* was complete or nearly so in all years (RI=0.98 to 1.0, Table 3.2). This finding, along with the absence of any first-generation hybrids among 459 wild-collected CAC and LM samples (Figure S3.3A), suggests new interspecific crosses are rare. In contrast, phenological isolation between

admixed individuals and *M. guttatus* was variable across the three study years: although the two groups were strongly isolated in 2012 (RI=0.88 and 0.99 depending on direction, Table 3.2) and in 2019 (RI=0.98 and 0.97, Table 3.2), they were much less so in 2022 (RI=0.44 and 0.78, Table 3.2).

What might explain the reduction in phenological isolation in 2022? Although median flowering time of each cohort was relatively stable across years, the duration of flowering was not: admixed plots in 2022 flowered later into the season than in 2019 or 2012 (Figure 3.4B,D). As a result, there was greater phenological overlap between cohorts in 2022 compared to 2019 or 2012 (e.g., plots S1_1 to S1_3 vs. plot S1_4; Figure 3.4B,D). We see this pattern in the individual-level data as well: in 2019, all marked flowers after May 15 were from the *M. guttatus* cohort, but in 2022, multiple flowers marked after May 15 were from substantially admixed individuals. Differences in absolute abundance likely play an important role as well: 2019 had much lower flower counts throughout the season in all plots compared to both 2012 and 2022 (Figure 3.4C). We also counted relatively more flowers in admixed than in *M. guttatus* plots in both 2012 and 2022 (Figure 3.4C), explaining the asymmetry of reproductive isolation between the *M. guttatus* and admixed cohorts in those years (i.e., higher probability of pollen flow from admixed to *M. guttatus* plots: Table 3.2).

The strength of phenological isolation between admixed individuals and *M. nasutus* also varied across years, depending on the direction of pollen flow. Admixed individuals were less likely to receive pollen from *M. nasutus* in 2022 (RI = 0.96) than in 2012 (RI = 0.83) or 2019 (RI=0.65), in part due to a second wave of late-flowering hybrids in 2022 (Figure 3.4B) that did not overlap with *M. nasutus*. In all three years, *M. nasutus* was poorly isolated from admixed individuals (RI=-0.51 to 0.48: Table 3.1), but these estimates are incomplete in 2012 and 2022

because our census of open flowers began after *M. nasutus* had already begun flowering (Figure 3.4B).

Measured reproductive barriers accurately predict offspring ancestry shifts

Our paired sets of maternal-offspring genotypes allow us to directly test whether the strong premating barriers we detect at CAC actually translate to observed offspring identities. Any deviation in offspring hybrid indices from maternal values implies incomplete assortative mating (postmating barriers could shift allele frequencies at particular loci but are unlikely to cause systematic shifts in offspring HI). For *M. nasutus*, we observed nearly complete assortative mating due to self-fertilization: at CAC only one of 45 offspring across 11 fruits was inferred as outcrossed by BORICE (Table 3.1), and even this was an intra-*M. nasutus* outcross. We do find one example at LM where 2 of 7 offspring in a single *M. nasutus* fruit have admixed ancestry (HI=0.687 and 0.627), implying a partially-admixed *M. guttatus*-like pollen parent. In the reciprocal direction, one offspring from a 2022 CAC *M. guttatus* fruit (maternal HI=0.089) and one offspring from a 2021 LM admixed fruit (maternal HI=0.316) had hybrid indices consistent with a *M. nasutus* pollen parent. We can therefore estimate that the rate of *M. nasutus* mating outside its own ancestry cohort is 2/67 or 3.0% maternally, and 2/1501=0.1% paternally.

Next, we asked to what extent mating system differences and phenological isolation shape patterns of assortative mating between hybrids and *M. guttatus* at CAC. We reasoned that assortative mating should be maximized in selfed offspring and, indeed, we observe only a slight deviation in the HIs of offspring inferred as selfed relative to maternal values (regression slope = -0.113 ± 0.031 , $p=0.005$, Figure 3.5A and Table S3.4). We attribute this slight relationship to uncertainty in either hybrid index estimation or the inference of selfing vs. outcrossing. For outcrossed offspring (which, by definition, have escaped the effects of mating system isolation),

we might expect to observe a stronger deviation between maternal and offspring HIs if assortative mating due to phenological isolation is incomplete. This is precisely what we observe: hybrid indices of outcrossed offspring show a clear deviation from maternal values (regression slope = -0.307 ± 0.024 , $p < 0.0001$, Figure 3.5B and Table S3.4) with *M. guttatus* parents on average producing more-admixed offspring and admixed parents producing less-admixed offspring.

Strikingly, we also find evidence that yearly variation in assortative mating mirrors patterns of phenological isolation at CAC. In 2019, when phenological isolation between *M. guttatus* and hybrids was nearly complete (RI > 0.98 in both directions), *M. guttatus* maternal plants almost never produced offspring with shifts in ancestry. In contrast, in 2022, when phenological isolation between these groups was at its lowest (RI: 0.34 and 0.79), both *M. guttatus* and admixed parents produced offspring with shifted HIs (Figure 3.5). In 2019 and 2021, most large shifts in hybrid ancestry occurred in the few admixed individuals with mostly *M. nasutus* ancestry (i.e., HI < 0.5), which are further from the population mean; they tended to produce offspring with much more *M. guttatus* ancestry, presumably by crossing with more abundant lower-HI hybrids (Figure 3.5B). These high-HI individuals were also more likely to self in 2022 (Table 3.1). The net result of these changes is that, in 2022, outcrossed offspring overall experienced a slight shift to higher levels of admixture compared to maternal samples (mean shift in HI = $+0.0172 \pm 0.0002$), whereas in 2019 and 2021, there was a slight shift to lower levels of admixture (mean shift in HI = -0.0333 ± 0.0002 and -0.0155 ± 0.0004).

Discussion

Here, we describe in detail the composition and mating dynamics within hybridizing populations of *Mimulus guttatus* and *Mimulus nasutus* during secondary contact. Admixture is prevalent in independent streams, with similar patterns of multi-generational, directional introgression from *M. nasutus* into *M. guttatus*. The distribution of hybrid ancestry across space is variable both at very fine (~50m) and coarser (~4km) scales, but stable across a decade of sampling. We measure partial reproductive isolation between three ancestry cohorts by mating system and phenology, both of which likely contribute to the maintenance of ancestry structure across space. We demonstrate substantial year-to-year variation in phenological isolation, which is likely associated with climatic variation. Using direct measurements of offspring ancestry composition, we confirm that variation in measured reproductive isolation predicts observed assortative mating. This is a rare direct confirmation of the effect of premating reproductive barriers on offspring outcomes. Fluctuations in reproductive isolation likely help maintain a mosaic of ancestry across the landscape, preventing either collapse into a single lineage or complete independence of cohorts. Our system demonstrates that the outcomes of hybridization can be dynamic and complex, particularly in scenarios of fluctuating, partial reproductive isolation.

Repeated cases of introgression across the landscape

While introgression between *M. nasutus* and *M. guttatus* has been detected previously at Catherine Creek and in a separate sympatric area in California (Brandvain et al., 2014; Kenney & Sweigart, 2016; Mantel & Sweigart, 2024; Sweigart & Willis, 2003; Zuellig & Sweigart, 2018), strong premating reproductive barriers between the two species imply that initial hybridization rates should be rare (Martin & Willis, 2007). Our data suggest that ongoing

admixture is common across the landscape, with independent admixed populations at Little Maui and Catherine Creek, despite no new hybridization events detected. This follows a trend in other systems where genomic signatures of introgression often co-occur with strong reproductive isolation (Machado et al., 2007; Sambatti et al., 2012), and a broader pattern of frequent signals of hybridization despite generally strong premating reproductive isolation across the tree of life (Christie et al., 2022; Goulet et al., 2017; Mallet et al., 2015). Theory shows that even occasional migration events can strongly influence allele frequencies (Clarke et al., 1997; Sambatti et al., 2012), and rare hybridization events can result in admixed populations if hybrids are persistent once formed. Our finding of weaker reproductive isolation between admixed groups compared to non-admixed progenitors means that, once a few admixed individuals are present, they may promote additional admixture, acting as a genetic ‘bridge’ between otherwise isolated populations and increasing the chance of adaptive introgression (Bettles et al., 2005; Gilman & Behm, 2011; Larson et al., 2013; Martinsen et al., 2001). Admixed groups might also function as a genetic ‘sieve’: multiple generations of selection can purge incompatible allele combinations, while recombination breaks up linkage between incompatibilities and potentially adaptive alleles (Simon et al., 2021; Xiong & Mallet, 2022; Zuellig & Sweigart, 2018).

An open question is how repeatable introgression patterns will be when hybridization occurs multiple times (Brice et al., 2021; Harrison & Larson, 2016; Langdon et al., 2024; Riquet et al., 2019; Simon et al., 2021). The overall pattern of directional introgression, with pervasive *M. nasutus* ancestry in majority-*M. guttatus* genomes but very little signature of introgression into *M. nasutus*, is consistent across streams. The high selfing rate of *M. nasutus* explains this directionality, which is a common pattern in selfer-outcrosser pairs (Nelson et al., 2021; Rifkin et al., 2019; Ruhsam et al., 2011; Sianta et al., 2024; Sweigart & Willis, 2003). But despite similar

asymmetries, each of our three streams have their own unique distribution of hybrid ancestry. This matches findings in other hybrid zones with a mosaic structure (Lepais et al., 2009; Rand & Harrison, 1989; Riquet et al., 2019; Schumer et al., 2018; Simon et al., 2021). Differences in the timing and extent of water availability (Sianta et al., 2024), the spatial distribution of microsites (Rand & Harrison, 1989), the relative abundance of each species (Lepais et al., 2009), or the presence of pollinators (Aldridge, 2005) could all influence these idiosyncratic patterns. In the future, expanding this work to additional streams and measuring ecological variables at higher resolution will help us understand which of these or other factors are most important. Overall, we see that a combination of consistent (i.e., selfing) and heterogeneous (i.e., flowering time and microhabitat) reproductive barriers can produce a patchwork of broadly similar but subtly different outcomes each time admixture occurs.

Complete organellar capture by hybridization

Hybridization often moves organellar genome haplotypes from one lineage into another, a phenomenon known as organellar capture (Comes & Abbott, 2001; Folk et al., 2017; Tsitrone et al., 2003). Typically, organellar capture is assessed in just a small number of samples at phylogenetic resolution. Our population-scale sample provides a unique window into hybridization history – mainly, that an entire sympatric area has a single maternal origin. An *M. nasutus* maternal origin is consistent with the asymmetric nature of gene flow in our system and the observation, by us and others (Martin & Willis, 2007) that *M. nasutus* is more often the maternal parent when hybridizing. It also corroborates our finding that all our *M. guttatus* samples have at least small amounts of nuclear *M. nasutus* ancestry. Still, the extent of capture in even our most *M. guttatus*-like samples is striking and gives us insight into the formation of our hybridizing populations. Not only was the initial hybridization directional, but hybrids must have

consistently remained as the seed parent across multiple generations of backcrossing, with nearby *M. guttatus* progenitors contributing primarily through pollen flow rather than seed dispersal. This pattern is consistent with pollen flow acting over longer distances than seed dispersal, so that most new seeds are from maternal plants from within a population, but pollen occasionally arrives from elsewhere. We might expect a similar pattern in other plant systems when pollen flow happens over longer distances than seed dispersal (Webb, 1998), or in animal systems for which males tend to disperse longer distances than females (Stephen Dobson, 1982).

It is possible that the *M. nasutus* organellar haplotype has some selective advantage within hybrid populations, perhaps due to segregating cytonuclear incompatibilities. Such interactions are common across eukaryotes (Barreto et al., 2018; Bogdanova et al., 2009; Gobron et al., 2013; Lee et al., 2008; Meiklejohn et al., 2013), including between populations of *M. guttatus* and *M. nasutus* (Fishman & Willis, 2006). But we stress that asymmetries in reproductive isolation and dispersal are sufficient to explain these patterns without needing to invoke selection.

Persistent spatial structure with fluctuating reproductive isolation

Repeated sampling across years allows us to see that each stream has a stable distribution of hybrid ancestry, suggesting a lack of severe hybrid breakdown or maladaptation, although the existence of weak postzygotic barriers (Mantel & Sweigart, 2024) helps explain genomic signatures of selection against *M. nasutus* ancestry at Catherine Creek (Kenney & Sweigart, 2016; Mantel & Sweigart, 2024). Ancestry levels are stable even at ~20-50m scales within a stream, much smaller than the scale of expected pollinator movement, suggesting that forces other than distance are maintaining isolation between cohorts. Mosaic hybrid zones with stark fine-scaled structure have been described in other systems (Ross & Harrison, 2002; Valbuena-

Carabaña et al., 2007) and may be common in plants (Abbott, 2017) but there are few studies of change over time in such systems. More broadly, hybrid zone studies have sometimes found stability across years (Moore & Buchanan, 1985; Sullivan, 1995), and other times found substantial shifts (Carney et al., 2000; Kleindorfer et al., 2014; Taylor et al., 2014), but the reasons for these patterns are underexplored. We therefore set out to understand what forces might lead to a fine-scaled stable distribution of admixture, and how they might be influenced by a fluctuating environment.

As expected, selfing is an important barrier isolating *M. nasutus* from both *M. guttatus* and admixed cohorts. Interestingly, it also provides a partial barrier between *M. guttatus* and admixed cohorts. For *M. guttatus*, our selfing rates agree with those of other studies, which range from about 25-50% (Colicchio et al., 2020; Dudash & Ritland, 1991; Leclerc-Potvin & Ritland, 1994; Ritland, 1989; Ritland & Ganders, 1987). Selfing rates in admixed *Mimulus* have not previously been documented; we find that they are generally closer to *M. guttatus* than *M. nasutus*, consistent with dominance of *M. guttatus* floral phenotypes (Fishman et al., 2002). We also see that selfing rates fluctuate slightly across years, possibly due to changes in pollinator abundance and timing (Karron et al., 2009), or in the size and number of flowers as a consequence of general plant health (Karron et al., 2004). One intriguing possibility is that the milder conditions in 2022 allowed more plants to have multiple flowers open simultaneously, leading to an increase in geitonogamous (between-flower) selfing, an important selfing mode in *M. guttatus* (Leclerc-Potvin & Ritland, 1994). But overall, selfing is probably a fairly consistent partial barrier across years.

Phenological isolation is also a strong barrier, confirming previous results (Kenney & Sweigart, 2016), and is particularly important between the *M. guttatus* and admixed cohorts that

are less isolated by mating system. However, phenological isolation is quite variable across years, a result confirmed by differential shifts in offspring ancestry across years. Phenology has a strong genetic component in these species: two major QTL contributing to differences in photoperiod response have been mapped to candidate genes (Fishman et al., 2014). Our results suggest that these genetic differences are only part of the picture, with isolation mediated by other factors. In particular, stronger isolation in drier (i.e., 2012, 2019) compared to wetter (i.e., 2022) seasons points to water availability as a key factor influencing the strength of phenological isolation, acting through changes in flowering duration and abundance. Precipitation amount and variability are frequently associated with shifts in flowering phenology across plant taxa (Bartlett et al., 2023; Ganjurjav et al., 2020; Matthews & Mazer, 2016; Sianta & Kay, 2021). With ongoing sampling at Catherine Creek, we will be able to test whether the correlation between precipitation and phenological isolation holds across time. Direct measurements of water availability in microsites across the growing season will also help confirm this relationship in the future.

Impact of fluctuating environments on hybrid zones

Fluctuating environmental conditions can provide a form of balancing selection, maintaining allelic diversity by favoring different alleles in different years (Abdul-Rahman et al., 2021; Delph & Kelly, 2014; Han et al., 2020). Similarly, environmental fluctuations may help maintain a variety of ancestry combinations after admixture: (Sianta et al., 2024) found a correlation between interannual variance in precipitation and the extent of introgression across replicate contact zones. Environmental fluctuations could influence the distribution of hybrid ancestry in two main ways: varying the strength of selection against hybrids (Tataru et al., 2023), or directly modulating the strength of premating reproductive isolation (Franks & Weis, 2009;

Seehausen et al., 2008). Our data support the importance of the latter effect, although the former may also play a role. Reproductive barriers that are sensitive to environmental cues, like phenology, may be important sources of fluctuating reproductive isolation across systems.

Climatic variability is projected to increase in many ecosystems with global climate change (Hunt & Elliott, 2004; Rind et al., 1989; Salinger, 2005). It is therefore imperative that we understand the effects of environmental variability on populations. In the context of climate change, hybridization is predicted to have both beneficial effects, such as increased genetic diversity and adaptive potential, and deleterious effects, such as swamping of rare taxa and homogenization of genetic diversity, each of which is likely to be context- and system-specific (Brauer et al., 2023; Chunco, 2014; Franks & Weis, 2009; Muhlfeld et al., 2014; Savage & Vellend, 2015; Vallejo-Marín & Hiscock, 2016). Our study suggests that climatic variability itself can impact the extent of reproductive isolation in hybridizing populations. In a warmer, more unpredictable world, fluctuating reproductive isolation may become more common, leading to an increase in complex, dynamic scenarios of hybridization like this one. To manage these scenarios, we need a better understanding of which factors influence reproductive isolation and hybridization, how they change across space and time, and how they impact the structure and composition of admixed populations.

References

- Abbott, R. J. (1992). Plant invasions, interspecific hybridization and the evolution of new plant taxa. *Trends in Ecology & Evolution*, 7(12), 401–405. [https://doi.org/10.1016/0169-5347\(92\)90020-C](https://doi.org/10.1016/0169-5347(92)90020-C)
- Abbott, R. J. (2017). Plant speciation across environmental gradients and the occurrence and nature of hybrid zones. *Journal of Systematics and Evolution*, 55(4), 238–258. <https://doi.org/10.1111/jse.12267>

- Abdul-Rahman, F., Tranchina, D., & Gresham, D. (2021). Fluctuating Environments Maintain Genetic Diversity through Neutral Fitness Effects and Balancing Selection. *Molecular Biology and Evolution*, 38(10), 4362–4375. <https://doi.org/10.1093/molbev/msab173>
- Adey, A., Morrison, H. G., Asan, Xun, X., Kitzman, J. O., Turner, E. H., Stackhouse, B., MacKenzie, A. P., Caruccio, N. C., Zhang, X., & Shendure, J. (2010). Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biology*, 11(12), R119. <https://doi.org/10.1186/gb-2010-11-12-r119>
- Aldridge, G. (2005). Variation in frequency of hybrids and spatial structure among Ipomopsis (Polemoniaceae) contact sites. *New Phytologist*, 167(1), 279–288. <https://doi.org/10.1111/j.1469-8137.2005.01413.x>
- Andolfatto, P., Davison, D., Erezyilmaz, D., Hu, T. T., Mast, J., Sunayama-Morita, T., & Stern, D. L. (2011). Multiplexed shotgun genotyping for rapid and efficient genetic mapping. *Genome Research*, 21(4), 610–617. <https://doi.org/10.1101/gr.115402.110>
- Arathi, H. S., & Kelly, J. K. (2004). Corolla Morphology Facilitates Both Autogamy and Bumblebee Pollination in *Mimulus guttatus*. *International Journal of Plant Sciences*, 165(6), 1039–1045. <https://doi.org/10.1086/423876>
- Arnold, M. L., & Kunte, K. (2017). *Adaptive Genetic Exchange: A Tangled History of Admixture and Evolutionary Innovation*. <https://doi.org/10.1016/j.tree.2017.05.007>
- Ayres, D. R., Zaremba, K., & Strong, D. R. (2004). Extinction of a Common Native Species by Hybridization with an Invasive Congener. *Weed Technology*, 18(1), 1288–1291. [https://doi.org/10.1614/0890-037X\(2004\)018\[1288:EOACNS\]2.0.CO;2](https://doi.org/10.1614/0890-037X(2004)018[1288:EOACNS]2.0.CO;2)
- Bank, C., Bürger, R., & Hermisson, J. (2012). The limits to parapatric speciation: Dobzhansky-Muller incompatibilities in a continent-island model. *Genetics*, 191(3), 845–863. <https://doi.org/10.1534/genetics.111.137513>
- Barreto, F. S., Watson, E. T., Lima, T. G., Willett, C. S., Edmands, S., Li, W., & Burton, R. S. (2018). Genomic signatures of mitonuclear coevolution across populations of *Tigriopus californicus*. *Nature Ecology & Evolution*, 2(8), 1250–1257. <https://doi.org/10.1038/s41559-018-0588-1>
- Bartlett, K. B., Austin, M. W., Beck, J. B., Zanne, A. E., & Smith, A. B. (2023). Beyond the usual climate? Factors determining flowering and fruiting phenology across a genus over 117 years. *American Journal of Botany*, 110(7), e16188. <https://doi.org/10.1002/ajb2.16188>
- Baskett, M. L., & Gomulkiewicz, R. (2011). Introgressive hybridization as a mechanism for species rescue. *Theoretical Ecology*, 4(2), 223–239. <https://doi.org/10.1007/s12080-011-0118-0>

- Behm, J. E., Ives, A. R., & Boughman, J. W. (2010). Breakdown in postmating isolation and the collapse of a species pair through hybridization. *The American Naturalist*, 175(1), 11–26. <https://doi.org/10.1086/648559>
- Bergland, A. O., Behrman, E. L., O'Brien, K. R., Schmidt, P. S., & Petrov, D. A. (2014). Genomic Evidence of Rapid and Stable Adaptive Oscillations over Seasonal Time Scales in *Drosophila*. *PLOS Genetics*, 10(11), e1004775. <https://doi.org/10.1371/journal.pgen.1004775>
- Bettles, C. M., Docker, M. F., Dufour, B., & Heath, D. D. (2005). Hybridization dynamics between sympatric species of trout: Loss of reproductive isolation. *Journal of Evolutionary Biology*, 18(5), 1220–1233. <https://doi.org/10.1111/j.1420-9101.2005.00935.x>
- Bogdanova, V. S., Galieva, E. R., & Kosterin, O. E. (2009). Genetic analysis of nuclear-cytoplasmic incompatibility in pea associated with cytoplasm of an accession of wild subspecies *Pisum sativum* subsp. *Elatius* (Bieb.) Schmahl. *Theoretical and Applied Genetics*, 118(4), 801–809. <https://doi.org/10.1007/s00122-008-0940-y>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Borge, T., Lindroos, K., Nadvornik, P., Syvanen, A. C., & Saetre, G. P. (2005). Amount of introgression in flycatcher hybrid zones reflects regional differences in pre and post-zygotic barriers to gene exchange. *Journal of Evolutionary Biology*, 18(6), 1416–1424. <https://doi.org/10.1111/j.1420-9101.2005.00964.x>
- Borzée, A., Kim, J. Y., Da Cunha, M. A. M., Lee, D., Sin, E., Oh, S., Yi, Y., & Jang, Y. (2016). Temporal and spatial differentiation in microhabitat use: Implications for reproductive isolation and ecological niche specification. *Integrative Zoology*, 11(5), 375–387. <https://doi.org/10.1111/1749-4877.12200>
- Brandvain, Y., Kenney, A. M., Flagel, L., Coop, G., & Sweigart, A. L. (2014). Speciation and Introgression between *Mimulus nasutus* and *Mimulus guttatus*. *PLoS Genetics*, 10(6), e1004410. <https://doi.org/10.1371/journal.pgen.1004410>
- Brauer, C. J., Sandoval-Castillo, J., Gates, K., Hammer, M. P., Unmack, P. J., Bernatchez, L., & Beheregaray, L. B. (2023). Natural hybridization reduces vulnerability to climate change. *Nature Climate Change*, 13(3), 282–289. <https://doi.org/10.1038/s41558-022-01585-1>
- Brice, C., Zhang, Z., Bendixsen, D., & Stelkens, R. (2021). Hybridization Outcomes Have Strong Genomic and Environmental Contingencies. *The American Naturalist*, 198(3), E53–E67. <https://doi.org/10.1086/715356>
- Broad Institute. (2019). *Picard Toolkit*, *GitHub repository* [Computer software]. <https://broadinstitute.github.io/picard/>

- Butlin, R. K., Saura, M., Charrier, G., Jackson, B., André, C., Caballero, A., Coyne, J. A., Galindo, J., Grahame, J. W., Hollander, J., Kemppainen, P., Martínez-Fernández, M., Panova, M., Quesada, H., Johannesson, K., & Rolán-Alvarez, E. (2014). Parallel evolution of local adaptation and reproductive isolation in the face of gene flow. *Evolution*, 68(4), 935–949. <https://doi.org/10.1111/evo.12329>
- Carney, S. E., Gardner, K. A., & Rieseberg, L. H. (2000). Evolutionary Changes Over the Fifty-Year History of a Hybrid Population of Sunflowers (*helianthus*). *Evolution*, 54(2), 462–474. <https://doi.org/10.1111/j.0014-3820.2000.tb00049.x>
- Chhatre, V. E., Evans, L. M., DiFazio, S. P., & Keller, S. R. (2018). Adaptive introgression and maintenance of a trispecies hybrid complex in range-edge populations of *Populus*. *Molecular Ecology*, 27(23), 4820–4838. <https://doi.org/10.1111/mec.14820>
- Christie, K., Fraser, L. S., & Lowry, D. B. (2022). The strength of reproductive isolating barriers in seed plants: Insights from studies quantifying premating and postmating reproductive barriers over the past 15 years. *Evolution*, evo.14565. <https://doi.org/10.1111/evo.14565>
- Chunco, A. J. (2014). Hybridization in a warmer world. *Ecology and Evolution*, 4(10), 2019–2031. <https://doi.org/10.1002/ece3.1052>
- Clarke, B. C., Johnson, M. S., Murray, J., Hewitt, G. M., Wragg, G. M., Clarke, B. C., & Grant, P. R. (1997). Clines in the genetic distance between two species of island land snails: How ‘molecular leakage’ can mislead us about speciation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 351(1341), 773–784. <https://doi.org/10.1098/rstb.1996.0072>
- Colicchio, J., Monnahan, P. J., Wessinger, C. A., Brown, K., Kern, J. R., & Kelly, J. K. (2020). Individualized mating system estimation using genomic data. *Molecular Ecology Resources*, 20(1), 333–347. <https://doi.org/10.1111/1755-0998.13094>
- Comes, H. P., & Abbott, R. J. (2001). Molecular phylogeography, reticulation, and lineage sorting in mediterranean *senecio* sect. *Senecio* (asteraceae). *Evolution*, 55(10), 1943–1962. <https://doi.org/10.1111/j.0014-3820.2001.tb01312.x>
- Corbett-Detig, R., & Nielsen, R. (2017). A Hidden Markov Model Approach for Simultaneously Estimating Local Ancestry and Admixture Time Using Next Generation Sequence Data in Samples of Arbitrary Ploidy. *PLOS Genetics*, 13(1), e1006529. <https://doi.org/10.1371/journal.pgen.1006529>
- Coyne, J. A., & Orr, H. A. (1997). Patterns of Speciation in *Drosophila* Revisited. *Evolution*, 51(1), 295. <https://doi.org/10.2307/2410984>
- Cribari-Neto, F., & Zeileis, A. (2010). Beta Regression in R. *Journal of Statistical Software*, 34, 1–24. <https://doi.org/10.18637/jss.v034.i02>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & Group, 1000

- Genomes Project Analysis. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156. <https://doi.org/10.1093/bioinformatics/btr330>
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, 10(2), giab008. <https://doi.org/10.1093/gigascience/giab008>
- Delph, L. F., & Kelly, J. K. (2014). On the importance of balancing selection in plants. *New Phytologist*, 201(1), 45–56. <https://doi.org/10.1111/nph.12441>
- Dobzhansky, Th. (1934). Studies on hybrid sterility. *Zeitschrift Für Zellforschung Und Mikroskopische Anatomie*, 21(2), 169–223. <https://doi.org/10.1007/BF00374056>
- Dole, J. A. (1992). Reproductive Assurance Mechanisms in Three Taxa of the *Mimulus Guttatus* Complex (scrophulariaceae). *American Journal of Botany*, 79(6), 650–659. <https://doi.org/10.1002/j.1537-2197.1992.tb14607.x>
- Dudash, M. R., & Ritland, K. (1991). Multiple Paternity and Self-Fertilization in Relation to Floral Age in *Mimulus Guttatus* (scrophulariaceae). *American Journal of Botany*, 78(12), 1746–1753. <https://doi.org/10.1002/j.1537-2197.1991.tb14539.x>
- Ebersole, J. P. (1985). Niche Separation of Two Damsel fish Species by Aggression and Differential Microhabitat Utilization. *Ecology*, 66(1), 14–20. <https://doi.org/10.2307/1941302>
- Farnitano, M. C., & Sweigart, A. L. (2023). Strong postmating reproductive isolation in *Mimulus* section *Eunanus*. *Journal of Evolutionary Biology*, 36(10), 1393–1410. <https://doi.org/10.1111/jeb.14219>
- Farnitano, M. C., & Sweigart, A. L. (2024). *Low-cost tagmentation library prep for low-coverage Illumina sequencing*. protocols.io. [dx.doi.org/10.17504/protocols.io.6qpvr8zrplmk/v1](https://doi.org/10.17504/protocols.io.6qpvr8zrplmk/v1)
- Fishman, L. (2020). *96-well CTAB-chloroform DNA extraction*. protocols.io. <https://dx.doi.org/10.17504/protocols.io.bgv6jw9e>
- Fishman, L., Kelly, A. J., & Willis, J. H. (2002). Minor quantitative trait loci underlie floral traits associated with mating system divergence in *Mimulus*. *Evolution*, 56(11), 2138–2155. <https://doi.org/10.1111/j.0014-3820.2002.tb00139.x>
- Fishman, L., & Sweigart, A. L. (2018). When Two Rights Make a Wrong: The Evolutionary Genetics of Plant Hybrid Incompatibilities. *Annual Review of Plant Biology*, 69(1), 707–731. <https://doi.org/10.1146/annurev-arplant-042817-040113>
- Fishman, L., Sweigart, A. L., Kenney, A. M., & Campbell, S. (2014). Major quantitative trait loci control divergence in critical photoperiod for flowering between selfing and outcrossing species of monkeyflower (*Mimulus*). *New Phytologist*, 201(4), 1498–1507. <https://doi.org/10.1111/nph.12618>

- Fishman, L., & Willis, J. H. (2006). A cytonuclear incompatibility causes anther sterility in *Mimulus* hybrids. *Evolution*, 60(7), 1372–1381. <https://doi.org/10.1111/j.0014-3820.2006.tb01216.x>
- Folk, R. A., Mandel, J. R., & Freudenstein, J. V. (2017). Ancestral Gene Flow and Parallel Organellar Genome Capture Result in Extreme Phylogenomic Discord in a Lineage of Angiosperms. *Systematic Biology*, 66(3), 320–337. <https://doi.org/10.1093/sysbio/syw083>
- Franks, S. J., & Weis, A. E. (2009). Climate change alters reproductive isolation and potential gene flow in an annual plant. *Evolutionary Applications*, 2(4), 481–488. <https://doi.org/10.1111/j.1752-4571.2009.00073.x>
- Ganjurjav, H., Gornish, E. S., Hu, G., Schwartz, M. W., Wan, Y., Li, Y., & Gao, Q. (2020). Warming and precipitation addition interact to affect plant spring phenology in alpine meadows on the central Qinghai-Tibetan Plateau. *Agricultural and Forest Meteorology*, 287, 107943. <https://doi.org/10.1016/j.agrformet.2020.107943>
- Gilman, R. T., & Behm, J. E. (2011). Hybridization, Species Collapse, and Species Reemergence After Disturbance to Premating Mechanisms of Reproductive Isolation. *Evolution*, 65(9), 2592–2605. <https://doi.org/10.1111/j.1558-5646.2011.01320.x>
- Gobron, N., Waszczak, C., Simon, M., Hiard, S., Boivin, S., Charif, D., Ducamp, A., Wenes, E., & Budar, F. (2013). A Cryptic Cytoplasmic Male Sterility Unveils a Possible Gynodioecious Past for *Arabidopsis thaliana*. *PLoS ONE*, 8(4). <https://doi.org/10.1371/journal.pone.0062450>
- Goodman, S. J., Barton, N. H., Swanson, G., Abernethy, K., & Pemberton, J. M. (1999). Introgression Through Rare Hybridization: A Genetic Study of a Hybrid Zone Between Red and Sika Deer (Genus *Cervus*) in Argyll, Scotland. *Genetics*, 152(1), 355–371. <https://doi.org/10.1093/genetics/152.1.355>
- Goulet, B. E., Roda, F., & Hopkins, R. (2017). Hybridization in plants: Old ideas, new techniques. *Plant Physiology*, 173(1), 65–78. <https://doi.org/10.1104/pp.16.01340>
- Grant, B. R., & Grant, P. R. (1996). High Survival of Darwin's Finch Hybrids: Effects of Beak Morphology and Diets. *Ecology*, 77(2), 500–509.
- Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M. H.-Y., Hansen, N. F., Durand, E. Y., Malaspinas, A.-S., Jensen, J. D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H. A., ... Pääbo, S. (2010). A draft sequence of the Neandertal genome. *Science (New York, N.Y.)*, 328(5979), 710–722. <https://doi.org/10.1126/science.1188021>
- Han, G., Wang, W., & Dong, Y. (2020). Effects of balancing selection and microhabitat temperature variations on heat tolerance of the intertidal black mussel *Septifer virgatus*. *Integrative Zoology*, 15(5), 416–427. <https://doi.org/10.1111/1749-4877.12439>

- Harrison, R. G., & Larson, E. L. (2016). Heterogeneous genome divergence, differential introgression, and the origin and structure of hybrid zones. *Molecular Ecology*, 25(11), 2454–2466. <https://doi.org/10.1111/mec.13582>
- Hasselman, D. J., Argo, E. E., McBride, M. C., Bentzen, P., Schultz, T. F., Perez-Umphrey, A. A., & Palkovacs, E. P. (2014). Human disturbance causes the formation of a hybrid swarm between two naturally sympatric fish species. *Molecular Ecology*, 23(5), 1137–1152. <https://doi.org/10.1111/mec.12674>
- Hendry, A. P., Berg, O. K., & Quinn, T. P. (1999). Condition Dependence and Adaptation-by-Time: Breeding Date, Life History, and Energy Allocation within a Population of Salmon. *Oikos*, 85(3), 499–514. <https://doi.org/10.2307/3546699>
- Hendry, A. P., Wenburg, J. K., Bentzen, P., Volk, E. C., & Quinn, T. P. (2000). Rapid Evolution of Reproductive Isolation in the Wild: Evidence from Introduced Salmon. *Science*, 290(5491), 516–518. <https://doi.org/10.1126/science.290.5491.516>
- Hood, G. R., Zhang, L., Hu, E. G., Ott, J. R., & Egan, S. P. (2019). Cascading reproductive isolation: Plant phenology drives temporal isolation among populations of a host-specific herbivore. *Evolution*, 73(3), 554–568. <https://doi.org/10.1111/evo.13683>
- Hunt, B. G., & Elliott, T. I. (2004). Interaction of climatic variability with climatic change. *Atmosphere-Ocean*, 42(3), 145–172. <https://doi.org/10.3137/ao.420301>
- Karron, J. D., Holmquist, K. G., Flanagan, R. J., & Mitchell, R. J. (2009). Pollinator visitation patterns strongly influence among-flower variation in selfing rate. *Annals of Botany*, 103(9), 1379–1383. <https://doi.org/10.1093/aob/mcp030>
- Karron, J. D., Mitchell, R. J., Holmquist, K. G., Bell, J. M., & Funk, B. (2004). The influence of floral display size on selfing rates in *Mimulus ringens*. *Heredity*, 92(3), 242–248. <https://doi.org/10.1038/sj.hdy.6800402>
- Kelly, J. K., & Willis, J. H. (2002). A Manipulative Experiment to Estimate Biparental Inbreeding in Monkeyflowers. *International Journal of Plant Sciences*, 163(4), 575–579. <https://doi.org/10.1086/340735>
- Kenney, A. M., & Sweigart, A. L. (2016). Reproductive isolation and introgression between sympatric *Mimulus* species. *Molecular Ecology*, 25(11), 2499–2517. <https://doi.org/10.1111/mec.13630>
- Kiang, Y. T., & Hamrick, J. L. (1978). Reproductive Isolation in the *Mimulus guttatus* M. nasutus Complex. *American Midland Naturalist*, 100(2), 269. <https://doi.org/10.2307/2424826>
- Kisel, Y., & Barraclough, T. G. (2010). Speciation Has a Spatial Scale That Depends on Levels of Gene Flow. *The American Naturalist*, 175(3), 316–334. <https://doi.org/10.1086/650369>

- Klein, E. K., Lagache-Navarro, L., & Petit, R. J. (2017). Demographic and spatial determinants of hybridization rate. *Journal of Ecology*, *105*(1), 29–38. <https://doi.org/10.1111/1365-2745.12674>
- Kleindorfer, S., O'Connor, J. A., Dudaniec, R. Y., Myers, S. A., Robertson, J., & Sulloway, F. J. (2014). Species collapse via hybridization in Darwin's tree finches. *The American Naturalist*, *183*(3), 325–341. <https://doi.org/10.1086/674899>
- Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, *15*(1), 356. <https://doi.org/10.1186/s12859-014-0356-4>
- Langdon, Q. K., Groh, J. S., Aguillon, S. M., Powell, D. L., Gunn, T., Payne, C., Baczenas, J. J., Donny, A., Dodge, T. O., Du, K., Schartl, M., Ríos-Cárdenas, O., Gutiérrez-Rodríguez, C., Morris, M., & Schumer, M. (2024). Swordtail fish hybrids reveal that genome evolution is surprisingly predictable after initial hybridization. *PLOS Biology*, *22*(8), e3002742. <https://doi.org/10.1371/journal.pbio.3002742>
- Larson, E. L., Andrés, J. A., Bogdanowicz, S. M., & Harrison, R. G. (2013). Differential introgression in a mosaic hybrid zone reveals candidate barrier genes. *Evolution*, *67*(12), 3653–3661. <https://doi.org/10.1111/evo.12205>
- Leclerc-Potvin, C., & Ritland, K. (1994). Modes of self-fertilization in *Mimulus guttatus* (Scrophulariaceae): A field experiment. *American Journal of Botany*, *81*(2), 199–205. <https://doi.org/10.1002/j.1537-2197.1994.tb15430.x>
- Lee, H. Y., Chou, J. Y., Cheong, L., Chang, N. H., Yang, S. Y., & Leu, J. Y. (2008). Incompatibility of Nuclear and Mitochondrial Genomes Causes Hybrid Sterility between Two Yeast Species. *Cell*, *135*(6), 1065–1073. <https://doi.org/10.1016/j.cell.2008.10.047>
- Leigh, J. W., & Bryant, D. (2015). popart: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110–1116. <https://doi.org/10.1111/2041-210X.12410>
- Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, *17*(4), 183–189. [https://doi.org/10.1016/S0169-5347\(02\)02497-7](https://doi.org/10.1016/S0169-5347(02)02497-7)
- Lepais, O., Petit, R. J., Guichoux, E., Lavabre, J. E., Alberto, F., Kremer, A., & Gerber, S. (2009). Species relative abundance and direction of introgression in oaks. *Molecular Ecology*, *18*(10), 2228–2242. <https://doi.org/10.1111/j.1365-294X.2009.04137.x>
- Levene, H. (1953). Genetic Equilibrium When More Than One Ecological Niche is Available. *The American Naturalist*, *87*(836), 331–333. <https://doi.org/10.1086/281792>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>

- Ling, T. C., Phokasem, P., Sinpoo, C., Yang, Y.-P., & Disayathanoowat, T. (2022). Microhabitat and Pollinator Differentiation Drive Reproductive Isolation between Two Sympatric *Salvia* Species (Lamiaceae). *Plants*, *11*(18), Article 18. <https://doi.org/10.3390/plants11182423>
- Lotsy, J. P. (1931). On the species of the taxonomist in its relation to evolution. *Genetica*, *13*(1), 1–16. <https://doi.org/10.1007/BF01725037>
- Lowry, D. B., & Willis, J. H. (2010). A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biology*, *8*(9). <https://doi.org/10.1371/journal.pbio.1000500>
- Lu, Z., Hofmeister, B. T., Vollmers, C., DuBois, R. M., & Schmitz, R. J. (2017). Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes. *Nucleic Acids Research*, *45*(6), e41. <https://doi.org/10.1093/nar/gkw1179>
- Machado, C. A., Haselkorn, T. S., & Noor, M. A. F. (2007). Evaluation of the Genomic Extent of Effects of Fixed Inversion Differences on Intraspecific Variation and Interspecific Gene Flow in *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, *175*(3), 1289–1306. <https://doi.org/10.1534/genetics.106.064758>
- Macnair, M. R., Macnair, V. E., & Martin, B. E. (1989). Adaptive speciation in *Mimulus*: An ecological comparison of *M. cupripetalus* with its presumed progenitor, *M. guttatus*. *New Phytologist*, *112*(2), 269–279. <https://doi.org/10.1111/j.1469-8137.1989.tb02383.x>
- Mallet, J., Besansky, N., & Hahn, M. W. (2015). How reticulated are species? *BioEssays*, *38*(2), 140–149. <https://doi.org/10.1002/bies.201500149>
- Malone, J. H., & Fontenot, B. E. (2008). Patterns of Reproductive Isolation in Toads. *PLoS ONE*, *3*(12), e3900. <https://doi.org/10.1371/journal.pone.0003900>
- Mantel, S. J., & Sweigart, A. L. (2019). Divergence in drought-response traits between sympatric species of *Mimulus*. *Ecology and Evolution*. <https://doi.org/10.1002/ece3.5549>
- Mantel, S. J., & Sweigart, A. L. (2024). Postzygotic barriers persist despite ongoing introgression in hybridizing *Mimulus* species. *Molecular Ecology*, *33*(4), e17261. <https://doi.org/10.1111/mec.17261>
- Manzoor, S. A., Griffiths, G., Obiakara, M. C., Esparza-Estrada, C. E., & Lukac, M. (2020). Evidence of ecological niche shift in *Rhododendron ponticum* (L.) in Britain: Hybridization as a possible cause of rapid niche expansion. *Ecology and Evolution*, *10*(4), 2040–2050. <https://doi.org/10.1002/ece3.6036>
- Martin, N. H., & Willis, J. H. (2007). Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric *Mimulus* species. *Evolution*, *61*(1), 68–82. <https://doi.org/10.1111/j.1558-5646.2007.00006.x>

- Martinsen, G. D., Whitham, T. G., Turek, R. J., & Keim, P. (2001). Hybrid Populations Selectively Filter Gene Introgression Between Species. *Evolution*, 55(7), 1325–1335. <https://doi.org/10.1111/j.0014-3820.2001.tb00655.x>
- Matthews, E. R., & Mazer, S. J. (2016). Historical changes in flowering phenology are governed by temperature \times precipitation interactions in a widespread perennial herb in western North America. *New Phytologist*, 210(1), 157–167. <https://doi.org/10.1111/nph.13751>
- Meiklejohn, C. D., Holmbeck, M. A., Siddiq, M. A., Abt, D. N., Rand, D. M., & Montooth, K. L. (2013). An Incompatibility between a Mitochondrial tRNA and Its Nuclear-Encoded tRNA Synthetase Compromises Development and Fitness in *Drosophila*. *PLOS Genetics*, 9(1), e1003238. <https://doi.org/10.1371/journal.pgen.1003238>
- Meisner, J., & Albrechtsen, A. (2018). Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data. *Genetics*, 210(2), 719–731. <https://doi.org/10.1534/genetics.118.301336>
- Melo, M. C., Grealy, A., Brittain, B., Walter, G. M., & Ortiz-Barrientos, D. (2014). Strong extrinsic reproductive isolation between parapatric populations of an Australian groundsel. *New Phytologist*, 203(1), 323–334. <https://doi.org/10.1111/nph.12779>
- Modliszewski, J. L., & Willis, J. H. (2012). Allotetraploid *Mimulus sookensis* are highly interfertile despite independent origins. *Molecular Ecology*, 21(21), 5280–5298. <https://doi.org/10.1111/j.1365-294X.2012.05706.x>
- Moore, W. S., & Buchanan, E. B. (1985). Stability of the northern flicker hybrid zone in historical times: Implications for adaptive speciation theory. *Evolution*, 39(1), 135–151. <https://doi.org/10.1111/j.1558-5646.1985.tb04086.x>
- Mower, J. P., Case, A. L., Floro, E. R., & Willis, J. H. (2012). Evidence against Equimolarity of Large Repeat Arrangements and a Predominant Master Circle Structure of the Mitochondrial Genome from a Monkeyflower (*Mimulus guttatus*) Lineage with Cryptic CMS. *Genome Biology and Evolution*, 4(5), 670–686. <https://doi.org/10.1093/gbe/evs042>
- Moyle, L. C., Olson, M. S., & Tiffin, P. (2004). Patterns of reproductive isolation in three angiosperm genera. *Evolution*, 58(6), 1195–1208. <https://doi.org/10.1111/j.0014-3820.2004.tb01700.x>
- Muhlfeld, C. C., Kovach, R. P., Jones, L. A., Al-Chokhachy, R., Boyer, M. C., Leary, R. F., Lowe, W. H., Luikart, G., & Allendorf, F. W. (2014). Invasive hybridization in a threatened species is accelerated by climate change. *Nature Climate Change*, 4(7), 620–624. <https://doi.org/10.1038/nclimate2252>
- Muller, H. J. (1942). Isolating mechanisms, evolution, and temperature. *Biological Symposia*, 6, 71–125.

- Musker, S. D., Ellis, A. G., Schlebusch, S. A., & Verboom, G. A. (2021). Niche specificity influences gene flow across fine-scale habitat mosaics in Succulent Karoo plants. *Molecular Ecology*, 30(1), 175–192. <https://doi.org/10.1111/mec.15721>
- Nelson, T. C., Stathos, A. M., Vanderpool, D. D., Finseth, F. R., Yuan, Y., & Fishman, L. (2021). Ancient and recent introgression shape the evolutionary history of pollinator adaptation and speciation in a model monkeyflower radiation (*Mimulus* section *Erythranthe*). *PLoS Genetics*, 17(2), e1009095. <https://doi.org/10.1371/journal.pgen.1009095>
- Nice, C. C., Gompert, Z., Fordyce, J. A., Forister, M. L., Lucas, L. K., & Buerkle, C. A. (2013). Hybrid speciation and independent evolution in lineages of alpine butterflies. *Evolution*, 67(4), 1055–1068. <https://doi.org/10.1111/evo.12019>
- Okonechnikov, K., Conesa, A., & García-Alcalde, F. (2016). Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics*, 32(2), 292–294. <https://doi.org/10.1093/bioinformatics/btv566>
- Orr, H. A., & Turelli, M. (2001). The evolution of postzygotic isolation: Accumulating Dobzhansky-Muller incompatibilities. *Evolution*, 55(6), 1085–1094. <https://doi.org/10.1111/j.0014-3820.2001.tb00628.x>
- Osmolovsky, I., Shifrin, M., Gamliel, I., Belmaker, J., & Sapir, Y. (2022). Eco-Geography and Phenology Are the Major Drivers of Reproductive Isolation in the Royal Irises, a Species Complex in the Course of Speciation. *Plants*, 11(23), Article 23. <https://doi.org/10.3390/plants11233306>
- Otis, J.-A., Thornton, D., Rutledge, L., & Murray, D. L. (2017). Ecological niche differentiation across a wolf-coyote hybrid zone in eastern North America. *Diversity and Distributions*, 23(5), 529–539. <https://doi.org/10.1111/ddi.12543>
- Oziolor, E. M., Reid, N. M., Yair, S., Lee, K. M., Guberman VerPloeg, S., Bruns, P. C., Shaw, J. R., Whitehead, A., & Matson, C. W. (2019). Adaptive introgression enables evolutionary rescue from extreme environmental pollution. *Science*, 364(6439), 455–457. <https://doi.org/10.1126/science.aav4155>
- Pfennig, K. S., Kelly, A. L., & Pierce, A. A. (2016). Hybridization as a facilitator of species range expansion. *Proceedings of the Royal Society B: Biological Sciences*, 283(1839), 20161329. <https://doi.org/10.1098/rspb.2016.1329>
- Picelli, S., Björklund, Å. K., Reinius, B., Sagasser, S., Winberg, G., & Sandberg, R. (2014). Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Research*, 24(12), 2033–2040. <https://doi.org/10.1101/gr.177881.114>
- Price, A. L., Tandon, A., Patterson, N., Barnes, K. C., Rafaels, N., Ruczinski, I., Beaty, T. H., Mathias, R., Reich, D., & Myers, S. (2009). Sensitive Detection of Chromosomal Segments of Distinct Ancestry in Admixed Populations. *PLOS Genetics*, 5(6), e1000519. <https://doi.org/10.1371/journal.pgen.1000519>

- PRISM Climate Group. (2024). *PRISM Time Series Data* [Dataset]. Oregon State University. <https://prism.oregonstate.edu>
- Quilodr  n, C. S., Tsoupas, A., & Currat, M. (2020). The Spatial Signature of Introgression After a Biological Invasion With Hybridization. *Frontiers in Ecology and Evolution*, 8. <https://www.frontiersin.org/articles/10.3389/fevo.2020.569620>
- R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rand, D. M., & Harrison, R. G. (1989). Ecological genetics of a mosaic hybrid zone: Mitochondrial, nuclear, and reproductive differentiation of crickets by soil type. *Evolution*, 43(2), 432–449. <https://doi.org/10.1111/j.1558-5646.1989.tb04238.x>
- Randi, E., & Lucchini, V. (2002). Detecting rare introgression of domestic dog genes into wild wolf (*Canis lupus*) populations by Bayesian admixture analyses of microsatellite variation. *Conservation Genetics*, 3(1), 29–43. <https://doi.org/10.1023/A:1014229610646>
- Rhymer, J. M., & Simberloff, D. (1996). Extinction by hybridization and introgression. *Annual Review of Ecology, Evolution, and Systematics*, 27(Volume 27, 1996), 83–109. <https://doi.org/10.1146/annurev.ecolsys.27.1.83>
- Rice, P., Longden, I., & Bleasby, A. (2000). EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics*, 16(6), 276–277. [https://doi.org/10.1016/S0168-9525\(00\)02024-2](https://doi.org/10.1016/S0168-9525(00)02024-2)
- Richards, T. J., & Ortiz-Barrientos, D. (2016). Immigrant inviability produces a strong barrier to gene flow between parapatric ecotypes of *Senecio laetus*. *Evolution*, 70(6), 1239–1248. <https://doi.org/10.1111/evo.12936>
- Rifkin, J. L., Castillo, A. S., Liao, I. T., & Rausher, M. D. (2019). Gene flow, divergent selection and resistance to introgression in two species of morning glories (*Ipomoea*). *Molecular Ecology*, 28(7), 1709–1729. <https://doi.org/10.1111/mec.14945>
- Rind, D., Goldberg, R., & Ruedy, R. (1989). Change in climate variability in the 21st century. *Climatic Change*, 14(1), 5–37. <https://doi.org/10.1007/BF00140173>
- Riquet, F., Liautard-Haag, C., Woodall, L., Bouza, C., Louisy, P., Hamer, B., Otero-Ferrer, F., Aublanc, P., B  duneau, V., Briard, O., El Ayari, T., Hochscheid, S., Belkhir, K., Arnaud-Haond, S., Gagnaire, P., & Bierne, N. (2019). Parallel pattern of differentiation at a genomic island shared between clinal and mosaic hybrid zones in a complex of cryptic seahorse lineages. *Evolution*, 73(4), 817–835. <https://doi.org/10.1111/evo.13696>
- Ritland, K. (1989). Correlated Matings in the Partial Selfer *Mimulus Guttatus*. *Evolution*, 43(4), 848–859. <https://doi.org/10.1111/j.1558-5646.1989.tb05182.x>

- Ritland, K., & Ganders, F. R. (1987). Covariation of selfing rates with parental gene fixation indices within populations of *mimulus guttatus*. *Evolution*, 41(4), 760–771. <https://doi.org/10.1111/j.1558-5646.1987.tb05851.x>
- Roda, F., Mendes, F. K., Hahn, M. W., & Hopkins, R. (2017). Genomic evidence of gene flow during reinforcement in Texas Phlox. *Molecular Ecology*, 26(8), 2317–2330. <https://doi.org/10.1111/mec.14041>
- Ross, C. L., & Harrison, R. G. (2002). A Fine-Scale Spatial Analysis of the Mosaic Hybrid Zone Between *Gryllus Firmus* and *Gryllus Pennsylvanicus*. *Evolution*, 56(11), 2296–2312. <https://doi.org/10.1111/j.0014-3820.2002.tb00153.x>
- Rosser, N., Seixas, F., Queste, L. M., Cama, B., Mori-Pezo, R., Kryvokhyzha, D., Nelson, M., Waite-Hudson, R., Goringe, M., Costa, M., Elias, M., Mendes Eleres de Figueiredo, C., Freitas, A. V. L., Joron, M., Kozak, K., Lamas, G., Martins, A. R. P., McMillan, W. O., Ready, J., ... Dasmahapatra, K. K. (2024). Hybrid speciation driven by multilocus introgression of ecological traits. *Nature*, 628(8009), 811–817. <https://doi.org/10.1038/s41586-024-07263-w>
- Ruhsam, M., Hollingsworth, P. M., & Ennos, R. A. (2011). Early evolution in a hybrid swarm between outcrossing and selfing lineages in *Geum*. *Heredity*, 107(3), 246–255. <https://doi.org/10.1038/hdy.2011.9>
- Salinger, M. J. (2005). Climate Variability and Change: Past, Present and Future – An Overview. *Climatic Change*, 70(1), 9–29. <https://doi.org/10.1007/s10584-005-5936-x>
- Sambatti, J. B. M., Strasburg, J. L., Ortiz-Barrientos, D., Baack, E. J., & Rieseberg, L. H. (2012). Reconciling Extremely Strong Barriers with High Levels of Gene Exchange in Annual Sunflowers. *Evolution*, 66(5), 1459–1473. <https://doi.org/10.1111/j.1558-5646.2011.01537.x>
- Savage, J., & Vellend, M. (2015). Elevational shifts, biotic homogenization and time lags in vegetation change during 40 years of climate warming. *Ecography*, 38(6), 546–555. <https://doi.org/10.1111/ecog.01131>
- Schumer, M., Powell, D. L., & Corbett-Detig, R. (2020). Versatile simulations of admixture and accurate local ancestry inference with mixnmatch and ancestryinfer. *Molecular Ecology Resources*, 20(4), 1141–1151. <https://doi.org/10.1111/1755-0998.13175>
- Schumer, M., Rosenthal, G. G., & Andolfatto, P. (2014). How common is homoploid hybrid speciation? *Evolution*, 68(6), 1553–1560. <https://doi.org/10.1111/evo.12399>
- Schumer, M., Xu, C., Powell, D. L., Durvasula, A., Skov, L., Holland, C., Blazier, J. C., Sankararaman, S., Andolfatto, P., Gil, †, Rosenthal, G., & Przeworski, M. (2018). Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science*, 360(6389), 656–660.

- Schwarzbach, A. E., Donovan, L. A., & Rieseberg, L. H. (2001). Transgressive character expression in a hybrid sunflower species. *American Journal of Botany*, 88(2), 270–277. <https://doi.org/10.2307/2657018>
- Scopece, G., Widmer, A., & Cozzolino, S. (2008). Evolution of Postzygotic Reproductive Isolation in a Guild of Deceptive Orchids. *The American Naturalist*, 171(3), 315–326. <https://doi.org/10.1086/527501>
- Seehausen, O., Takimoto, G., Roy, D., & Jokela, J. (2008). Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Molecular Ecology*, 17(1), 30–44. <https://doi.org/10.1111/j.1365-294X.2007.03529.x>
- Selz, O. M., & Seehausen, O. (2019). Interspecific hybridization can generate functional novelty in cichlid fish. *Proceedings of the Royal Society B: Biological Sciences*, 286(1913), 20191621. <https://doi.org/10.1098/rspb.2019.1621>
- Servedio, M. R., & Hermisson, J. (2020). The evolution of partial reproductive isolation as an adaptive optimum. *Evolution*, 74(1), 4–14. <https://doi.org/10.1111/evo.13880>
- Shuker, D., Underwood, K., King, T. M., & Butlin, R. K. (2005). Patterns of male sterility in a grasshopper hybrid zone imply accumulation of hybrid incompatibilities without selection. *Proceedings of the Royal Society B: Biological Sciences*, 272(1580), 2491–2497. <https://doi.org/10.4135/9781412972093.n399>
- Sianta, S. A., & Kay, K. M. (2021). Parallel evolution of phenological isolation across the speciation continuum in serpentine-adapted annual wildflowers. *Proceedings. Biological Sciences*, 288(1948), 20203076. <https://doi.org/10.1098/rspb.2020.3076>
- Sianta, S. A., Moeller, D. A., & Brandvain, Y. (2024). The extent of introgression between incipient *Clarkia* species is determined by temporal environmental variation and mating system. *Proceedings of the National Academy of Sciences*, 121(12), e2316008121. <https://doi.org/10.1073/pnas.2316008121>
- Simon, A., Fraïsse, C., El Ayari, T., Liautard-Haag, C., Strelkov, P., Welch, J. J., & Bierne, N. (2021). How do species barriers decay? Concordance and local introgression in mosaic hybrid zones of mussels. *Journal of Evolutionary Biology*, 34(1), 208–223. <https://doi.org/10.1111/jeb.13709>
- Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating Individual Admixture Proportions from Next Generation Sequencing Data. *Genetics*, 195(3), 693. <https://doi.org/10.1534/genetics.113.154138>
- Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, 236(4803), 787–792. <https://doi.org/10.1126/science.3576198>
- Sobel, J. M., & Chen, G. F. (2014). Unification of methods for estimating the strength of reproductive isolation. *Evolution*, 68(5), 1511–1522. <https://doi.org/10.1111/evo.12362>

- Stebbins, G. L. (1959). The Role of Hybridization in Evolution. *Proceedings of the American Philosophical Society*, 103(2), 231–251.
- Stephen Dobson, F. (1982). Competition for mates and predominant juvenile male dispersal in mammals. *Animal Behaviour*, 30(4), 1183–1192. [https://doi.org/10.1016/S0003-3472\(82\)80209-1](https://doi.org/10.1016/S0003-3472(82)80209-1)
- Suarez-Gonzalez, A., Lexer, C., & Cronk, Q. C. B. (2018). Adaptive introgression: A plant perspective. *Biology Letters*, 14(3), 20170688. <https://doi.org/10.1098/rsbl.2017.0688>
- Sullivan, B. K. (1995). Temporal stability in hybridization between *Bufo microscaphus* and *Bufo woodhousii* (Anura: Bufonidae): Behavior and morphology. *Journal of Evolutionary Biology*, 8(2), 233–247. <https://doi.org/10.1046/j.1420-9101.1995.8020233.x>
- Sweigart, A. L., Martin, N. H., & Willis, J. H. (2008). Patterns of nucleotide variation and reproductive isolation between a *Mimulus* allotetraploid and its progenitor species. *Molecular Ecology*, 17(8), 2089–2100. <https://doi.org/10.1111/j.1365-294X.2008.03707.x>
- Sweigart, A. L., & Willis, J. H. (2003). Patterns of nucleotide diversity in two species of *mimulus* are affected by mating system and asymmetric introgression. *Evolution*, 57(11), 2490–2506. <https://doi.org/10.1111/j.0014-3820.2003.tb01494.x>
- Tataru, D., Wheeler, E. C., & Ferris, K. G. (2023). Spatially and temporally varying selection influence species boundaries in two sympatric *Mimulus*. *Proceedings of the Royal Society B: Biological Sciences*, 290(1992), 20222279. <https://doi.org/10.1098/rspb.2022.2279>
- Taylor, S. A., Curry, R. L., White, T. A., Ferretti, V., & Lovette, I. (2014). Spatiotemporally consistent genomic signatures of reproductive isolation in a moving hybrid zone. *Evolution*, 68(11), 3066–3081. <https://doi.org/10.1111/evo.12510>
- Troth, A., Puzey, J. R., Kim, R. S., Willis, J. H., & Kelly, J. K. (2018). Selective trade-offs maintain alleles underpinning complex trait variation in plants. *Science*, 361(6401), 475–478. <https://doi.org/10.1126/science.aat5760>
- Tsitronis, A., Kirkpatrick, M., & Levin, D. A. (2003). A model for chloroplast capture. *Evolution*, 57(8), 1776–1782. <https://doi.org/10.1111/j.0014-3820.2003.tb00585.x>
- Ungerer, M. C., Baird, S. J., Pan, J., & Rieseberg, L. H. (1998). Rapid hybrid speciation in wild sunflowers. *Proceedings of the National Academy of Sciences of the United States of America*, 95(20), 11757–11762. <https://doi.org/10.1073/pnas.95.20.11757>
- Valbuena-Carabaña, M., González-Martínez, S. C., Hardy, O. J., & Gil, L. (2007). Fine-scale spatial genetic structure in mixed oak stands with different levels of hybridization. *Molecular Ecology*, 16(6), 1207–1219. <https://doi.org/10.1111/j.1365-294X.2007.03231.x>

- Vallejo-Marín, M., & Hiscock, S. J. (2016). Hybridization and hybrid speciation under global change. *New Phytologist*, 211(4), 1170–1187. <https://doi.org/10.1111/nph.14004>
- Van der Auwera, G., & O'Connor, B. (2020). *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra (1st Edition)*. O'Reilly Media.
- Webb, C. J. (1998). The Selection of Pollen and Seed Dispersal in Plants. *Plant Species Biology*, 13(1), 57–67. <https://doi.org/10.1111/j.1442-1984.1998.tb00248.x>
- Whitener, M. R., Mangelson, H., & Sweigart, A. L. (2024). Patterns of genomic variation reveal a single evolutionary origin of the wild allotetraploid *Mimulus sookensis*. *Evolution*, qpae079. <https://doi.org/10.1093/evolut/qpae079>
- Woodsdalek, J. E. (1916). Causes of sterility in the mule. *The Biological Bulletin*, 30(1), 1-[56]-1. <https://doi.org/10.2307/1536434>
- Wolfe, L. M., Blair, A. C., & Penna, B. M. (2007). Does intraspecific hybridization contribute to the evolution of invasiveness?: An experimental test. *Biological Invasions*, 9(5), 515–521. <https://doi.org/10.1007/s10530-006-9046-0>
- Xiong, T., & Mallet, J. (2022). On the impermanence of species: The collapse of genetic incompatibilities in hybridizing populations. *Evolution*, 76(11), 2498–2512. <https://doi.org/10.1111/evo.14626>
- Yan, L.-J., Burgess, K. S., Zheng, W., Tao, Z.-B., Li, D.-Z., & Gao, L.-M. (2019). Incomplete reproductive isolation between *Rhododendron* taxa enables hybrid formation and persistence. *Journal of Integrative Plant Biology*, 61(4), 433–448. <https://doi.org/10.1111/jipb.12718>
- Zuellig, M. P., & Sweigart, A. L. (2018). A two-locus hybrid incompatibility is widespread, polymorphic, and active in natural populations of *Mimulus*. *Evolution*, 72(11), 2394–2405. <https://doi.org/10.1111/evo.13596>

Table 3.1. Mating system estimation.

Cohort*	Year	Maternal		Offspring			Selfing Rate
		Families	Fruits	Outcrossed	Selfed	Ambiguous	
<i>M. guttatus</i> (HI<0.15)	2019	25	27	112	38	13	0.253
	2021	15	27	145	11	17	0.071
	2022	35	35	142	67	6	0.321
Admixed (HI 0.15 - 0.8)	2019	26	46	176	49	17	0.218
	2021	6	12	54	20	5	0.27
	2022	26	29	123	68	1	0.356
<i>M. nasutus</i> (HI>0.95)	2019	8	8	1	24	0	0.96
	2021	0	0	-	-	-	-
	2022	3	3	0	20	0	1

*Determined by maternal hybrid index; includes families from CAC_S1 and CAC_S2

Table 3.2. Measurements of premating reproductive isolation.

Year	Direction* (ovule ← pollen)	Mating system isolation	Phenological isolation	Combined reproductive isolation[†]
2012	G ← N	-	1	-
2012	N ← G	-	1	-
2012	G ← A	-	0.879	-
2012	A ← G	-	0.989	-
2012	N ← A	-	-0.51	-
2012	A ← N	-	0.826	-
2019	G ← N	0.262	1.000	1.000
2019	N ← G	0.960	1.000	1.000
2019	G ← A	0.262	0.983	0.987
2019	A ← G	0.209	0.980	0.984
2019	N ← A	0.960	0.435	0.977
2019	A ← N	0.209	0.747	0.802
2021	G ← N	0.068	-	-
2021	N ← G	-	-	-
2021	G ← A	0.068	-	-
2021	A ← G	0.294	-	-
2021	N ← A	-	-	-
2021	A ← N	0.294	-	-
2022	G ← N	0.318	0.998	0.999
2022	N ← G	1.000	0.980	1.000
2022	G ← A	0.318	0.351	0.557
2022	A ← G	0.405	0.787	0.873
2022	N ← A	1.000	-0.535	1.000
2022	A ← N	0.405	0.956	0.974

Values are scaled from -1 (complete disassortative mating) to 1 (complete assortative mating), with 0 indicating random mating.

*G=*M. guttatus*, A=Admixed, N=*M. nasutus*.

[†] Combined effects of phenological and mating system isolation.

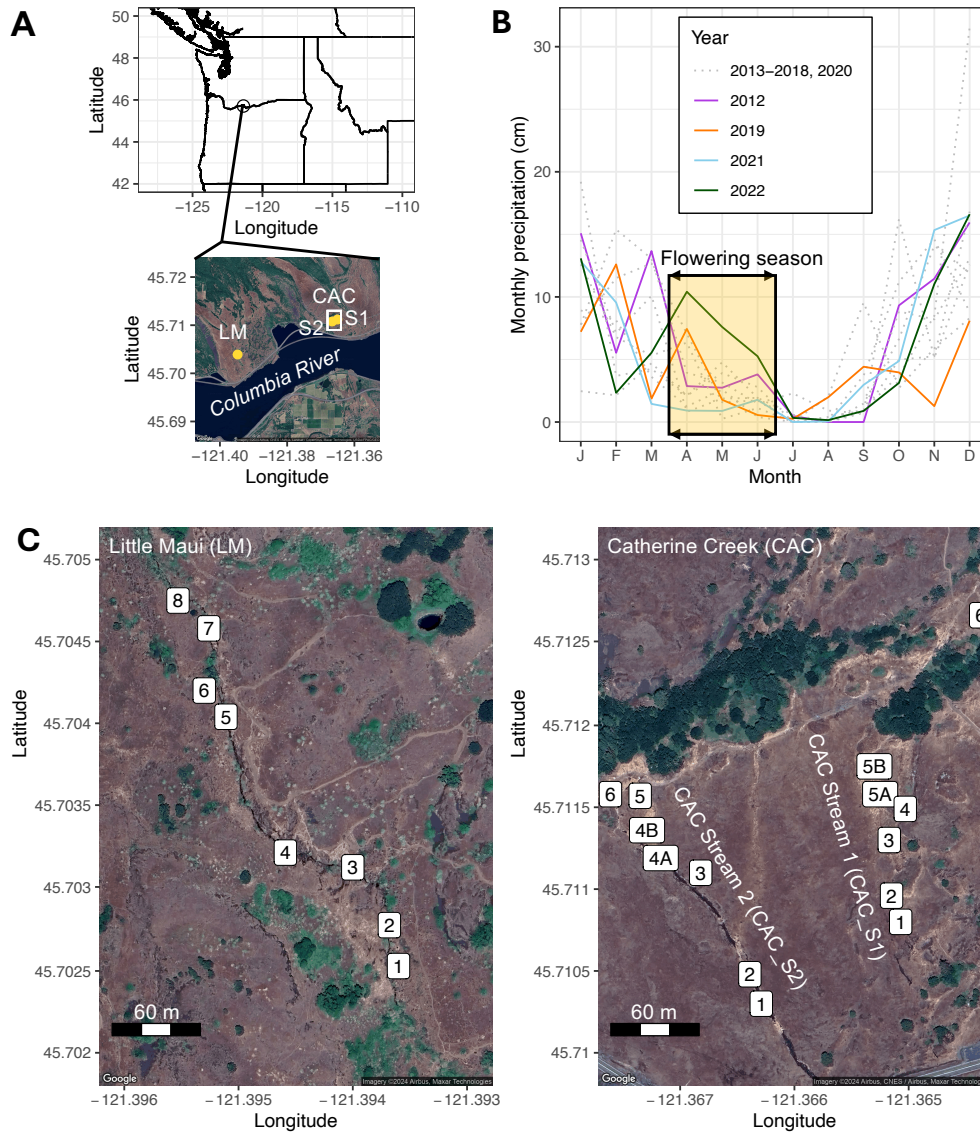


Figure 3.1. Sampling locations and local variation in precipitation. A) Location of three sampling locations in the Columbia River Gorge area, Washington, USA. LM=Little Maui stream, CAC=Catherine Creek site with two parallel streams, S1 and S2. Approximate distance between CAC_S1 and CAC_S2 is ~120m, between CAC and LM is ~4km. B) Interpolated monthly precipitation totals for a 4km grid square covering the CAC and LM sites, from (PRISM Climate Group, 2024). C) Sample plots within each stream. A 0.5mx0.5m square was placed at each plot for flower counts and collections. Exact square placement varied slightly across years, and not all plot locations were sampled in each year. Map imagery from Google, ©2024 Airbus, CNES / Airbus, Landsat / Copernicus, Maxar Technologies, USDA/FPAC/GEO.

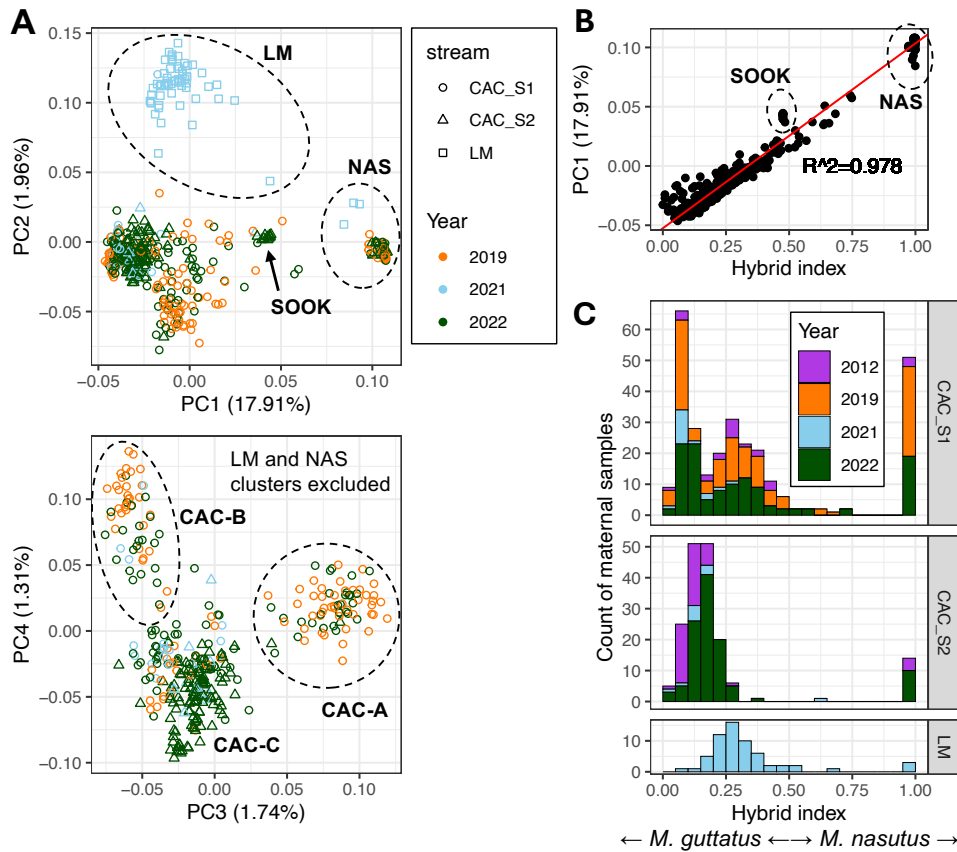


Figure 3.2. Directional admixture shapes population structure in replicate streams. A) Genomic PCA based on genotype likelihoods at 19,633 variant sites. PC1 separates *M. nasutus* samples (NAS) from *M. guttatus* and admixed samples, while PC2 separates Little Maui (LM) samples from Catherine Creek (CAC) samples. PC3 differentiates an admixed group found primarily in plots S1_1 and S1_2 (cluster CAC-A: note that these plots were not sampled in 2021), while PC4 separates a less-admixed group found primarily in plot S1_4 (cluster CAC-B). The remaining cluster (CAC-C) includes individuals from both CAC streams, which are not differentiated by these PC axes. Variation within Catherine Creek is not structured by year. NAS and LM-A clusters were removed in the PC3-PC4 panel in order to show finer population structure. B) PC axis 1 correlates strongly with hybrid index (the proportion of *M. nasutus* genomic ancestry, determined by local ancestry inference). NAS=*M. nasutus* individuals, SOOK=11 individuals of the allopolyploid species *M. sookensis*. C) Hybrid index, the proportion of sites with *M. nasutus* ancestry across the genome for each individual, is distributed differently in each replicate stream, but this pattern is consistent across years. Hybrid index of 0.0 indicates *M. guttatus*, 1.0 indicates *M. nasutus*. Note that LM was only sampled in 2021; CAC_S2 was only sampled in 2012, 2021, and 2022. Data from 2012 is from (Kenney & Sweigart, 2016).

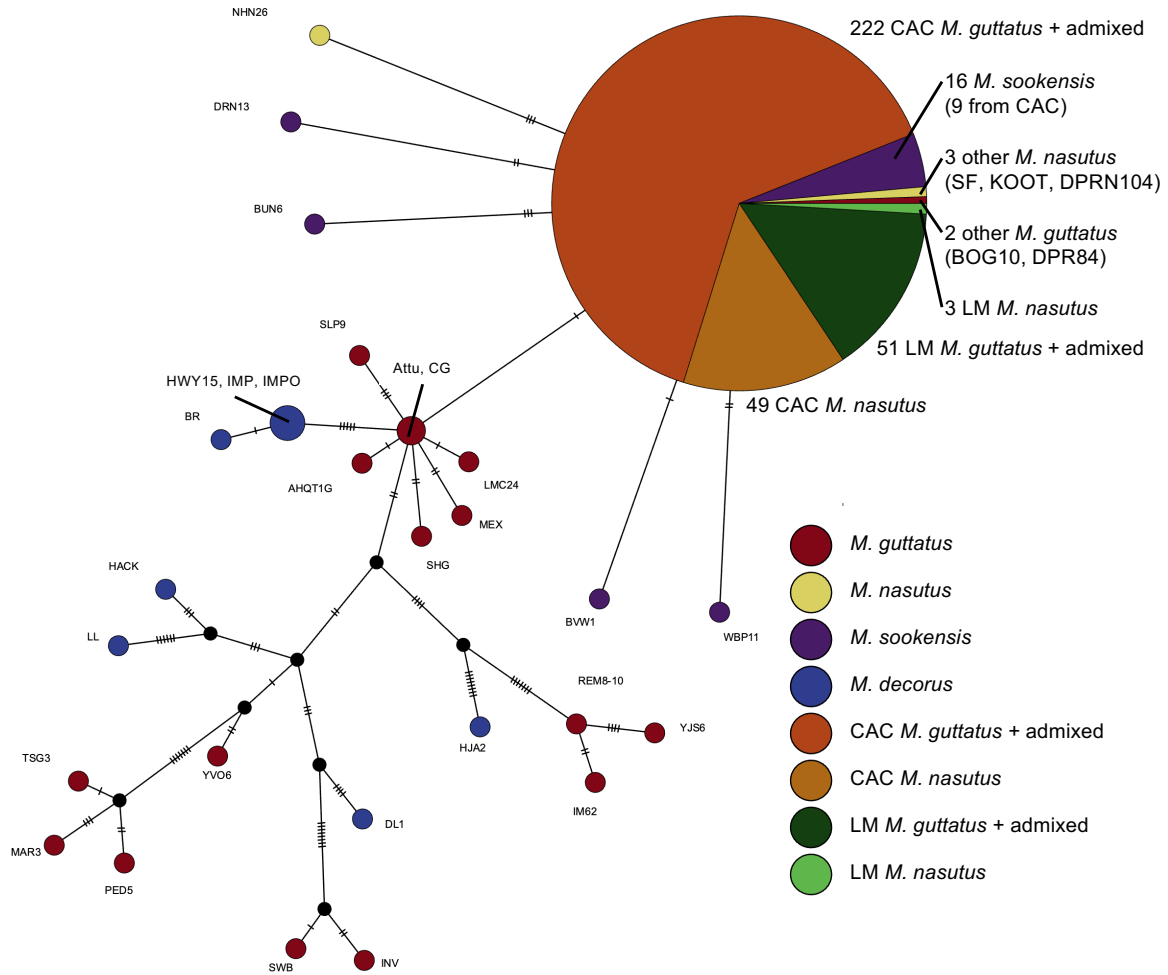
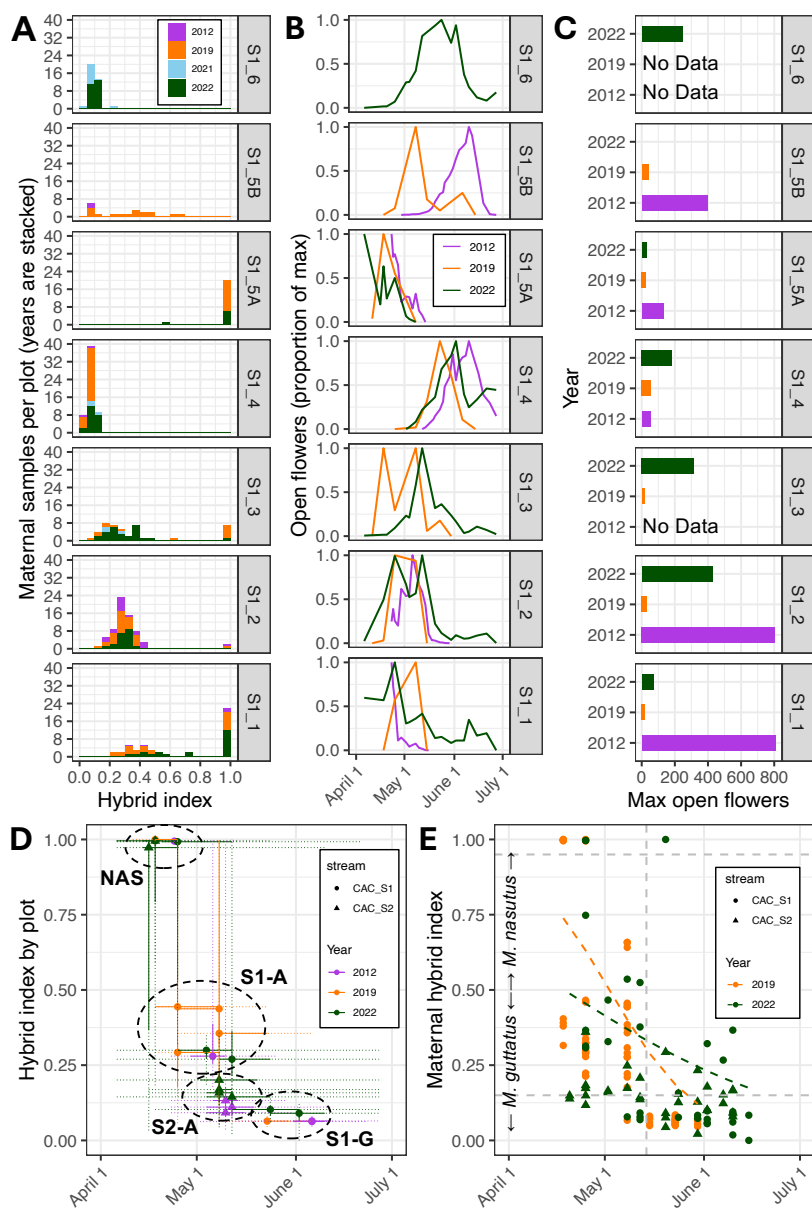


Figure 3.3. Complete chloroplast capture of a *M. nasutus* haplotype in sympatric *M. guttatus*. NJ-Net haplotype network of chloroplast variation built from 280 CAC maternal samples, 54 LM maternal samples, and 41 additional samples from the *M. guttatus* species complex (Whitener et al., 2024), using 102 total variant sites (43 parsimony-informative). A single chloroplast haplotype is present in all maternal samples from Catherine Creek, including *M. guttatus*, admixed, and *M. nasutus* samples. All *M. nasutus* samples from across the range share this haplotype or a close derivative, as do samples from *M. sookensis* (a polyploid with *M. nasutus* as maternal parent). *M. guttatus* haplotypes are more variable, with only two non-CAC samples sharing the *M. nasutus* haplotype, at least one of which (DPR84) is from another sympatric site with known introgression.

Figure 3.4. Fine-scaled spatial and phenological structure across years. A) Histogram of maternal samples sequenced from each plot within CAC_S1, binned by hybrid index, with sampling years in stacked bars. Hybrid index of 0.0 indicates *M. guttatus*, 1.0 indicates *M. nasutus*. Seven plots along CAC_S1 vary in the distribution of *M. guttatus*, admixed, and *M. nasutus* individuals, but plots have a consistent distribution across years. B) Counted open flowers throughout the flowering season in seven plots within CAC_1 during the 2012, 2019, and 2022 seasons, scaled to the maximum number of open flowers counted within that plot in that year. Plots with *M. nasutus* individuals (S1 plots 1-3 and 5A) had early flowers, admixed plots (plots S1_1 to S1_3 and S1_5B) had intermediate peak flowering, and plots with *M. guttatus* (plots S1_4 and S1_6) had later peak flowering. Peak flowering was similar but not identical across years. Tails of open flowers indicate that in 2022, admixed plots (plots S1_1 to S1_3 and S1_5B) continued flowering for much longer than in 2019, resulting in more overlap with *guttatus* plots (S1_4 and S1_6). C) Maximum number of open flowers on any one day for each plot. In 2019, there were far fewer open flowers throughout the growing season than in 2021 or 2022. D) Median flower date for a plot is associated with median hybrid index of that plot. Each point represents the median date out of all counted flowers and the median hybrid index of sequenced individuals from within that plot. Solid lines represent the interquartile range of counted flowers and hybrid indices, while dotted lines represent the full range of flowers and indices for each plot. Circles indicate four main groups: *M. nasutus* plots from both CAC_S1 and CAC_S2, *M. guttatus* plots from CAC_S1, admixed plots from CAC_S1, and admixed plots from CAC_S2. E) Date of marked open flowers is associated with hybrid index of the marked individual. Random sets of flowers were marked when open throughout the season and those individuals were later sampled for sequencing; these are not necessarily the first open flowers for an individual.



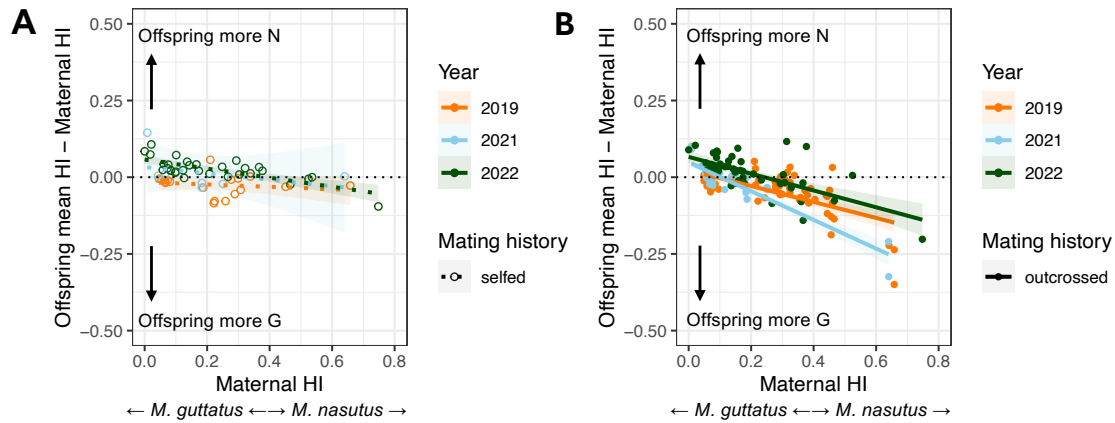


Figure 3.5. Assortative mating agrees with patterns of reproductive isolation. A)

Comparison of offspring hybrid index to maternal hybrid index (HI, the proportion of *M. nasutus* ancestry across the genome) for selfed offspring. Each point represents the mean within a single fruit across all selfed offspring of the difference (Offspring HI – Maternal HI), with positive values indicating greater *M. nasutus* ancestry in offspring relative to maternal plants. Lines indicate individual linear model fits for each year, with shading around each line indicating 95% confidence intervals. The thin black dotted line shows the 1-to-1 null expectation for selfed fruits, where offspring hybrid index is expected to match maternal HI. B) Equivalent plot for outcrossed offspring, showing a larger deviation from maternal ancestry, with the direction of deviation typically in the direction of the population mean.

CHAPTER IV

CORRELATED GENOMIC PATTERNS OF INTROGRESSION ACROSS SPACE DESPITE
CONTRASTING HYBRIDIZATION HISTORIES

¹Farnitano M.C., K. Karoly, V.A. Sotola, and A.L. Sweigart. Correlated genomic patterns of introgression across geographic space despite contrasting hybridization histories. To be submitted to *Molecular Ecology*.

Abstract

Patterns of introgression at genomic resolution have revolutionized our understanding of reproductive isolation and the porosity of species boundaries. Yet, we still know little about the forces that shape these patterns, including how they differ across populations under different ecological, demographic, and reproductive conditions. The genetic mechanisms underlying reproductive isolation have been investigated using both trait mapping approaches and hybrid ancestry patterns, but these methods are rarely compared directly. Through a comparison of geographically disparate contact zones, we investigate the scope of parallelism in patterns of genomic ancestry within the model species pair *Mimulus guttatus* and *Mimulus nasutus*. In addition, we take advantage of known candidate loci and QTL for premating and postmating reproductive barriers between these species to ask how well trait mapping approaches predict patterns of ancestry in wild hybridizing populations. We find that the extent and timing of introgression varies considerably both among and between geographic areas. Parallelism in patterns of genomic ancestry is strongest for geographically proximal groups but remains significant even for populations separated by ~1000km. We find evidence for the selective filtering of introgressed alleles from more admixed to less admixed groups, providing a window into the adaptive introgression process. Finally, we see little correspondence between known candidate loci for premating or postmating isolation and outliers for introgression. We discuss reasons for this lack of correspondence, which argue for the complementarity of different approaches to studying reproductive isolation. Overall, our findings highlight the complexity of factors shaping hybridization patterns.

Introduction

Hybridization between related lineages is a widespread phenomenon across the tree of life (Abbott et al., 2013; Lotsy, 1931; Mallet et al., 2015; Stebbins, 1959). Population-scale genomic sequencing in the last few decades have enabled us to describe hybridization patterns in detail, including how patterns of introgression vary across the genome in a variety of different scenarios (Baack & Rieseberg, 2007; Goulet et al., 2017; Sedghifar et al., 2016; Taylor & Larson, 2019). Some species have small, dispersed segments of introgressed ancestry left over from historical hybridization events (Edelman et al., 2019; Green et al., 2010; Nelson et al., 2021; Randi & Lucchini, 2002). Others form hybrid swarms where two parental genomes form a well-mixed continuum of ancestry (Behm et al., 2010; Hasselman et al., 2014; Kleindorfer et al., 2014; Ruhsam et al., 2011). A common and well-studied phenomenon is tension zones, where hybridization occurs in a narrow zone of contact and segments of introgressed ancestry diffuse outward into allopatric populations (Barton & Hewitt, 1985, 1989; Macholán et al., 2007; Martinsen et al., 2001; Moore & Buchanan, 1985; Streisfeld & Kohn, 2005). Other cases of hybridization follow a more mosaic pattern, in which hybridization occurs in multiple scattered locations across a heterogeneous landscape (Curry, 2015; Howard et al., 1993; Rand & Harrison, 1989; Valbuena-Carabaña et al., 2007). As detecting introgression in genomic data has become commonplace, evolutionary biologists are beginning to shift their focus to uncovering the root causes of variation in introgression, both among populations and among loci across the genome.

Demographic and stochastic forces will generate variation in introgression across individuals and across the genome. Relative population sizes, expansions or bottlenecks, and dispersal rates help determine the relative contribution of two species to admixed genomes (Bourret et al., 2022; Klein et al., 2017; Quilodrán et al., 2020; Valbuena-Carabaña et al., 2007).

Mating dynamics can lead to asymmetries in the direction of introgression (Kenney & Sweigart, 2016; Pickup et al., 2019; Ruhsam et al., 2011; Sianta et al., 2024). The time since hybridization, and whether hybridization happened in a single historical pulse or continuously over time, will affect the size distribution of introgressed ancestry blocks (Menon et al., 2021). The effective population size and levels of inbreeding in the admixed population will influence how often introgressed blocks are shared among individuals (Pickup et al., 2019). Beyond these demographic forces, selection is expected to have strong effects on the distribution of introgression (Sachdeva & Barton, 2018; Sedghifar et al., 2016).

One way that selection can influence introgression patterns in a predictable way is if there is pervasive, polygenic selection against ancestry from the minor parent (Juric et al., 2016; Schumer et al., 2018). This could be due to higher genetic load in that species (Juric et al., 2016; Nouhaud et al., 2022), intrinsic incompatibilities between genomes (Schumer et al., 2018), or local adaptation and coadaptation favoring alleles from the majority parent in its native habitat (Hohenlohe et al., 2012). If this form of selection is acting, we expect to see a pattern of reduced introgression in low-recombination regions, with small blocks of introgression retained in higher-recombination regions. This is because, in regions of low recombination, selection against detrimental alleles will remove longer blocks of introgression before they have a chance to be broken up by recombination. Blocks in high recombination regions may more readily escape linkage with detrimental alleles and persist as small segments of neutral or beneficial minor parent ancestry. This correlation of introgression with recombination rate has been detected in diverse systems (Burri et al., 2015; Kenney & Sweigart, 2016; S. H. Martin et al., 2019; Nouhaud et al., 2022; Schumer et al., 2018). However, these patterns may in some cases be confounded by linked selection operating within species, which can produce higher

differentiation among species in low recombination areas even when no hybridization has occurred (Cruickshank & Hahn, 2014).

In addition to genome-wide patterns, specific loci may be under selection in admixed populations, leading to detectable signals of either reduced or elevated introgression. Areas of reduced introgression have been associated with genes controlling intrinsic incompatibilities (Moran et al., 2024; Powell et al., 2020; Runemark et al., 2018). Locally adapted alleles could also resist introgression (Jones et al., 2012; Lowry & Willis, 2010), though we have fewer examples of candidate genes for these traits. In contrast, areas of elevated minor parent ancestry might indicate cases of adaptive introgression (Chhatre et al., 2018; Menon et al., 2021; Pardo-Diaz et al., 2012; Racimo et al., 2015), in which the minor parent allele is favored either globally or in a specific admixed environment. Expectations involving premating reproductive barrier loci are less clear: in theory, assortative mating loci should resist introgression (Streisfeld & Kohn, 2005; Wessinger et al., 2023), but mating-related loci are also documented crossing species boundaries and fixing in new lineages (Nelson et al., 2021; Stankowski & Streisfeld, 2015).

Despite extensive theoretical work predicting patterns of introgression, and an increasing number of systems with genomic ancestry data, a number of questions remain. One unresolved question is how replicable these genomic ancestry patterns are when hybridization occurs multiple times. A few studies have shown remarkable similarity in the genomic distribution of introgression across independent cases of hybridization of the same species pair (Chaturvedi et al., 2020; Nouhaud et al., 2022). Species pairs with higher divergence prior to hybridization appear to have more repeatable patterns in secondary contact than more closely related species pairs, likely due to the presence of more strong, fixed incompatibility loci with predictable effects (Langdon et al., 2024). Variation across space can be driven by demographic factors such

as population expansion or invasion (Quilodr  n et al., 2020), or distance from the center of a contact zone (Simon et al., 2021). Ecological factors likely result in differences as well (Brice et al., 2021).

We might expect that loci identified in mapping studies as important for reproductive isolating barriers would have predictable responses in ancestry frequencies within wild populations. However, there are few direct tests of the correspondence between these two types of data. Partly, this is because there are few study systems for which both types of data exist for the same organisms. Where differentiation in hybrid zones has been associated with known adaptive or barrier loci, it is usually in the context of chromosomal inversions that suppress recombination (Dean et al., 2024; Machado et al., 2007) or genes on sex chromosomes (Maroja et al., 2009; Trier et al., 2014). A recent analysis of the *Mus musculus musculus* \times *M. musculus domesticus* hybrid zone and corresponding laboratory crosses found that wild and laboratory studies identified distinct and rarely overlapping sets of loci (Frayer & Payseur, 2024). A multispecies analysis in cichlids also found low correspondence between different methods of detecting putative incompatibilities (Feller et al., 2024).

The *Mimulus guttatus* \times *M. nasutus* study system is ideal for addressing broad questions about the nature of hybridization and introgression. Hybridization has been detected in multiple distinct geographic areas (Brandvain et al., 2014; Kenney & Sweigart, 2016; Zuellig & Sweigart, 2018a), allowing for comparative tests of genomic introgression patterns across these replicates. Furthermore, multiple premating and postmating reproductive incompatibilities have been mapped to causal genes (Zuellig & Sweigart, 2018b) or quantitative trait loci (Fishman et al., 2014; Mantel & Sweigart, 2024), setting us up to test the correspondence between introgression patterns in hybrid populations and mapped reproductive isolation traits. In this study, we take

advantage of sequencing data generated in Chapter III from many hybrid individuals in the CAC and LM admixed populations in Washington, USA. We combine those data with new data from a second, completely independent area of hybridization, near the Don Pedro Reservoir in north-central California, USA. By comparing patterns of genomic ancestry within and among these distinct areas, we are able to address the repeatability of the processes generating introgression patterns across populations of the same species. We ask whether either broad genome-wide patterns of admixture, or specific patterns at key loci, are shared among populations separated by geographic space. Then, by comparing our data with known candidate loci and mapped QTL, we ask whether known reproductive isolation loci are predictive of these introgression patterns. Each of these questions is key to our understanding of the selective processes occurring during and after hybridization, and our answers have implications for introgression across the tree of life.

Materials and Methods

Study sites and sample collection

For this study, we sequenced a total of 949 individuals across two study areas: the Catherine Creek (CAC) and Little Maui (LM) locations near the Columbia River in Washington, USA (northern area), detailed in Chapter III, and the Don Pedro Reservoir (DPR) and nearby locations (southern area) in the foothills of the Sierra Nevada mountains in California, USA (Figure 4.1A, Table S4.1). We used 510 previously sequenced wild-growing individuals from Catherine Creek streams 1 and 2, collected across 2019, 2021, and 2022, plus 64 previously sequenced wild-growing individuals from Little Maui collected in 2021, for a total of 574 individuals from the northern area; sequencing details and population structure of these

individuals are described in Chapter III. In addition, we collected 117 wild-growing samples from the main DPR location, plus seeds from 258 wild maternal plants growing throughout the surrounding area (Figure 4.1B). Of these 258 seed collections, 37 were from higher-elevation (>1000 m) locations, 152 from other sympatric locations (i.e., we collected at least one *M. guttatus* and one *M. nasutus* at that location), and 62 samples from allopatric sites where only *M. guttatus* was sampled (Figure 4.1B). Collection locations and information are provided in Tables S4.1-S4.2.

Southern area samples were collected in the field from April to June of 2021. At the DPR main site, plants were marked throughout the growing season in a manner analogous to CAC sampling, in order to capture the range of phenological and ancestry variation at the site. Samples from the surrounding area were collected in single or a few visits to each location during the DPR growing season (Table S4.1) and may not be representative of the full range of seasonal variation at these locations. We attempted to sample both *M. nasutus* and *M. guttatus* and any possible hybrids or admixed individuals, including sampling both senesced and actively flowering individuals where available, but sampling was not exhaustive and may have been biased by the date(s) of sampling for any particular location.

Tissue from wild-growing DPR individuals was dried with silica gel beads in the field, then flash-frozen in liquid nitrogen at the time of DNA extraction. For wild-collected seeds, a subset of seeds from each maternal family were germinated in the UGA Botany greenhouses, where bud or leaf tissue was collected from one germinant per maternal family and stored at -80C, then flash-frozen in liquid nitrogen at the time of DNA extraction. For all samples, DNA was extracted using a CTAB protocol with phenol-chloroform extraction (Fishman, 2020). Samples were prepped for low-coverage whole-genome short-read sequencing using the same

approach followed for CAC and LM samples (see Chapter III), i.e., a tagmentation library prep (Farnitano & Sweigart, 2024) with bulk-prepped Tn5 tagmentation enzyme (Lu et al., 2017). Samples were sequenced on an Illumina NovaSeq 6000 machine at the Duke University Sequencing and Genomic Technologies Core Facility.

Local ancestry assignment

We conducted local ancestry inference to estimate *M. guttatus* vs. *M. nasutus* ancestry across the genome in each sample from the Columbia River and Don Pedro Reservoir areas, following the approach detailed in Chapter III, which uses the *ancestryinfer* pipeline and *ancestry_HMM* program (Corbett-Detig & Nielsen, 2017; Schumer et al., 2020). Briefly, samples were aligned to the *M. guttatus* IM62v3 reference genome and the *M. nasutus* SFv2 reference genomes (<https://phytozome-next.jgi.doe.gov>), keeping only reads that aligned exactly once to each genome. We then extracted read counts per sample at each of 208,560 ancestry-informative markers with large allele frequency differences between *M. nasutus* and allopatric *M. guttatus* reference lines (see Chapter III for creation of ancestry-informative marker panel). Ancestry calls at each marker were obtained with *ancestry_HMM*, and we used a posterior probability cutoff of 0.9 to assign genotypes at each site, with lower probabilities set to missing. Samples with fewer than 25,000 called ancestry-informative markers were excluded from the analysis, leaving a total of 877 individuals across both sample areas (92% of sequenced individuals; at least 66% of individuals from each collection location, Table S4.2). We calculated hybrid index (HI, the number of called sites with *M. nasutus* ancestry plus half the number of heterozygous sites, divided by the total number of called sites) and ancestry heterozygosity (the proportion of heterozygous sites out of all called sites) for each individual.

Genomic PCA analysis

To assess population structure and variation across northern and southern samples, we generated a SNP marker panel by starting with a previously obtained list of 3,493,514 SNPs polymorphic within a set of 36 diverse *M. guttatus* and *M. nasutus* lines (details in Chapter III), then randomly downsampling to one marker per 10kb, resulting in a panel of 17,514 SNP markers. We used *angsd* v0.940 (Korneliussen et al., 2014) with the GATK likelihood method to estimate genotype likelihoods at each marker for each of the northern and southern individuals. We used these likelihoods to conduct a principal components analysis (PCA) in *PCAngsd* v1.10 (Meisner & Albrechtsen, 2018).

Assessing parallelism in local ancestry

To assess parallelism across the genome in local ancestry, we divided samples in eight sample groups based on geography and hybrid index: 1) DPR area allopatric *M. guttatus* (n=61 samples); 2) DPR area sympatric *M. guttatus* (n=191); 3) CAC *M. guttatus* (n=184); 4) CAC *admixed* (n=213); 5) LM *admixed* (n=56); 6) DPR area sympatric *admixed* (n=22); 7) DPR area sympatric *M. nasutus* (n=75); and 8) CAC+LM *M. nasutus* (n=68). We used hybrid indices of less than 0.15 to define *M. guttatus*, 0.15-0.85 for *admixed*, and greater than 0.85 to define *M. nasutus*. DPR sympatric groups included any lower- or higher-elevation southern location where both *M. guttatus* and *M. nasutus* or *admixed* individuals were sampled, but we excluded high elevation samples from the *M. nasutus* group because of concerns about their comparability (5 individuals). Northern (CAC and LM) *M. nasutus* samples were combined due to the small sample size at LM (3 individuals), and the small number of LM *M. guttatus* samples (2 individuals) were not used.

For each sample group, we used ancestry calls from local ancestry inference to calculate the ancestry frequency at each of the 208,560 ancestry-informative markers (the proportion of called samples with *M. nasutus* ancestry at that site, treating heterozygous ancestry as half). For these calculations, only samples with a called genotype were included at each marker; any marker with less than 50% of samples called in a given sample group was set to missing for that group (647 to 8105 markers removed per sample group, or 0.3-3.9% of markers).

To reduce the noise associated with uncertainty or genotype errors at individual markers, we binned each chromosome into 50kb windows and calculated ancestry frequencies for each group within each window. We used the weighted average of ancestry frequencies across all markers in the window, weighted by the number of called samples at each marker. For this analysis, markers where fewer than 50% of samples in the group had ancestry calls were still included in the weighted average. Then, for each sample group, we retained windows with at least 50% of samples called in at least five markers, setting any window below this threshold to missing data (655 to 1370 windows removed per sample group, or 15-31% of windows).

Correlation between ancestry frequencies and analysis of variance

To determine the correlation between ancestry frequency across groups, we excluded markers where at least one group had a missing ancestry frequency, leaving 179,983 markers (86% of the full marker set), then used the *cor* function in the R package ‘stats’ to obtain correlation (r) values. We also determined the correlation between ancestry frequency using 50kb windows, excluding any windows where at least one group was set to missing; 2,739 windows were included in this analysis (63% of possible windows, covering ~137 Mb or 40.4% of the reference genome length).

We tested whether recombination rate and gene density were correlated with ancestry frequency. We used recombination rates in 50kb windows from a linkage map estimated using recombinant inbred lines derived from three independent CAC *M. guttatus* x CAC *M. nasutus* crosses (Mantel & Sweigart, 2024). We also used the gene annotation from the *M. guttatus* IM62v3 reference genome (<https://phytozome-next.jgi.doe.gov>) to calculate gene density in 50kb windows, defined as the number of annotated genes with start sites within each window. Correlation (r) values were obtained as above between ancestry frequency of each sample group and gene density and recombination rate.

We next tested whether correlations between ancestry frequencies of different sample groups remained after accounting for their covariation with gene density and recombination rate. For each pair of sample groups reciprocally, we ran a linear model on all 50kb windows without missing data using the *lm* function in the R package ‘stats’ with the following formula, setting ancestry frequency of one sample group to be the dependent variable and ancestry frequency of the second sample group to be a predictor:

$$(Dependent\ variable) \sim (Gene\ density) + (Recombination\ rate) + (Predictor\ group)$$

We then used the *anova* function in the R package ‘stats’ to estimate the percent variance in the dependent variable explained by gene density, recombination rate, the predictor group, and residual variance. Note that, due to the ordering of variables in the model, only variance that is independent of gene density and recombination rate will be included in the ‘predictor group’ variance.

Selecting ancestry frequency outliers

We identified genomic regions that were outliers for *M. nasutus* frequency in each sample group, using two complementary approaches. First, we calculated the 5th and 95th

percentile of *M. nasutus* ancestry across 50kb windows for each sample group, excluding any windows set to missing as above. We chose 50kb windows at or below the 5th percentile of *M. nasutus* ancestry within each sample group (“low-*M. nasutus* outlier windows”), and windows at or above the 95th percentile of *M. nasutus* ancestry (“high-*M. nasutus* outlier windows”) (Table S4.3). As a complementary approach to outlier assignment, we calculated the 5th and 95th percentiles of ancestry frequency across individual ancestry-informative markers within each sample group, excluding markers with genotype calls in <50% of the samples in that group. We then identified unbroken stretches of at least 10 markers at or below the 5th percentile (“low-*M. nasutus* outlier segments”) or unbroken stretches of at least 10 markers at or above the 95th percentile (“high-*M. nasutus* outlier segments”) with no other non-missing markers between them; markers set to missing were allowed inside a segment (Table S4.4).

We tested whether “outlier segments” and “outlier windows” captured similar regions of the genome. Between 80.6% and 99.8% of high-*M. nasutus* outlier windows, depending on sample group, contained high-*M. nasutus* outlier segments (Table S4.5). In the reverse direction, 77.3 to 99.3% of non-missing windows containing high-*M. nasutus* outlier segments were also called as outlier windows (Table S4.5). For low-*M. nasutus* outliers, 86.7% to 99.4% of outlier windows contained outlier segments, while 56.4% to 91.8% of non-missing windows containing outlier segments were also called as outlier windows (Table S4.5).

Overlapping ancestry frequency outliers

We investigated whether ancestry outliers were shared between sample groups more often than expected by chance. First, we determined the extent of overlap across sample groups of outlier 50kb windows, compared to a null expectation of random outlier placement. For each sample group, we created 100 permutations of the low-*M. nasutus* outlier windows, and 100

permutations of the high-*M. nasutus* outlier window. Each permutation was created by randomly selecting X windows, with replacement, from the complete list of non-missing windows, where X is the number of true outlier windows. Then, for each pair of sample groups, we ran the following analysis, separately for low-*M. nasutus* windows and high-*M. nasutus* windows. First, we calculated the number of outlier windows that were shared between groups (the true overlap). Second, we calculated the number of outlier windows from one group that were shared with each of the 100 random permutations of the other group (the permuted overlaps). We gave each comparison a Z-score comparing the true overlap to the mean and standard deviation of the permuted overlaps, calculated as follows:

$$Z = \frac{(\text{True overlap}) - \text{Mean}(\text{Permuted overlap})}{\sqrt{\text{Variance}(\text{Permuted overlap})}}$$

This z-score represents the degree to which the true overlap in outlier windows between groups exceeds the expected overlap if outlier windows were randomly distributed. We also calculated, for low-*M. nasutus* and high-*M. nasutus* outliers, the number of 50kb outlier windows that overlapped across each combination of 2 or more sample groups, versus the number of 50kb outlier windows that were unique to one sample group.

We also determined the extent to which ancestry outlier segments overlapped across sample groups, compared to a null expectation of random placement of these segments. For each sample group, we generated 100 random permutations of the low-*M. nasutus* outlier segments, and 100 random permutations of the high-*M. nasutus* outlier segments. To generate these random permutations, we took each true outlier segment and, retaining its original length (in number of called sites), we moved it to a random start position chosen from all non-missing markers; if the new start position would cause the segment to extend past the last called site in a chromosome, we chose a new start position. We followed this procedure for each outlier segment, creating a

list of new segments with the same length distribution as the true outlier segments; this list represents one permutation. After obtaining these permutations for each sample group, we ran the following analysis on each pair of sample groups, separately for low-*M. nasutus* and high-*M. nasutus* outliers. First, we calculated the total length, in base pairs, of overlap between true outlier segments of the two groups, divided by the total length of the outlier segments in the first group (true overlap proportion). Then, for each permutation of the second group, we calculated the length of overlap, in base pairs, between the true outlier segments of the first group and the permuted outlier segments of the second group, divided by the total length of the outlier segments in the first group (permuted overlap proportions). Then, we gave each true overlap proportion a z-score, calculated as above.

For all overlap comparisons of high-*M. nasutus* outlier windows and segments, we did not include the two *M. nasutus* sample groups (from CAC + LM and from DPR), because the distribution of ancestry in these groups was one-inflated: 65% to 88% of windows have an ancestry frequency of 1.0; marker segments with an ancestry frequency of 1.0 likewise cover 69% to 85% of the reference genome (Table S4.3-S4.4). Similarly, for all overlap comparisons of low-*M. nasutus* outlier segments and windows, we did not include DPR allopatric *M. guttatus*, because the distribution of ancestry was zero-inflated: 37% of windows have an ancestry frequency of 0.0; marker segments with an ancestry frequency of 0.0 likewise cover 51% of the reference genome (Table S4.3-S4.4).

Extent of overlap between ancestry outliers and known QTL regions.

We tested whether genomic regions associated with known premating or postmating reproductive isolation barriers followed predictable patterns of ancestry within each sample group. For premating barrier loci, we focused on two major quantitative trait loci (QTL)

affecting photoperiod divergence between *M. guttatus* and *M. nasutus*. These QTL, on chromosomes 07 and 08, were identified previously in a cross involving lines from the CAC site, as well as in another interspecific cross; they have been fine-mapped to a single strong candidate gene at the chromosome 07 locus, and a small number of candidates at the chromosome 08 locus (Fishman et al., 2014). We identified the two 50kb windows containing these candidate genes using the *M. guttatus* IM62v3 reference genome annotation (<https://phytozome-next.jgi.doe.gov>) and examined the ancestry frequency of these windows in each of the eight sample groups, as well as whether they were classified in high-*M. nasutus* or low-*M. nasutus* outlier windows for any sample group.

For postmating reproductive isolation, we first looked at a pair of genetic loci which interact to produce hybrid lethality in certain crosses between *M. guttatus* and *M. nasutus*. These loci, *hll3* on chromosome 13 and *hll4* on chromosome 14, were identified in crosses involving lines from the DPR area, and have been shown to be polymorphic and to cross species boundaries (Zuellig & Sweigart, 2018a); they have been mapped to single paralogous genes at each locus (Zuellig & Sweigart, 2018b). We identified the two 50kb windows containing these genes using the *M. guttatus* IM62v3 reference genome annotation (<https://phytozome-next.jgi.doe.gov>) and examined the ancestry frequency of these windows in our eight sample groups, as well as whether they were classified in our high-*M. nasutus* or low-*M. nasutus* outlier windows for any sample groups.

Next, we looked at patterns of ancestry at a larger set of potential reproductive incompatibility loci. We selected regions from a recent mapping study of *M. guttatus* x *M. nasutus* recombinant inbred lines derived from Catherine Creek samples (Mantel & Sweigart, 2024), which used three separate mapping populations to identify quantitative trait loci (QTL)

associated with male and female sterility, as well as regions of transmission ratio distortion. We included nine QTL where the *M. guttatus* allele was associated with reduced male or female fertility, and two QTL where the *M. nasutus* allele was associated with reduced male or female fertility. A final fertility QTL was excluded because its predicted effect was ambiguous – there was a main effect of the *M. guttatus* allele reducing fertility, but a multilocus interaction in which the *M. nasutus* allele was associated with reduced fertility. In addition, we selected twelve loci (six pairs) with significant multilocus transmission ratio distortion (TRD), in which an NN/GG two-locus genotype combination was underrepresented. These 23 loci are potential candidates for postmating reproductive incompatibilities that may be under selection in admixed populations. We divided these loci into two groups based on whether we expect the *M. guttatus* or *M. nasutus* allele to be selected against in admixed population: for fertility, whichever allele was associated with lower fertility, and for the TRD loci, whichever allele was in the underrepresented two-locus genotype. We refer to these sets as the N- loci (*M. nasutus* allele involved in the trait or interaction; lower expected *M. nasutus* ancestry frequency) and the G- loci (*M. guttatus* allele involved in the trait or interaction; higher expected *M. nasutus* ancestry frequency). For each fertility QTL, we used the 1.5-LOD-drop region to define the ends of the QTL region; for each TRD locus, we used the 95% confidence LD region to define the ends of the region.

For each of the N- loci and G- loci, we obtained the median and distribution of ancestry frequency across ancestry sites contained within the boundaries of the locus, separately for each sample group. We compared these distributions with the median, 5th, and 95th percentile of ancestry frequency genome-wide. Then, to directly test whether these loci appeared as outliers for ancestry frequency in our data, we calculated for each sample group the length in bp of

overlap between the low-*M. nasutus* or high-*M. nasutus* outlier segments and the set of N- or G- loci. Using an approach analogous to our previous outlier overlap permutations tests, we determined the length of overlap between each of 100 random permutations of the outlier segments and the N- or G- locus sets. We used the mean and standard deviation of these permutation overlaps to calculate a z-score for the extent of overlap between outlier regions and N- or G- loci compared to an expectation of randomly placed outlier regions.

Statistical analyses

All statistical analyses were conducted using R version 4.3.1 (R Core Team, 2023) and the tidyverse packages, v2.0.0 (Wickham & RStudio, 2023).

Results

Dramatic differences in admixture patterns across geographic areas

We set out to make comparisons between contact zones in the Columbia River Gorge, WA (CAC+LM, northern populations) and a different set of contact zones in the Sierra Nevada foothills and mountains, CA (DPR area, southern populations) (Figure 4.1A-B). Two primary axes of genome-wide variation delineate northern from southern areas (PC1, 21.2% of variation, Figure 4.1C) and, within both areas, a continuum of ancestry from *M. guttatus* to *M. nasutus* (PC2, 18.8% of variation, Figure 4.1C). While variation between northern and southern *M. guttatus* is much greater, northern and southern *M. nasutus* are still cleanly defined by PC1 (Figure 4.1C). Within the southern area, allopatric populations have low admixture, (mean HI 0.01, Figure 4.1D), although a few samples do have detectable *M. nasutus* ancestry (two of 61 individuals have HI>0.05, max HI=0.09). Surprisingly, southern sympatric populations showed very different patterns of admixture compared to northern populations – both at a regional scale

and at individual locations (Figure 4.1B-D). At lower elevations, sympatric southern populations typically have much smaller average admixture proportions than our highly admixed northern sites (Figure 4.1B-C): only 3 of 260 individuals have HI between 0.15 and 0.85, compared to 269 of 521 northern individuals (Figure 4.1D, Table S4.2). At higher elevations (>1000m), however, we see much higher admixture proportions (19 of 33 individuals have HI between 0.15 and 0.85, Table S4.2). In fact, we see much more recent admixture at high elevation southern populations than at either of the northern populations: we sampled 7 putative first-generation hybrids at high elevation (ancestry heterozygosity >0.75 and HI 0.4-0.6) compared to only one at the northern locations (Figures 4.1C-D and S4.1, northern data in Chapter III). At these high elevation populations, admixture is also more symmetrical, with both majority-*M. nasutus* and majority-*M. guttatus* admixed individuals, in contrast with the highly asymmetrical nature of admixture in northern and low-elevation southern populations (Figures 4.1D, S4.1).

Patterns of variation in ancestry across the genome

To facilitate contrasts between geographic areas, we divided samples into eight groups based on geography and ancestry proportions (Figure 4.2A, Methods). Ancestry frequencies are highly heterogeneous across the genome within each group (Figure 4.2A-B, additional chromosomes in Figures 4.5, S4.2, and S4.3). Within *M. nasutus* samples, a few genomic regions with putative *M. guttatus* introgression are apparent, though *M. nasutus* ancestry frequency never reaches below 0.75 and the majority of windows have no *M. guttatus* ancestry. In *M. guttatus* groups, small peaks of elevated *M. nasutus* introgression are interspersed among longer stretches of near-zero *M. nasutus* frequency; these peaks are most apparent in CAC *M. guttatus* but can be found in both southern sympatric and southern allopatric *M. guttatus* as well. Within northern admixed groups, *M. nasutus* frequency is more variable, but still often features long stretches of

low *M. nasutus* frequency punctuated by shorter regions of high *M. nasutus* introgression. The southern *admixed* group, which has the smallest sample size and the most recent hybridization, also had the most stochastic ancestry frequencies, with less defined peaks and troughs. We see no evidence of introgression that has swept to fixation (or near-fixation) in any sample group, though such sweeps may be missed if they have reached high frequency in the allopatric *M. guttatus* we used to define ancestry markers.

Ancestry frequencies are correlated across the genome between all *M. guttatus* and *admixed* groups, both for 50kb windows (Figure 4.3A) and for individual markers (Figure S4.4). One potential reason for this correlation might be shared patterns of polygenic linked selection against *M. nasutus* ancestry. We tested for this effect by measuring the correlation between ancestry frequency (in 50kb windows) and both recombination rate and gene density as proxies for the strength of linked selection. Gene density and recombination rate are correlated with each other ($r=0.114$, Table 4.1), matching previous results in *Mimulus* (Hellsten et al., 2013; Kenney & Sweigart, 2016); our correlation may be weaker due to lack of window coverage in the regions of lowest gene density. Consistent with the effects of linked selection, ancestry frequencies in *admixed* and sympatric *M. guttatus* groups are positively correlated with gene density ($r=0.284$ - 0.356 , Table 4.1); this correlation is weaker but still detectable for *M. nasutus* and allopatric *M. guttatus* groups ($r=0.083$ - 0.129 , Table 4.1). Recombination rate is also correlated with ancestry frequency in *admixed* and sympatric *M. guttatus* groups, though the relationship was weaker than for gene density ($r=0.061$ to 0.087 , Table 4.1). Despite these strong associations, patterns of gene density and recombination rate do not completely account for the correlations between ancestry frequencies of different sample groups (Figure 4.3B), suggesting that other factors such as shared

migration history or selective effects at particular alleles also play a role in shared ancestry patterns.

Shared ancestry patterns within geographic areas

Similarities between ancestry patterns in higher-admixture vs. lower-admixture sample groups might indicate a sharing of introgressed alleles via gene flow, or similar selection pressures acting in parallel after hybridization. Within geographic areas, ancestry frequencies in both windows and individuals markers are strongly correlated between higher-admixture and lower-admixture groups (Figures 4.3A-B and S4.4, purple and blue outlines), suggesting connectivity across these groups within an area. The strongest ancestry correlation occurs between CAC *admixed* and CAC *M. guttatus* groups ($r=0.81$, Figure 4.3A; variance explained 53.0-53.6%, Figure 4.3B), which we know are incompletely isolated by premating barriers (Chapter III). CAC and LM, two populations separated by ~4km, are also strongly correlated ($r=0.63$, Figure 4.3A; variance explained 30.1-31.1%, Figure 4.3B), indicating either gene flow of introgressed *M. nasutus* alleles between these populations or a shared genetic basis for selection on ancestry. Correlations are also strong between southern sympatric *M. guttatus* and southern allopatric *M. guttatus* ($r=0.53$, Figure 4.3A; variance explained 23.0-25.4%, Figure 4.3B), arguing for gene flow from sympatry to allopatry leading to the spread of *M. nasutus* alleles. These correlations are weaker between DPR area *admixed* populations and other DPR groups ($r=0.15-0.35$, Figure 4.3A; variance explained 1.4-6.6%, Figure 4.3B), likely because many of these admixed individuals are more recent hybrids and are from high elevations, geographically isolated from the other sampled sympatric and allopatric locations.

If blocks of *M. nasutus* introgression are shared within and among nearby populations via migration or parallel selection, we should expect outliers for high *M. nasutus* ancestry to co-

occur in the same genomic regions in different sample groups. We do in fact see this pattern: between all CAC and LM *M. guttatus* and *admixed* groups, high-*M. nasutus* outliers are more often overlapping than randomly permuted outliers, whether we use outlier windows ($z=15.15-32.34$, Figure 4.4A, dashed box) or outlier segments ($z=4.85-15.89$, Figure S4.5A, dashed box). Similarly, for comparisons between southern sympatric and allopatric *M. guttatus*, both outlier windows ($z=12.99-17.04$, Figure 4.4A, dotted box) and outlier segments ($z=6.31-8.34$, Figure S4.5A, dotted box) overlap across sample groups more often than randomly permuted outliers. These shared regions are potential candidates for adaptive introgression within each geographic area.

Genomic regions with low *M. nasutus* ancestry, in contrast, might represent genetic incompatibilities or other traits resistant to introgression; sharing of outliers for low *M. nasutus* ancestry would indicate similar genetic architecture for these incompatibilities across sample groups. Across sample groups within the northern area, sharing of low-*M. nasutus* outlier windows (Figure 4.3B, dashed box) or outlier segments (Figure S4.5B, dashed box) is weaker than for high-*M. nasutus* outliers, but is still significant for some comparisons. The weaker signal could indicate a more diffuse pattern of genome-wide negative selection that does not reliably highlight particular loci, or it could indicate a highly polymorphic genetic basis for incompatibility, with the relevant loci not always shared even between nearby populations.

Shared ancestry patterns between distant geographic areas

Even for comparisons across northern and southern populations (~1000km distance), ancestry frequencies are correlated across genomic windows and markers (Figures 4.3A and S4.4, red and yellow outlines) for *M. guttatus* and *admixed* sample groups. Once again, these correlations remain after accounting for gene density and recombination rate, though they

explain a smaller percentage of variance than within-area comparisons (variance explained 1.4 to 12.2% between geographic areas, red and yellow outlines, vs. 1.4 to 53.6% within geographic areas, blue and purple outlines, Figure 4.3B). Migration is unlikely to contribute to these cross-area correlations, so they implicate shared genetic factors and selective pressures operating in parallel in these two geographic areas.

Correlations in ancestry frequency between northern and southern populations could be due to diffuse global patterns, sharing of high-*M. nasutus* outliers (candidates for parallel adaptive introgression), sharing of low-*M. nasutus* outliers (candidates for shared incompatibilities), or some combination of the above. We find that sharing of high-*M. nasutus* outlier windows is significant for all northern vs. southern comparisons of *M. guttatus* and *admixed* groups ($z=3.82-11.87$, Figure 4.4A); many but not all of these comparisons have significant overlap in high-*M. nasutus* outlier segments as well ($z=1.72-4.72$, Figure S4.5A). Sharing of low-*M. nasutus* outlier windows ($z=1.44-13.48$, Figure 4.4B) or outlier segments ($z=1.46-10.28$, Figure S4.5B) is also significant for many, but not all, southern sympatric *M. guttatus* vs. northern population comparisons, but is generally weaker than for high-*M. nasutus* outliers. These patterns suggest substantial parallelism in the selective forces driving introgression patterns in these two distinct areas.

Potential introgression of M. guttatus ancestry into M. nasutus

Overall, we see little correlation in ancestry frequency between *M. nasutus* and other sample groups (Figure 4.3A-B), consistent with a lack of substantial introgression from *M. guttatus* into *M. nasutus*. However, ancestry frequency within southern sympatric *M. nasutus* ancestry does correlate slightly with ancestry in the southern *admixed* group across genomic windows or markers ($r=0.18$, Figures 4.3A and S4.4), suggesting some alleles may be filtering

into *M. nasutus* through occasional outcrossing with admixed individuals. This pattern matches our finding of more symmetrical hybridization patterns in the high elevation southern populations. In support of this possibility, we find that outlier windows for low *M. nasutus* ancestry (i.e., nonzero *M. guttatus* ancestry) within both *M. nasutus* groups do overlap more often than chance with low-*M. nasutus* outliers in the southern sympatric *admixed* group ($z=2.92-12.37$, Figure 4.4B), though these comparisons are not significant in our outlier segment analysis ($z=-0.15$ to 1.77 , Figure S4.5B). The fact that northern and southern *M. nasutus* show a similar pattern of sharing with the southern *admixed* group suggests this pattern may be driven by parallel selective pressures at particular loci, rather than migration; however, we caution that idiosyncratic biases in local ancestry inference at particular markers could also result in a similar pattern.

Potential adaptive introgression outliers are shared across multiple groups

Outlier windows for high *M. nasutus* ancestry that are shared across more than one group are potential candidates for adaptive introgression into *M. guttatus*. Interestingly, there were many more 50kb windows than expected by chance that overlapped across more than two sample groups (Figure 4.4C, Table 4.2). For example, 34 windows were shared between all three northern *M. guttatus* and *admixed* groups, and 13 of these are also shared with southern sympatric *M. guttatus*; these 13 correspond to six stretches of one or more continuous windows across four chromosomes. At least 45% of high-*M. nasutus* outlier windows from any one group were shared with at least one other group, for a total of 228 windows shared across at least two groups (Table 4.2). 104 of those windows were shared across at least one northern and one southern group. One window, Chr04:400,001-450,000, was a high-*M. nasutus* outlier in all six sample groups.

Shared outlier windows for low *M. nasutus* ancestry might indicate regions of reproductive incompatibility that are detrimental in hybrids and therefore resistant to introgression. Most of these outliers were unique to a sample group or shared with only two sample groups; 28 windows were shared between three to four groups and no windows were shared across five or more groups (Figure 4.4D, Table 4.2). However, of the 284 shared windows, 169 were shared between a northern and a southern group. We might expect lower overlap of these windows due to the longer stretches of low introgression throughout the genome, which do not highlight particular loci; still, this pattern suggests that some genomic regions may be under similar selection against introgression across northern and southern groups.

Genetically mapped reproductive isolation loci do not predict natural introgression patterns

Two major candidate loci on chromosomes 07 and 08 (Fishman et al., 2014) responsible for differences in critical photoperiod, an important premating reproductive barrier between *M. guttatus* and *M. nasutus*, did not show obvious patterns of high or low introgression across sample groups. However, the first of these loci, on Chr07, was within a region of substantial missing data, and therefore the appropriate windows were filtered out of our dataset. The second locus, on Chr08, was an outlier for high *M. nasutus* introgression in the DPR *admixed* group, but not in any other sample groups; note that 680 windows (15.6%) are outliers for at least one sample group, so this may be a chance occurrence. Still, we did see an intriguing pattern wherein the larger genomic region surrounding both loci had ancestry frequencies below the genome-wide mean value at CAC (in both *M. guttatus* and *admixed* cohorts) but above the genome-wide mean value at the LM and DPR *admixed* cohorts (Figure 4.5A-B).

Two candidate loci (Zuellig & Sweigart, 2018a, 2018b), jointly responsible for a lethal genetic incompatibility in certain crosses of *M. guttatus* and *M. nasutus*, showed patterns of ancestry opposite to our expectations of negative selection in admixed populations. The 50kb windows containing these two loci were not high-*M. nasutus* or low-*M. nasutus* outliers in any sample group. The area around the *hl13* locus, which in test crosses has a nonfunctional *M. nasutus* allele, has a higher *M. nasutus* ancestry frequency than the genome-wide average in all *admixed* and *M. guttatus* sample groups (Figure 4.5C). Furthermore, the area around the *hl14* locus, which in test crosses has a functional *M. nasutus* allele but a nonfunctional *M. guttatus* allele, has a lower *M. nasutus* ancestry frequency than the genome-wide average in both LM and DPR *admixed* groups, and is near the genome-wide average for CAC groups (Figure 4.5D).

Similarly, sterility loci showed no consistent pattern of ancestry across groups; in fact, QTL for which *M. guttatus* alleles were associated with reduced fertility (G- loci) were more likely to have a median *M. nasutus* ancestry frequency below the genome-wide median for all three tested northern groups, the opposite direction of our expectation (6 of 9 loci below the median in all three groups, Figure 4.6A). Only one sterility locus (on Chr02) had a median frequency above the genome-wide 95th percentile in only one sample group (the southern sympatric group, Figure 4.6A).

For transmission ratio distortion loci, the pattern was more suggestive: loci with expected detrimental effects from the *M. nasutus* allele (N- loci) did tend to have reduced *M. nasutus* ancestry frequencies across all three northern *M. guttatus* and *admixed* groups (4 to 6 of 6 loci with median frequencies below the genome-wide median, Figure 4.6B), but not in the southern groups. Still this pattern was not strong: none of these loci had medians below the 5th percentile in any sample group (Figure 4.6B). Loci with the opposite expected effect (G- loci) were equally

likely to have higher or lower *M. nasutus* ancestry frequencies, suggesting that selection against incompatibility alleles from the minor parent (N- loci) may be stronger or more efficient than selection against incompatibility alleles in the majority parent (G- loci).

For most sample groups, N- and G- loci did not overlap significantly with either high-*M. nasutus* or low-*M. nasutus* outliers (Figure 4.7). There was a slight trend of higher overlap with low-*M. nasutus* outliers compared to high-*M. nasutus* outliers for both N- and G- loci. For N- loci, this trend is in the expected direction if selection is eliminating incompatible *M. nasutus* alleles. For G- loci, this trend is opposite to our expectations of selection against *M. guttatus* alleles, but may suggest that minor parent incompatibility alleles are more likely to be removed than majority-parent incompatibilities. Overall, the lack of significant overlap indicates that selection against ancestry is not strong enough at these loci to drive ancestry patterns in our populations.

Discussion

Here, we examine genome-wide patterns of interspecific ancestry across two sampling areas separated by ~1000 km, in order to address two of the biggest open questions in the study of hybridization: 1) how repeatable are patterns of ancestry across independent cases of hybridization, and 2) how well do premating and postmating reproductive barrier loci from genetic mapping studies predict actual patterns of gene flow in wild populations. We find that geographically proximate locations have more shared ancestry patterns than more distant geographic areas, but that even distant areas have substantial similarity in genomic ancestry patterns. These shared patterns appear despite major differences in the extent and timing of introgression. However, there are also important differences in genomic patterns across areas,

including at presumably influential candidate loci for reproductive barrier traits. We find that overall, mapped reproductive trait loci were not significant outliers for ancestry in wild populations. Epistatic incompatibility loci showed more consistent patterns than either single-locus fertility loci or premating traits, but were still not reliable predictors of ancestry. Our data show that a combination of shared and population-specific processes are likely driving introgression patterns across populations, and that these processes do not map cleanly to charismatic traits from laboratory crosses. *Mimulus guttatus* and *Mimulus nasutus* appear to be partially isolated by a highly polygenic suite of loci that vary across space, rather than a few distinct ‘islands of speciation’. Similarly, signatures of increased admixture are dispersed throughout the genome, with no ‘smoking-gun’ cases of rapid adaptive introgression, though shared signals across populations provide potential candidates for further exploration. This study highlights the power of combining data across multiple replicate populations, and demonstrates that multiple genetic approaches are necessary to disentangle the complexities of species divergence.

Sympatric populations differ widely in the extent and timing of hybridization

We previously showed that our northern CAC and LM admixed populations have substantial directional introgression resulting in stable populations of hybrids, in which few or no individuals of *M. guttatus* ancestry are free of *M. nasutus* introgression (Chapter III). The sympatric populations from our southern DPR area tell a very different story. At low elevations, samples from the DPR area carried much less introgression than the typical CAC or LM *M. guttatus* individual. This finding matches previous work from DPR showing that hybridization is rare (N. H. Martin & Willis, 2007); still, other studies have detected some introgression in this area (Sweigart & Willis, 2003; Zuellig & Sweigart, 2018a), raising the question of why hybrids

at DPR have not persisted in the same way compared to northern locations. Perhaps a patchier microhabitat structure facilitates the persistence of hybrids at CAC and LM in an intermediate niche that doesn't exist reliably at DPR. Another possibility is that the intermediate phenology of hybrids at CAC has helped to isolate and maintain a hybrid swarm once formed, while the more variable phenological conditions of DPR (N. H. Martin & Willis, 2007 and Farnitano *pers. obs.*) are not ideal for hybrid persistence, at least in some years.

At higher elevations, we find a distinct pattern of frequent recent (first- and early-generation) hybrids between *M. guttatus* and *M. nasutus*, which are extremely rare at both CAC and DPR sites. It is unclear whether we are capturing an early snapshot of hybrid swarm formation at high elevations, which will eventually settle into a dynamic more similar to CAC, or whether these patterns are highlighting fundamental differences in the patterns of premating isolation and hybrid survival at high elevations. Ecological differences at these locations are numerous, including temperature, snowpack, the timing of snowmelt and water flow, and the makeup of surrounding plant communities. Phenological cues, pollinator communities, and demographic factors likely also differ. Furthermore, a third reproductively compatible lineage, *Mimulus laciniatus*, is also found in these areas and could be interacting with *M. guttatus* and *M. nasutus* in more complex ways (Tataru et al., 2023). Future studies examining reproductive barriers and their environmental correlates in these higher elevation populations would be a useful point of contrast with our work in the CAC area.

Genome-wide effects of chromosome structure on introgression

Multiple independent datasets in both plants and animals have shown that patterns of introgression are associated with recombination rate across the genome (Burri et al., 2015; Juric et al., 2016; Kenney & Sweigart, 2016; S. H. Martin et al., 2019; Nouhaud et al., 2022; Schumer

et al., 2018). This pattern is thought to be driven by linked selection removing parent alleles from hybrid populations: areas of low recombination are removed more completely, while neutral or adaptive alleles in higher recombination areas can become decoupled from incompatible or deleterious alleles and persist. We find support for this association in our data, though it is stronger with gene density than with recombination: windows with more genes tend to have less *M. nasutus* introgression. The recombination map we used may not have the resolution necessary to detect a stronger relationship; in addition, many windows in the lowest recombination regions were excluded from our analysis due to low marker density and poor alignment coverage, which may reduce our ability to detect this signal. Gene density is known to be strongly correlated with recombination rate in *Mimulus* (Kenney & Sweigart, 2016), and may even be causally associated with recombination breakpoint formation (Hellsten et al., 2013). Interestingly, while gene density is a strong predictor of introgression patterns, it does not fully explain patterns of ancestry correlation across groups, suggesting that other factors are necessary to explain the sharing of ancestry patterns in different ancestry groups or geographic groups.

Introgression at particular genomic loci filters from admixed into allopatric populations

Within locations, ancestry patterns in *M. guttatus* groups were highly correlated with patterns in nearby more-admixed groups despite lower overall introgression, suggesting that introgression may be filtering back from *admixed* populations into *M. guttatus*. In particular, the *M. guttatus* group at CAC has relatively distinct, clean peaks of admixture on multiple chromosomes; these peaks are usually also elevated in the CAC *admixed* group but surrounded by much more noise, consistent with a filtering of only certain introgression segments from *admixed* into less admixed individuals. Similarly, allopatric *M. guttatus* from the DPR area share regions of elevated introgression with sympatric *M. guttatus*, suggesting that migration from

sympatry to allopatry may bring *M. nasutus* alleles into allopatry, but only after they have survived selection in sympatry. This process of a genomic sieve, or selective filter for introgression, has been discussed as a way for neutral or adaptive introgression at certain loci to spread despite a general pattern of reproductive incompatibility (Chhatre et al., 2018; Martinsen et al., 2001; Simon et al., 2021). While we do not detect any strong sweeps of adaptive introgression (i.e., no *M. nasutus* alleles have risen to fixation or very high frequency in *M. guttatus*), the loci that have ‘escaped’ into *M. guttatus* may still be interesting from an adaptive perspective.

Ancestry correlations across distant locations indicate a shared basis for introgression patterns

We also find that genomic ancestry patterns have significant correlations across locations.

Admixed populations at CAC and LM, two locations ~4km apart that are presumably independent but may share occasional migrants, have very similar patterns of genomic ancestry, including many shared outliers for low and high *M. nasutus* introgression. These similarities could be due to a combination of migration and shared history, common patterns due to shared genome architecture, or similar selective forces at individual loci with shared alleles. Our ANOVA approach argues that shared genome architecture is important but only part of the story, and that either shared history or shared selection pressures contribute to these patterns.

Similarity between DPR *M. guttatus* and CAC or LM groups cannot be attributed to migration and shared history due to the vast distances separating them. Still, we find significant correlations and overlap of high-introgression outliers between southern and northern sites that appear to go beyond a shared genome architecture. These similarities are substantially weaker than in the CAC vs. LM comparisons, so greater geographic distance does appear to reduce the amount of shared overlap. Different combinations of alleles at each location, different

environmental selective pressures, and different demographic conditions could all contribute to the reduction in pattern sharing. But the remaining correlation is intriguing and argues for at least some parallel selective forces acting during hybridization at the two sympatric locations.

Overlap of high-*M. nasutus* introgression outliers (candidates for adaptive introgression) among groups was much more consistent than overlap of low-introgression outliers (candidates for assortative mating or incompatibilities). This may be an artifact of the asymmetrical distribution of ancestry frequencies: most chromosomes have long stretches of relatively low introgression, punctuated by sharper peaks of higher introgression. This means that low-introgression outliers will be buried in large swathes of similar ancestry and may not be reliably localized to particular regions, while high-introgression outliers are more likely to indicate real (and shared) ancestry peaks. In addition, incompatibility loci or other sources of selection against introgression may be highly polygenic (Guerrero et al., 2017; Pritchard et al., 2010; Schumer et al., 2018), making it harder to detect the effect of a few clear outliers within a genome-wide pattern. Sources of adaptive introgression or positive selection may be polygenic as well: we detect many outlier peaks but no obvious sweeps to fixation; but these loci may still be relatively rare compared to negatively selected loci.

A few other patterns are worth noting. The DPR area *admixed* population has much noisier ancestry frequencies than any other group; this comes from a combination of low sample size (22 individuals in the group) and the higher proportion of early-generation hybrids, which have had fewer generations of recombination and selection to establish consistent frequencies. As a result, correlations and outlier overlaps with this group are generally lower than in other comparisons, though we still detect some shared patterns. We also do not expect to see a strong pattern of ancestry correlation in comparisons involving *M. nasutus* groups, both because of the

low amounts of introgression in that direction, and because selective pressures in a majority-*M. nasutus* background may act differently compared to groups with a majority-*M. guttatus* background. However, we do detect some sharing between southern admixed and *M. nasutus* groups, which may indicate potential candidates for adaptive introgression from *M. guttatus* into *M. nasutus* at these loci. This is particularly intriguing because our population data show more symmetrical admixture at the southern high-elevation locations where those *admixed* individuals are primarily found.

Candidate genes for premating and postmating barriers are not strong outliers for introgression

The premating and postmating reproductive barrier candidate genes we tested are not typically found in high-*M. nasutus* or low-*M. nasutus* outlier regions. In fact, they often showed patterns opposite to their predicted effects. For instance, the hybrid lethality locus *h113* on chromosome 13, which harbors a nonfunctional *M. guttatus* allele, shows depressed *M. nasutus* ancestry surrounding the locus in all *M. guttatus* and *admixed* populations. The corresponding locus *h114* on chromosome 14, which has a nonfunctional *M. nasutus* allele, has slightly elevated *M. nasutus* ancestry in LM and within DPR area *admixed* individuals. We know that both incompatible alleles are polymorphic within species, so incompatible combinations may not be possible in all populations. Furthermore, the incompatible combination is doubly recessive, and with low levels of hybridization, selection may simply not be strong enough to affect ancestry patterns at these loci.

Premating loci on chromosomes 07 and 08, which affect flowering phenology via photoperiod response, do not have consistent directional effects across populations. This suggests that phenology, while important for isolation, does not strongly alter hybrid fitness. Population structure and reproductive isolation may end up as a byproduct of the ancestry

outcomes at premating loci, rather than a driver of their ancestry patterns. Interestingly, these loci have different directional effects in CAC vs. LM populations, two populations with contrasting population structure and phenology distributions. CAC has more variability in flowering time, with distinct flowering cohorts, while LM appears to be a more continuous admixed population with less flowering time differentiation between *M. nasutus* and *admixed* groups (Chapter III, Sweigart pers. obs.). The allelic distributions and effects of these loci in the DPR area are also not known, but since photoperiod cues occur at a different time of the season, we expect that different alleles may be present within *M. guttatus* compared to the northern area. The relative roles of these genetic factors and other environmental inputs on population-level phenology differences require further attention.

Sterility loci do not have a consistent effect on introgression patterns. Since these loci were identified using recombinant inbred lines (Mantel & Sweigart, 2024), and mostly affected *M. guttatus* genotypes, we suspect that some of these may have recessive effects that were hidden from selection in diverse, outcrossing admixed populations. Furthermore, these loci were often detected in only one of multiple *M. guttatus* inbred lines (Mantel & Sweigart, 2024), which suggests that selection could replace them with other *M. guttatus* alleles instead of *M. nasutus* alleles; we would not see this form of selection in our ancestry frequency data.

Transmission ratio distortion loci, which represent multilocus interactions between *M. nasutus* and *M. guttatus* ancestry, have a more consistent, but still underwhelming, effect. *M. nasutus* interacting alleles were the most underrepresented at CAC, and somewhat at LM, but not at southern locations. The *M. guttatus* interactors were not reliably over- or underrepresented even at CAC, suggesting that selection against incompatibilities will more efficiently remove the minor parent ancestry alleles (*M. nasutus*) compared to the majority parent (*M. guttatus*). These

loci were mapped using lines sourced from CAC, so a stronger signal there is logical, but suggests that incompatibility interactions may be highly localized and not often shared even among geographically close sympatric sites. Even within CAC, these incompatibility loci, like the sterility loci, are often not shared among *M. guttatus* lines (Mantel & Sweigart, 2024), highlighting that incompatibilities between *M. guttatus* and *M. nasutus* are likely due to the combined effects of many rare or restricted loci rather than fixed differences at the species level. Other incompatibilities mapped in these species have been geographically restricted as well (Fishman & Willis, 2006; Sweigart et al., 2007; Wright et al., 2013; Zuellig & Sweigart, 2018a). The relatively young age of *M. guttatus* - *M. nasutus* divergence means that incompatibilities may not have had time to fix; in swordtails, repeatability of introgression is related to the time since divergence (Langdon et al., 2024). The enormous genetic diversity and large geographic range of *M. guttatus* (Brandvain et al., 2014; Twyford et al., 2020) might also make range-wide fixation of alleles very slow; for instance, diverse *S. saccharomyces* strains show a variety of context-dependent outcomes after hybridization (Brice et al., 2021).

A recent analysis of the classic *Mus musculus* hybrid zone found a similar result to ours, that natural variation is not necessarily predicted by the results of crossing studies (Frayer & Payseur, 2024). They discuss a number of possibilities to explain this discrepancy, some of which are relevant for *Mimulus*: incompatibilities we notice in the lab may not be the most important for natural populations, and reproductive incompatibility may be polymorphic and polygenic, diluting the signal at any particular locus. A third possibility discussed, that long-standing incompatibilities are purged over historical time in hybrid zones and therefore do not show effects in contemporary populations, is not a likely explanation of our data, because the loci we tested were discovered in the very populations we are investigating. Together, our results

show that discordance between crossing studies and natural variation may be a wider phenomenon, not a quirk of one particular study system.

Most of the outlier regions we find, even those shared among populations, are not tied to known premating or postmating loci. Presumably at least some of these have important selective effects in admixed populations, though others may be the result of drift or be affected by directional selection unrelated to species differences. Future work could use these regions as a starting point to identify additional genes that are important to speciation and hybridization in this system. We stress that neither trait mapping nor admixture scanning is sufficient to fully understand the genetic basis of species barriers; instead, we can learn a great deal by comparing and contrasting the findings from multiple independent approaches.

Through our data, we can assemble a more complete picture of admixture between *Mimulus guttatus* and *Mimulus nasutus*, providing a comparison with other studied hybrid systems. Hybridization is dynamic and varied across the landscape: we find evidence of recent hybridization in one area, an older stable swarm of admixture at another, and minimal hybridization despite sympatry in a third area. Genomic patterns are variable across loci but correlated across geography, suggesting similar influences on locus-by-locus introgression across these different areas of sympatry. While there is evidence for both pervasive selection against introgression, and regions of relaxed or positive selection, we do not find clearly defined ‘islands of divergence’ as detected in other systems, nor do we see particular loci that have swept through entire populations via strong adaptive introgression. For both selection against incompatibility and potential adaptive introgression, the core signal appears to be highly polygenic and dispersed throughout the genome. As such, we don’t see strong effects of any particular known candidate

loci on introgression patterns. Comparative work in a larger number of actively hybridizing systems will help determine whether these patterns are expected or exceptional.

References

- Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C. A., Buggs, R., Butlin, R. K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S. H., Hermansen, J. S., Hewitt, G., Hudson, A. G., Jiggins, C., ... Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, 26(2), 229–246. <https://doi.org/10.1111/j.1420-9101.2012.02599.x>
- Baack, E. J., & Rieseberg, L. H. (2007). A genomic view of introgression and hybrid speciation. *Current Opinion in Genetics & Development*, 17(6), 513–518. <https://doi.org/10.1016/j.gde.2007.09.001>
- Barton, N. H., & Hewitt, G. M. (1985). Analysis of Hybrid Zones. *Annual Review of Ecology and Systematics*, 16, 113–148.
- Barton, N. H., & Hewitt, G. M. (1989). Adaptation, speciation and hybrid zones. *Nature*, 341(6242), 497–503. <https://doi.org/10.1038/341497a0>
- Behm, J. E., Ives, A. R., & Boughman, J. W. (2010). Breakdown in postmating isolation and the collapse of a species pair through hybridization. *The American Naturalist*, 175(1), 11–26. <https://doi.org/10.1086/648559>
- Bourret, S. L., Kovach, R. P., Cline, T. J., Strait, J. T., & Muhlfeld, C. C. (2022). High dispersal rates in hybrids drive expansion of maladaptive hybridization. *Proceedings of the Royal Society B: Biological Sciences*, 289(1986), 20221813. <https://doi.org/10.1098/rspb.2022.1813>
- Brandvain, Y., Kenney, A. M., Flagel, L., Coop, G., & Sweigart, A. L. (2014). Speciation and Introgression between *Mimulus nasutus* and *Mimulus guttatus*. *PLoS Genetics*, 10(6), e1004410. <https://doi.org/10.1371/journal.pgen.1004410>
- Brice, C., Zhang, Z., Bendixsen, D., & Stelkens, R. (2021). Hybridization Outcomes Have Strong Genomic and Environmental Contingencies. *The American Naturalist*, 198(3), E53–E67. <https://doi.org/10.1086/715356>
- Burri, R., Nater, A., Kawakami, T., Mugal, C. F., Olason, P. I., Smeds, L., Suh, A., Dutoit, L., Bureš, S., Garamszegi, L. Z., Hogner, S., Moreno, J., Qvarnström, A., Ružić, M., Sæther, S.-A., Sætre, G.-P., Török, J., & Ellegren, H. (2015). Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of *Ficedula* flycatchers. *Genome Research*, 25(11), 1656–1665. <https://doi.org/10.1101/gr.196485.115>

- Chaturvedi, S., Lucas, L. K., Buerkle, C. A., Fordyce, J. A., Forister, M. L., Nice, C. C., & Gompert, Z. (2020). Recent hybrids recapitulate ancient hybrid outcomes. *Nature Communications*, 11(1), Article 1. <https://doi.org/10.1038/s41467-020-15641-x>
- Chhatre, V. E., Evans, L. M., DiFazio, S. P., & Keller, S. R. (2018). Adaptive introgression and maintenance of a trispecies hybrid complex in range-edge populations of *Populus*. *Molecular Ecology*, 27(23), 4820–4838. <https://doi.org/10.1111/mec.14820>
- Corbett-Detig, R., & Nielsen, R. (2017). A Hidden Markov Model Approach for Simultaneously Estimating Local Ancestry and Admixture Time Using Next Generation Sequence Data in Samples of Arbitrary Ploidy. *PLOS Genetics*, 13(1), e1006529. <https://doi.org/10.1371/journal.pgen.1006529>
- Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, 23(13), 3133–3157. <https://doi.org/10.1111/mec.12796>
- Curry, C. M. (2015). An Integrated Framework for Hybrid Zone Models. *Evolutionary Biology*, 42(3), 359–365. <https://doi.org/10.1007/s11692-015-9332-9>
- Dean, L. L., Whiting, J. R., Jones, F. C., & MacColl, A. D. C. (2024). Reproductive isolation in a three-way contact zone. *Molecular Ecology*, 33(5), e17275. <https://doi.org/10.1111/mec.17275>
- Edelman, N. B., Frandsen, P. B., Miyagi, M., Clavijo, B., Davey, J., Dikow, R. B., García-Accinelli, G., Van Belleghem, S. M., Patterson, N., Neafsey, D. E., Challis, R., Kumar, S., Moreira, G. R. P., Salazar, C., Chouteau, M., Counterman, B. A., Papa, R., Blaxter, M., Reed, R. D., ... Mallet, J. (2019). Genomic architecture and introgression shape a butterfly radiation. *Science*, 366(6465), 594–599. <https://doi.org/10.1126/science.aaw2090>
- Farnitano, M. C., & Sweigart, A. L. (2024). *Low-cost tagmentation library prep for low-coverage Illumina sequencing*. protocols.io. [dx.doi.org/10.17504/protocols.io.6qpvr8zrplmk/v1](https://doi.org/10.17504/protocols.io.6qpvr8zrplmk/v1)
- Feller, A. F., Peichel, C. L., & Seehausen, O. (2024). Testing for a role of postzygotic incompatibilities in rapidly speciated Lake Victoria cichlids. *Evolution*, 78(4), 652–664. <https://doi.org/10.1093/evolut/qpae007>
- Fishman, L. (2020). *96-well CTAB-chloroform DNA extraction*. protocols.io. <https://dx.doi.org/10.17504/protocols.io.bgv6jw9e>
- Fishman, L., Sweigart, A. L., Kenney, A. M., & Campbell, S. (2014). Major quantitative trait loci control divergence in critical photoperiod for flowering between selfing and outcrossing species of monkeyflower (*Mimulus*). *New Phytologist*, 201(4), 1498–1507. <https://doi.org/10.1111/nph.12618>

- Fishman, L., & Willis, J. H. (2006). A cytonuclear incompatibility causes anther sterility in *Mimulus* hybrids. *Evolution*, 60(7), 1372–1381. <https://doi.org/10.1111/j.0014-3820.2006.tb01216.x>
- Frayser, M. E., & Payseur, B. A. (2024). Do genetic loci that cause reproductive isolation in the lab inhibit gene flow in nature? *Evolution*, 78(6), 1025–1038. <https://doi.org/10.1093/evolut/qpae044>
- Goulet, B. E., Roda, F., & Hopkins, R. (2017). Hybridization in plants: Old ideas, new techniques. *Plant Physiology*, 173(1), 65–78. <https://doi.org/10.1104/pp.16.01340>
- Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M. H.-Y., Hansen, N. F., Durand, E. Y., Malaspina, A.-S., Jensen, J. D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H. A., ... Pääbo, S. (2010). A draft sequence of the Neandertal genome. *Science (New York, N.Y.)*, 328(5979), 710–722. <https://doi.org/10.1126/science.1188021>
- Guerrero, R. F., Muir, C. D., Josway, S., & Moyle, L. C. (2017). Pervasive antagonistic interactions among hybrid incompatibility loci. *PLOS Genetics*, 13(6), e1006817. <https://doi.org/10.1371/journal.pgen.1006817>
- Hasselman, D. J., Argo, E. E., McBride, M. C., Bentzen, P., Schultz, T. F., Perez-Umphrey, A. A., & Palkovacs, E. P. (2014). Human disturbance causes the formation of a hybrid swarm between two naturally sympatric fish species. *Molecular Ecology*, 23(5), 1137–1152. <https://doi.org/10.1111/mec.12674>
- Hellsten, U., Wright, K. M., Jenkins, J., Shu, S., Yuan, Y., Wessler, S. R., Schmutz, J., Willis, J. H., & Rokhsar, D. S. (2013). Fine-scale variation in meiotic recombination in *Mimulus* inferred from population shotgun sequencing. *Proceedings of the National Academy of Sciences*, 110(48), 19478–19482. <https://doi.org/10.1073/pnas.1319032110>
- Hijmans, R. J., Barbosa, M., Ghosh, A., & Mandel, A. (2024). *geodata: Download Geographic Data* (Version 0.6-2) [Computer software]. <https://cran.r-project.org/web/packages/geodata/index.html>
- Hohenlohe, P. A., Bassham, S., Currey, M., & Cresko, W. A. (2012). Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1587), 395–408. <https://doi.org/10.1098/rstb.2011.0245>
- Howard, D. J., Waring, G. L., Tibbets, C. A., & Gregory, P. G. (1993). Survival of Hybrids in a Mosaic Hybrid Zone. *Evolution*, 47(3), 789–800. <https://doi.org/10.2307/2410184>
- Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M. C., White, S., Birney, E., Searle, S., Schmutz, J., Grimwood, J., Dickson, M. C., Myers, R. M., Miller, C. T., Summers, B. R., Knecht, A. K., ... Kingsley, D. M. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, 484(7392), 55–61. <https://doi.org/10.1038/nature10944>

- Juric, I., Aeschbacher, S., & Coop, G. (2016). The Strength of Selection against Neanderthal Introgression. *PLOS Genetics*, *12*(11), e1006340. <https://doi.org/10.1371/journal.pgen.1006340>
- Kenney, A. M., & Sweigart, A. L. (2016). Reproductive isolation and introgression between sympatric *Mimulus* species. *Molecular Ecology*, *25*(11), 2499–2517. <https://doi.org/10.1111/mec.13630>
- Klein, E. K., Lagache-Navarro, L., & Petit, R. J. (2017). Demographic and spatial determinants of hybridization rate. *Journal of Ecology*, *105*(1), 29–38. <https://doi.org/10.1111/1365-2745.12674>
- Kleindorfer, S., O'Connor, J. A., Dudaniec, R. Y., Myers, S. A., Robertson, J., & Sulloway, F. J. (2014). Species collapse via hybridization in Darwin's tree finches. *The American Naturalist*, *183*(3), 325–341. <https://doi.org/10.1086/674899>
- Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, *15*(1), 356. <https://doi.org/10.1186/s12859-014-0356-4>
- Langdon, Q. K., Groh, J. S., Aguillon, S. M., Powell, D. L., Gunn, T., Payne, C., Baczenas, J. J., Donny, A., Dodge, T. O., Du, K., Schartl, M., Ríos-Cárdenas, O., Gutiérrez-Rodríguez, C., Morris, M., & Schumer, M. (2024). Swordtail fish hybrids reveal that genome evolution is surprisingly predictable after initial hybridization. *PLOS Biology*, *22*(8), e3002742. <https://doi.org/10.1371/journal.pbio.3002742>
- Lotsy, J. P. (1931). On the species of the taxonomist in its relation to evolution. *Genetica*, *13*(1), 1–16. <https://doi.org/10.1007/BF01725037>
- Lowry, D. B., & Willis, J. H. (2010). A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biology*, *8*(9). <https://doi.org/10.1371/journal.pbio.1000500>
- Lu, Z., Hofmeister, B. T., Vollmers, C., DuBois, R. M., & Schmitz, R. J. (2017). Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes. *Nucleic Acids Research*, *45*(6), e41. <https://doi.org/10.1093/nar/gkw1179>
- Machado, C. A., Haselkorn, T. S., & Noor, M. A. F. (2007). Evaluation of the Genomic Extent of Effects of Fixed Inversion Differences on Intraspecific Variation and Interspecific Gene Flow in *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, *175*(3), 1289–1306. <https://doi.org/10.1534/genetics.106.064758>
- Macholán, M., Munclinger, P., Šugerková, M., Dufková, P., Bímová, B., Božíková, E., Zima, J., & Piálek, J. (2007). Genetic analysis of autosomal and x-linked markers across a mouse hybrid zone. *Evolution*, *61*(4), 746–771. <https://doi.org/10.1111/j.1558-5646.2007.00065.x>

- Mallet, J., Besansky, N., & Hahn, M. W. (2015). How reticulated are species? *BioEssays*, 38(2), 140–149. <https://doi.org/10.1002/bies.201500149>
- Mantel, S. J., & Sweigart, A. L. (2024). Postzygotic barriers persist despite ongoing introgression in hybridizing *Mimulus* species. *Molecular Ecology*, 33(4), e17261. <https://doi.org/10.1111/mec.17261>
- Maroja, L. S., Andrés, J. A., & Harrison, R. G. (2009). Genealogical discordance and patterns of introgression and selection across a cricket hybrid zone. *Evolution*, 63(11), 2999–3015. <https://doi.org/10.1111/j.1558-5646.2009.00767.x>
- Martin, N. H., & Willis, J. H. (2007). Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric *Mimulus* species. *Evolution*, 61(1), 68–82. <https://doi.org/10.1111/j.1558-5646.2007.00006.x>
- Martin, S. H., Davey, J. W., Salazar, C., & Jiggins, C. D. (2019). Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLoS Biology*, 17(2), 1–28. <https://doi.org/10.1371/journal.pbio.2006288>
- Martinsen, G. D., Whitham, T. G., Turek, R. J., & Keim, P. (2001). Hybrid Populations Selectively Filter Gene Introgression Between Species. *Evolution*, 55(7), 1325–1335. <https://doi.org/10.1111/j.0014-3820.2001.tb00655.x>
- Meisner, J., & Albrechtsen, A. (2018). Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data. *Genetics*, 210(2), 719–731. <https://doi.org/10.1534/genetics.118.301336>
- Menon, M., Bagley, J. C., Page, G. F. M., Whipple, A. V., Schoettle, A. W., Still, C. J., Wehenkel, C., Waring, K. M., Flores-Renteria, L., Cushman, S. A., & Eckert, A. J. (2021). Adaptive evolution in a conifer hybrid zone is driven by a mosaic of recently introgressed and background genetic variants. *Communications Biology*, 4(1), Article 1. <https://doi.org/10.1038/s42003-020-01632-7>
- Moore, W. S., & Buchanan, E. B. (1985). Stability of the northern flicker hybrid zone in historical times: Implications for adaptive speciation theory. *Evolution*, 39(1), 135–151. <https://doi.org/10.1111/j.1558-5646.1985.tb04086.x>
- Moran, B. M., Payne, C. Y., Powell, D. L., Iverson, E. N. K., Donny, A. E., Banerjee, S. M., Langdon, Q. K., Gunn, T. R., Rodriguez-Soto, R. A., Madero, A., Baczenas, J. J., Kleczko, K. M., Liu, F., Matney, R., Singhal, K., Leib, R. D., Hernandez-Perez, O., Corbett-Detig, R., Frydman, J., ... Schumer, M. (2024). A lethal mitonuclear incompatibility in complex I of natural hybrids. *Nature*, 626(7997), 119–127. <https://doi.org/10.1038/s41586-023-06895-8>
- Nelson, T. C., Stathos, A. M., Vanderpool, D. D., Finseth, F. R., Yuan, Y., & Fishman, L. (2021). Ancient and recent introgression shape the evolutionary history of pollinator adaptation and speciation in a model monkeyflower radiation (*Mimulus* section

- Erythranthe). *PLoS Genetics*, 17(2), e1009095.
<https://doi.org/10.1371/journal.pgen.1009095>
- Nouhaud, P., Martin, S. H., Portinha, B., Sousa, V. C., & Kulmuni, J. (2022). Rapid and predictable genome evolution across three hybrid ant populations. *PLOS Biology*, 20(12), e3001914. <https://doi.org/10.1371/journal.pbio.3001914>
- Pardo-Diaz, C., Salazar, C., Baxter, S. W., Merot, C., & Figueiredo-Ready, W. (2012). Adaptive Introgression across Species Boundaries in Heliconius Butterflies. *PLoS Genet*, 8(6), 1002752. <https://doi.org/10.1371/journal.pgen.1002752>
- Pickup, M., Brandvain, Y., Fraïsse, C., Yakimowski, S., Barton, N. H., Dixit, T., Lexer, C., Cereghetti, E., & Field, D. L. (2019). Mating system variation in hybrid zones: Facilitation, barriers and asymmetries to gene flow. *New Phytologist*, 224(3), 1035–1047. <https://doi.org/10.1111/nph.16180>
- Powell, D. L., Garcia-Olazabal, M., Keegan, M., Reilly, P., Du, K., Díaz-Loyo, A. P., Banerjee, S., Blakkan, D., Reich, D., Andolfatto, P., Rosenthal, G. G., Schartl, M., & Schumer, M. (2020). Natural hybridization reveals incompatible alleles that cause melanoma in swordtail fish. *Science*, 368(May), 731–736. <https://doi.org/10.1101/2019.12.12.874586>
- Pritchard, J. K., Pickrell, J. K., & Coop, G. (2010). The Genetics of Human Adaptation: Hard Sweeps, Soft Sweeps, and Polygenic Adaptation. *Current Biology*, 20(4), R208–R215. <https://doi.org/10.1016/j.cub.2009.11.055>
- Quilodrán, C. S., Tsoupas, A., & Currat, M. (2020). The Spatial Signature of Introgression After a Biological Invasion With Hybridization. *Frontiers in Ecology and Evolution*, 8. <https://www.frontiersin.org/articles/10.3389/fevo.2020.569620>
- R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Racimo, F., Sankararaman, S., Nielsen, R., & Huerta-Sánchez, E. (2015). Evidence for archaic adaptive introgression in humans. *Nature Reviews Genetics*, 16(6), 359–371. <https://doi.org/10.1038/nrg3936>
- Rand, D. M., & Harrison, R. G. (1989). Ecological genetics of a mosaic hybrid zone: Mitochondrial, nuclear, and reproductive differentiation of crickets by soil type. *Evolution*, 43(2), 432–449. <https://doi.org/10.1111/j.1558-5646.1989.tb04238.x>
- Randi, E., & Lucchini, V. (2002). Detecting rare introgression of domestic dog genes into wild wolf (*Canis lupus*) populations by Bayesian admixture analyses of microsatellite variation. *Conservation Genetics*, 3(1), 29–43. <https://doi.org/10.1023/A:1014229610646>
- Ruhsam, M., Hollingsworth, P. M., & Ennos, R. A. (2011). Early evolution in a hybrid swarm between outcrossing and selfing lineages in Geum. *Heredity*, 107(3), 246–255. <https://doi.org/10.1038/hdy.2011.9>

- Runemark, A., Trier, C. N., Eroukhmanoff, F., Hermansen, J. S., Matschiner, M., Ravinet, M., Elgvin, T. O., & Sætre, G.-P. (2018). Variation and constraints in hybrid genome formation. *Nature Ecology & Evolution*, 2(3), Article 3. <https://doi.org/10.1038/s41559-017-0437-7>
- Sachdeva, H., & Barton, N. H. (2018). Introgression of a Block of Genome Under Infinitesimal Selection. *Genetics*, 209(4), 1279–1303. <https://doi.org/10.1534/genetics.118.301018>
- Schumer, M., Powell, D. L., & Corbett-Detig, R. (2020). Versatile simulations of admixture and accurate local ancestry inference with mixnmatch and ancestryinfer. *Molecular Ecology Resources*, 20(4), 1141–1151. <https://doi.org/10.1111/1755-0998.13175>
- Schumer, M., Xu, C., Powell, D. L., Durvasula, A., Skov, L., Holland, C., Blazier, J. C., Sankararaman, S., Andolfatto, P., Gil, †, Rosenthal, G., & Przeworski, M. (2018). Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science*, 360(6389), 656–660.
- Sedghifar, A., Brandvain, Y., & Ralph, P. (2016). Beyond clines: Lineages and haplotype blocks in hybrid zones. *Molecular Ecology*, 25(11), 2559–2576. <https://doi.org/10.1111/mec.13677>
- Sianta, S. A., Moeller, D. A., & Brandvain, Y. (2024). The extent of introgression between incipient *Clarkia* species is determined by temporal environmental variation and mating system. *Proceedings of the National Academy of Sciences*, 121(12), e2316008121. <https://doi.org/10.1073/pnas.2316008121>
- Simon, A., Fraïsse, C., El Ayari, T., Liautard-Haag, C., Strelkov, P., Welch, J. J., & Bierne, N. (2021). How do species barriers decay? Concordance and local introgression in mosaic hybrid zones of mussels. *Journal of Evolutionary Biology*, 34(1), 208–223. <https://doi.org/10.1111/jeb.13709>
- Stankowski, S., & Streisfeld, M. A. (2015). Introgressive hybridization facilitates adaptive divergence in a recent radiation of monkeyflowers. *Proceedings of the Royal Society B: Biological Sciences*, 282(1814), 20151666. <https://doi.org/10.1098/rspb.2015.1666>
- Stebbins, G. L. (1959). The Role of Hybridization in Evolution. *Proceedings of the American Philosophical Society*, 103(2), 231–251.
- Streisfeld, M. A., & Kohn, J. R. (2005). Contrasting patterns of floral and molecular variation across a cline in *Mimulus aurantiacus*. *Evolution*, 59(12), 2548–2559. <https://doi.org/10.1111/j.0014-3820.2005.tb00968.x>
- Sweigart, A. L., Mason, A. R., & Willis, J. H. (2007). Natural variation for a hybrid incompatibility between two species of *Mimulus*. *Evolution*, 61(1), 141–151. <https://doi.org/10.1111/j.1558-5646.2007.00011.x>

- Sweigart, A. L., & Willis, J. H. (2003). Patterns of nucleotide diversity in two species of *Mimulus* are affected by mating system and asymmetric introgression. *Evolution*, 57(11), 2490–2506. <https://doi.org/10.1111/j.0014-3820.2003.tb01494.x>
- Tataru, D., Wheeler, E. C., & Ferris, K. G. (2023). Spatially and temporally varying selection influence species boundaries in two sympatric *Mimulus*. *Proceedings of the Royal Society B: Biological Sciences*, 290(1992), 20222279. <https://doi.org/10.1098/rspb.2022.2279>
- Taylor, S. A., & Larson, E. L. (2019). Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nature Ecology & Evolution*, 3(2), Article 2. <https://doi.org/10.1038/s41559-018-0777-y>
- Trier, C. N., Hermansen, J. S., Sætre, G.-P., & Bailey, R. I. (2014). Evidence for Mito-Nuclear and Sex-Linked Reproductive Barriers between the Hybrid Italian Sparrow and Its Parent Species. *PLOS Genetics*, 10(1), e1004075. <https://doi.org/10.1371/journal.pgen.1004075>
- Twyford, A. D., Wong, E. L. Y., & Friedman, J. (2020). Multi-level patterns of genetic structure and isolation by distance in the widespread plant *Mimulus guttatus*. *Heredity*, 125(4), 227–239. <https://doi.org/10.1038/s41437-020-0335-7>
- Valbuena-Carabaña, M., González-Martínez, S. C., Hardy, O. J., & Gil, L. (2007). Fine-scale spatial genetic structure in mixed oak stands with different levels of hybridization. *Molecular Ecology*, 16(6), 1207–1219. <https://doi.org/10.1111/j.1365-294X.2007.03231.x>
- Wessinger, C. A., Katzer, A. M., Hime, P. M., Rausher, M. D., Kelly, J. K., & Hileman, L. C. (2023). A few essential genetic loci distinguish *Penstemon* species with flowers adapted to pollination by bees or hummingbirds. *PLOS Biology*, 21(9), e3002294. <https://doi.org/10.1371/journal.pbio.3002294>
- Wickham, H., & RStudio. (2023). *tidyverse: Easily Install and Load the “Tidyverse”* (Version 2.0.0) [Computer software]. <https://cran.r-project.org/web/packages/tidyverse/index.html>
- Wright, K. M., Lloyd, D., Lowry, D. B., Macnair, M. R., & Willis, J. H. (2013). Indirect Evolution of Hybrid Lethality Due to Linkage with Selected Locus in *Mimulus guttatus*. *PLoS Biology*, 11(2), e1001497. <https://doi.org/10.1371/journal.pbio.1001497>
- Zuellig, M. P., & Sweigart, A. L. (2018a). A two-locus hybrid incompatibility is widespread, polymorphic, and active in natural populations of *Mimulus*. *Evolution*, 72(11), 2394–2405. <https://doi.org/10.1111/evo.13596>
- Zuellig, M. P., & Sweigart, A. L. (2018b). Gene duplicates cause hybrid lethality between sympatric species of *Mimulus*. *PLoS Genetics*, 14(4), e1007130. <https://doi.org/10.1101/201392>

Table 4.1. Correlations between ancestry frequencies and gene density or recombination.

<i>M. nasutus</i> Ancestry frequency Sample group	Correlation (r)		% variance explained (ANOVA)	
	Gene density	Recombination rate	Gene density	Recombination rate [†]
CAC area <i>M. nasutus</i>	0.0937	0.0282	0.88***	0.03
DPR sympatric <i>M. nasutus</i>	0.0831	-0.0024	0.69***	0.01
DPR sympatric <i>admixed</i>	0.3334	0.0609	11.11***	0.05
LM <i>admixed</i>	0.3460	0.0644	11.97***	0.06
CAC <i>admixed</i>	0.3011	0.0869	9.07***	0.28*
CAC <i>M. guttatus</i>	0.2839	0.0845	8.06***	0.27*
DPR sympatric <i>M. guttatus</i>	0.3272	0.0854	10.71***	0.23*
DPR allopatric <i>M. guttatus</i>	0.1293	0.0317	1.67***	0.03
Gene density vs. Recombination rate		0.1142		
***p<0.001, *p<0.05 after Bonferroni correction				
[†] Added to model after gene density; if added first, variance explained is significant at p<0.001 for all <i>admixed</i> and sympatric <i>M. guttatus</i> groups				

Table 4.2. Summary of overlap between outlier windows across sample groups.

Bottom 5% ancestry frequency		Top 5% ancestry frequency	
Number of 50kb windows	Sample groups with window as outlier*	Number of 50kb windows	Sample groups with window as outlier [†]
3440	0	3689	0
645	1	452	1
256	2	142	2
21	3	46	3
7	4	30	4
0	5	9	5
0	6	1	6
0	7		
169	Both northern and southern groups	104	Both northern and southern groups
*Out of seven tested sample groups		[†] Out of six tested sample groups.	

Figure 4.1. Geographic distribution of collection locations and admixture proportions. A) Location of northern (CAC and LM) and southern (DPR area) collection areas. Details of the CAC and LM sites can be found in Chapter III. B) Map of collection locations in the southern DPR area, including three high-elevation sites with higher admixture levels. Classification into ancestry groups is based on hybrid index: $HI < 0.15$ are counted as *M. guttatus*, $HI > 0.85$ as *M. nasutus*, and $HI 0.15-0.85$ as *admixed*. Pie chart areas are proportional to the number of samples from each location. Elevation data is from the Shuttle Radar Topography Mission (SRTM) dataset, downloaded using the ‘elevation_3s’ function in the R package ‘geodata’ (Hijmans et al., 2024). C) Genomic PCA of all northern and southern samples using 17,514 SNP markers. Samples are colored by hybrid index as in (B) and (C). PC1 separates *M. guttatus* samples and, to a lesser extent, *M. nasutus* samples, into northern and southern groups; PC2 separates samples by *M. guttatus* vs. *M. nasutus* ancestry. A cluster of seven putative F1 hybrids from high-elevation southern locations are circled. D) Distribution of ancestry proportions (hybrid index, with 0=*M. guttatus* and 1=*M. nasutus*) for samples from each collection area. DPR area allopatric group includes any location from the southern region where *M. nasutus* or *admixed* individuals were not sampled; all other southern locations are treated as sympatric.

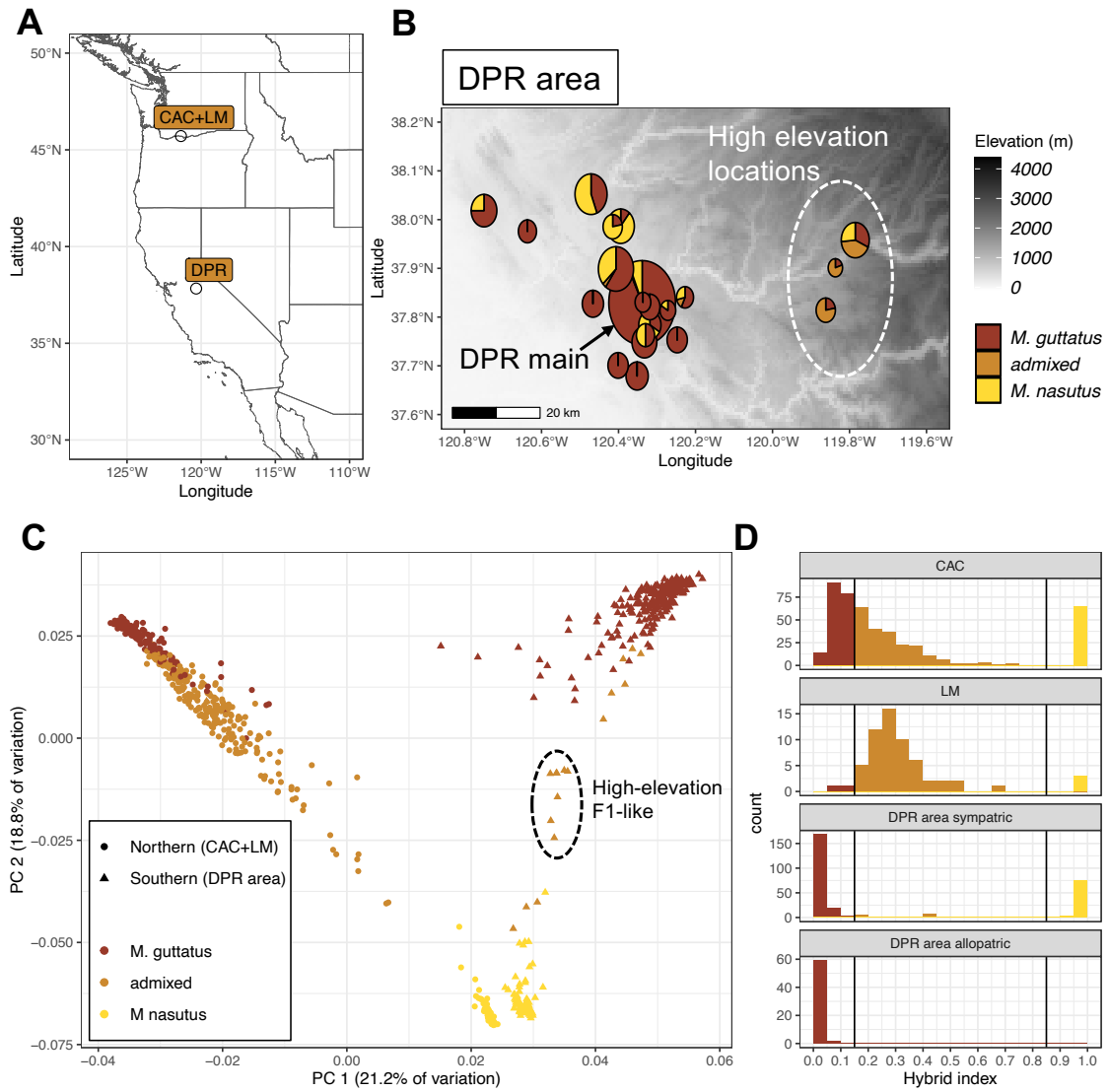
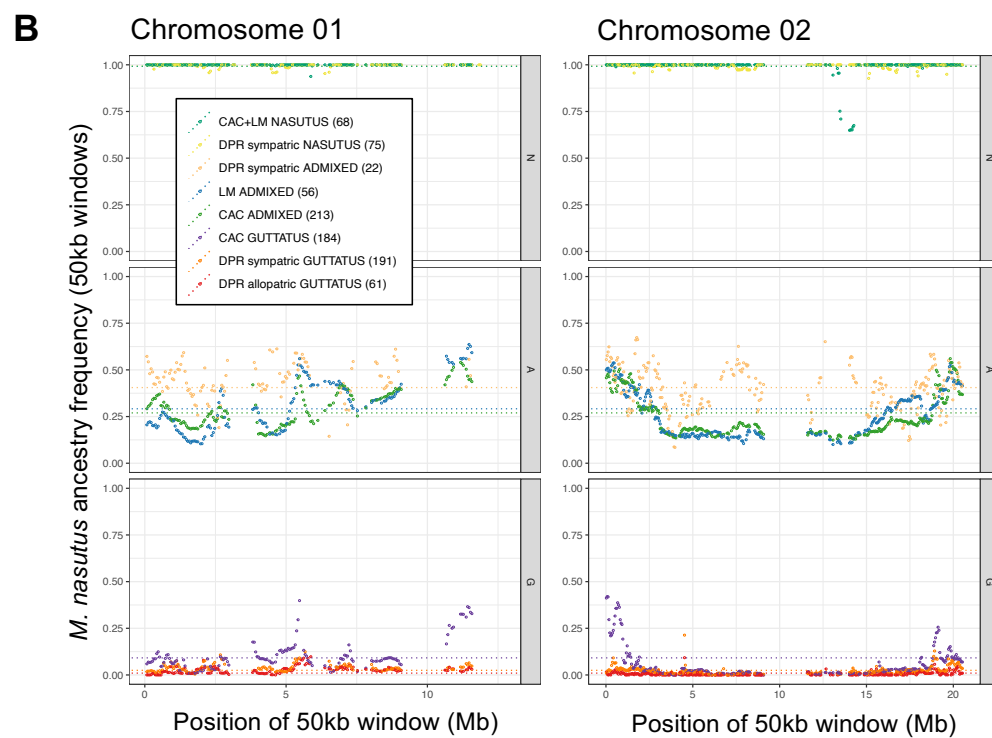
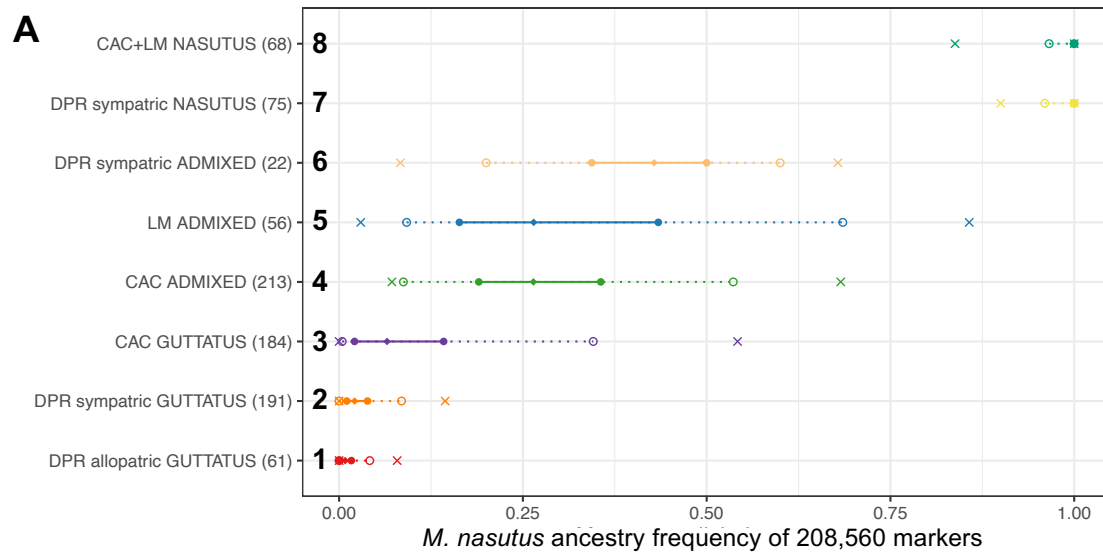


Figure 4.2. Distributions of ancestry frequencies across the genome for eight sample groups. A) Range of ancestry frequencies for 208,560 ancestry loci across the genome in each of eight sample groups, equivalent to the proportion of *M. nasutus* alleles relative to *M. guttatus* alleles in the population at a given locus. Diamonds indicate median ancestry frequency across all loci, with additional points representing the 25%/75% quantiles (closed circles), 5%/95% quantiles (open circles), and minimum/maximum values (X's). Numbers in parentheses indicate the sample sizes of each sample group. B-C) Ancestry frequencies in 50kb windows along chromosome 01 (panel B) and chromosome 02 (panel C) for each of eight sample groups. Each color represents a sample group; groups are organized into *M. guttatus* (G), *admixed* (A), and *M. nasutus* (N) cohorts. Horizontal dotted lines indicate the mean ancestry frequency for each sample group. Plots for the remaining chromosomes are provided in Figures 4.5 and S4.2-S4.3.



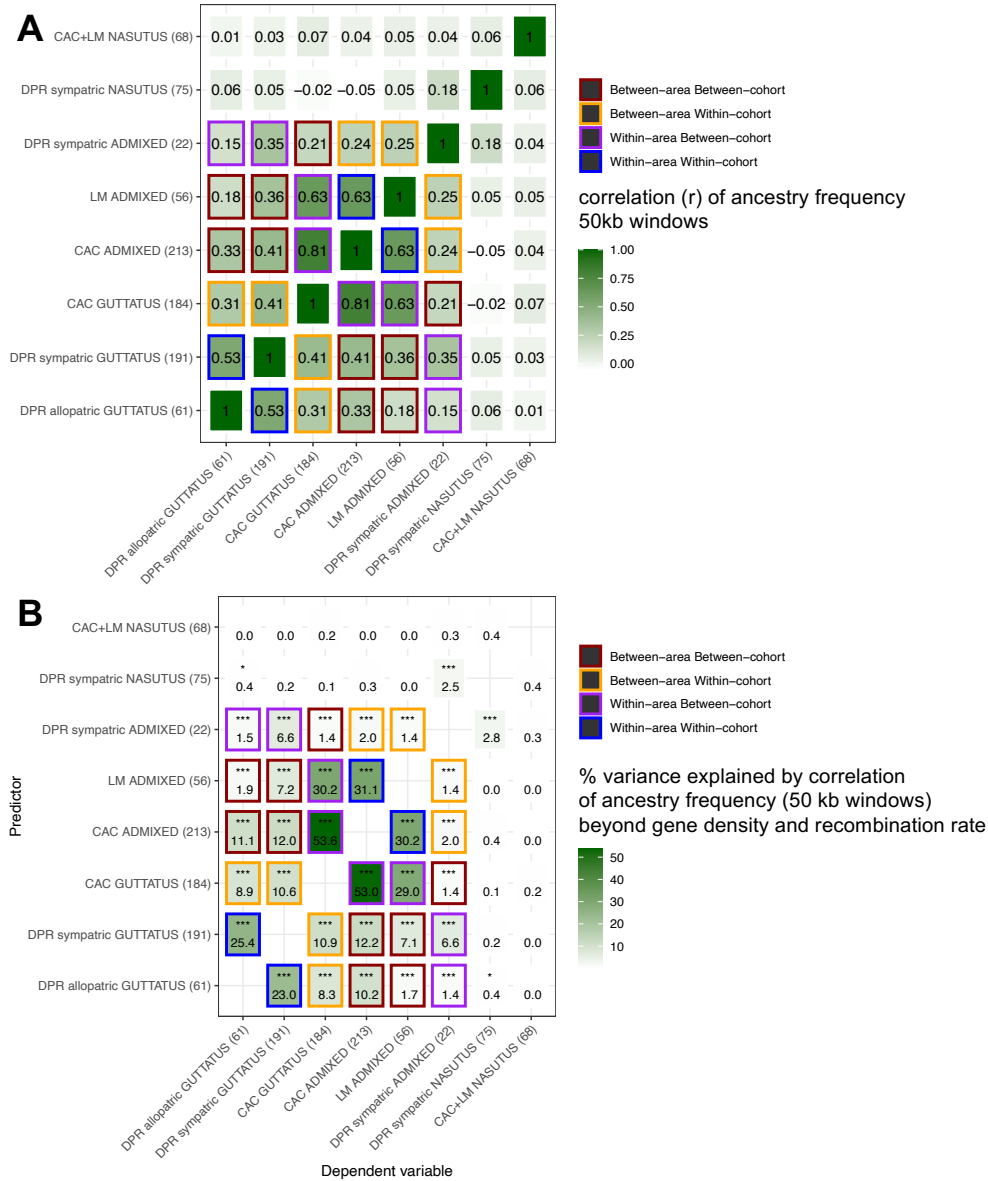


Figure 4.3. Correlations in ancestry frequencies across the genome between sample groups.

A) Correlation (r) values between ancestry frequencies of 2,739 50kb windows across the genome for each pair of sample groups. All included windows have at least 5 sites with at least 50% of samples called in every sample group. Colored outlines indicate whether a comparison is being made within or between areas (northern vs. southern) and within or between cohorts (*M. guttatus* vs. *admixed*); comparisons involving *M. nasutus* groups are not outlined. Numbers in parentheses indicate sample size of each group. B) Percent of variance in ancestry frequency of one sample group (dependent variable) that is explained by the ancestry frequency of another sample group (predictor), after accounting for variance explained by gene density and recombination rate. Ancestry frequency is calculated in 2,739 50kb windows. Variance is estimated using an ANOVA from a linear model fit of Dependent variable ~ Gene density + Recombination rate + Predictor. Separate models are run for each pair of groups.

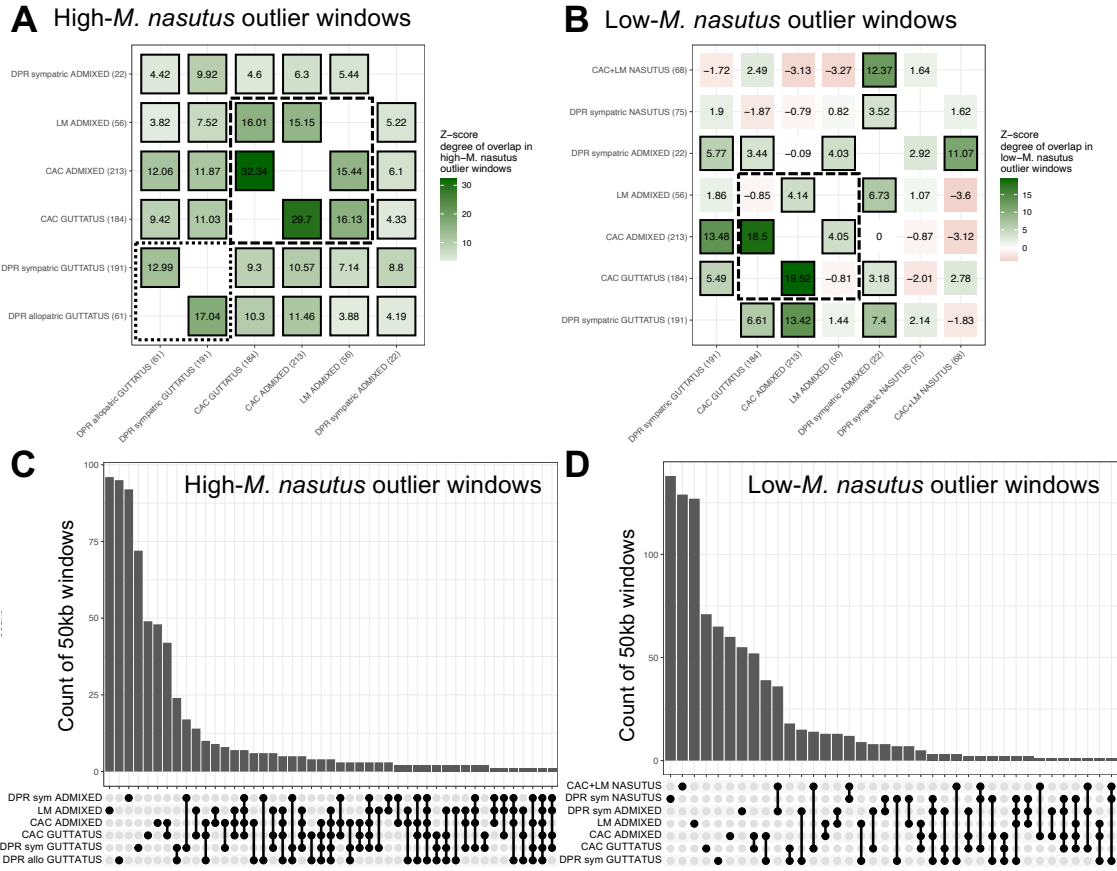


Figure 4.4. Overlap in ancestry outlier windows across sample groups.

A) Overlap in high-*M. nasutus* outlier windows. Outliers are defined as windows with ancestry frequency higher than the genome-wide 95% quantile value for a given sample group, with overlap scored as the number of windows that are outliers for both sample groups. Each value is a z-score representing to what degree two sets of windows overlap more often than expected by chance. Z-scores are calculated as the deviation of the true number of overlapping windows from the mean of the overlap between one sample group (y-axis) and 100 random permutations of the second group (x-axis), scaled by the standard deviation of the permuted values. Comparisons with a p-value < 0.05, after Bonferroni correction for 30 tests, are outlined in solid black. As a reference point, the dashed box indicates northern – northern *M. guttatus* and *admixed* comparisons; the dotted box indicates southern *M. guttatus sympatric* – *allopatric* comparisons.

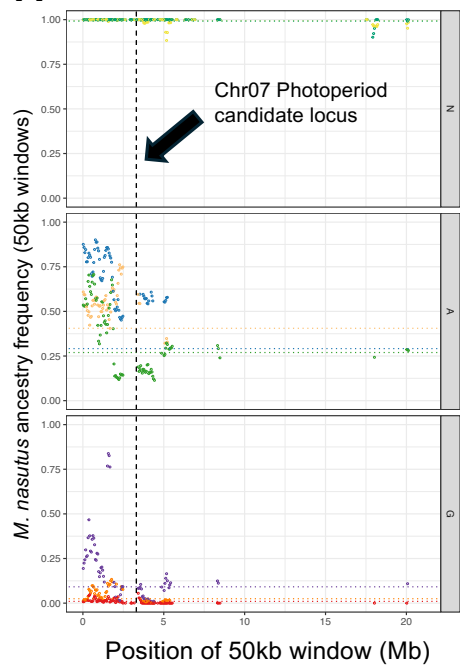
B) Overlap in low-*M. nasutus* outlier windows, defined as windows with ancestry frequency lower than the genome-wide 5% quantile value for a given sample group. Each value is a z-score as in (A). Comparisons with a p-value < 0.05, after Bonferroni correction for 42 tests, are outlined in black. As a reference point, the dashed box indicates northern – northern *M. guttatus* and *admixed* comparisons.

C) Upset plot showing the number of high-*M. nasutus* outlier windows that are unique or shared between each sample group. Each column shows a set of one or more sample groups, indicated by the connected dots, and the bars indicate the number of windows shared by exactly that set of sample groups.

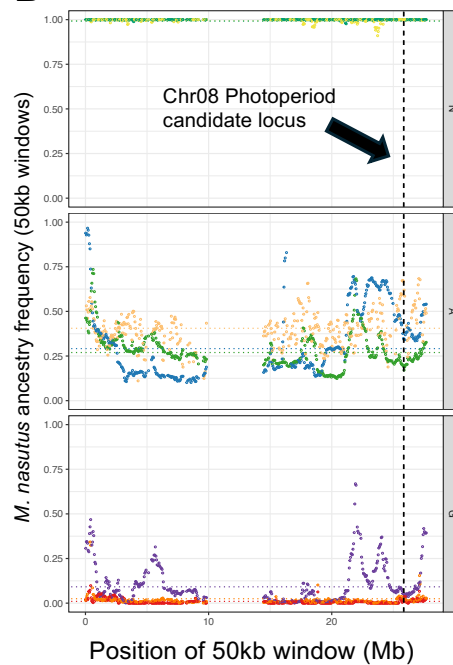
D) Upset plot as in (C), showing the number of low-*M. nasutus* outlier windows that are unique or shared between each set of sample groups.

Figure 4.5. Distributions of ancestry frequencies around reproductive barrier candidate loci. *M. nasutus* ancestry frequencies in 50kb windows along chromosome 07 (panel A), chromosome 08 (panel B), chromosome 13 (panel C), and chromosome 14 (panel D) for eight sample groups, organized into *M. guttatus* (G), *admixed* (A), and *M. nasutus* (N) cohorts. Horizontal dotted lines indicate the mean ancestry frequency for each sample group. Vertical dashed lines indicates the location of the Chr07 and Chr08 photoperiod candidate loci identified in (Fishman et al., 2014), and the *hll3* and *hll4* lethality incompatibility pair identified in (Zuellig & Sweigart, 2018a, 2018b). The expected directions of selection for *hll3* (G-, or increased *M. nasutus* frequency) and *hll4* (N-, or decreased *M. nasutus* frequency) are noted.

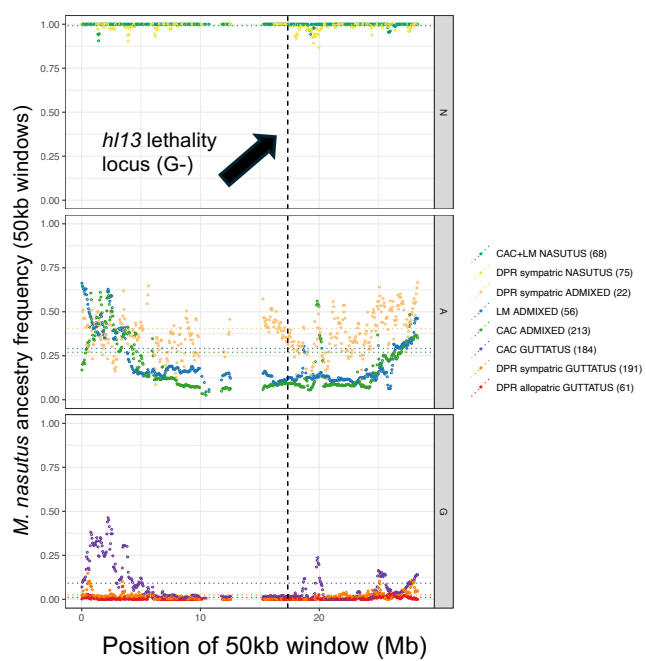
A Chromosome 07



B Chromosome 08



C Chromosome 13



D Chromosome 14

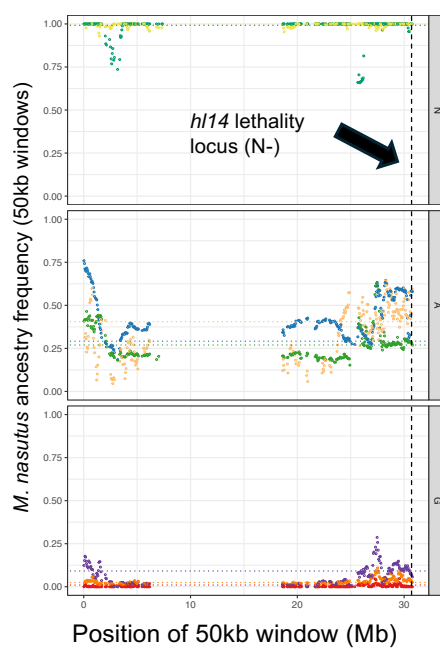
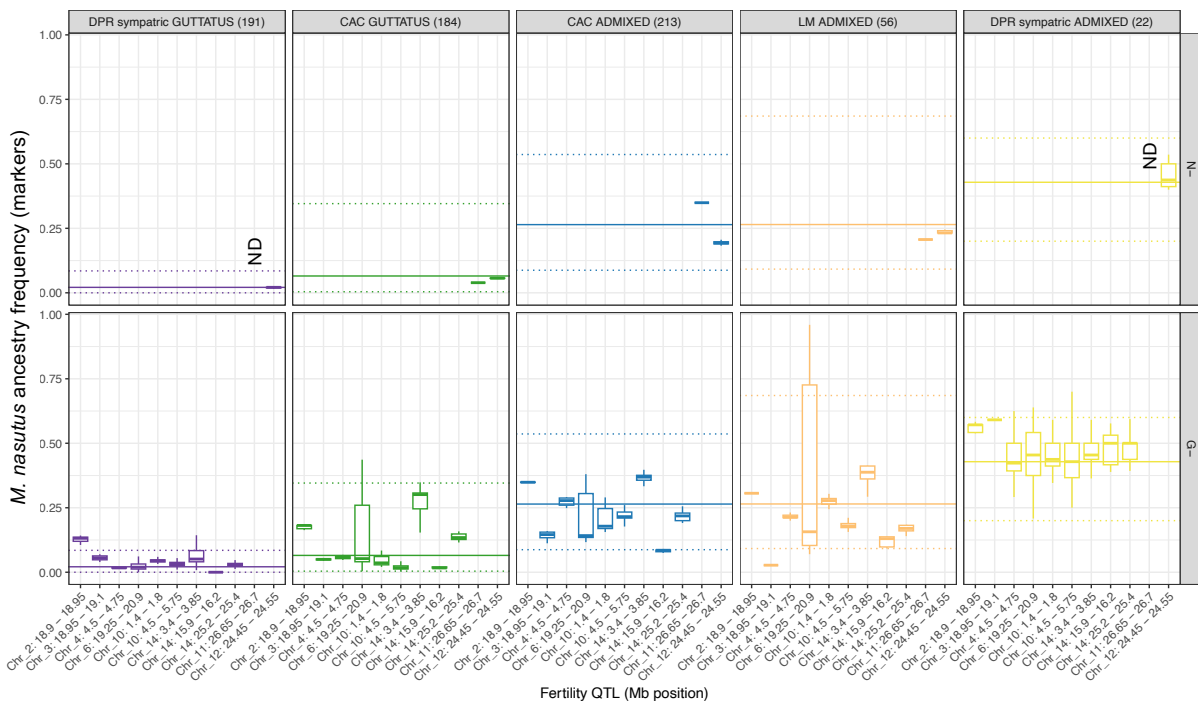
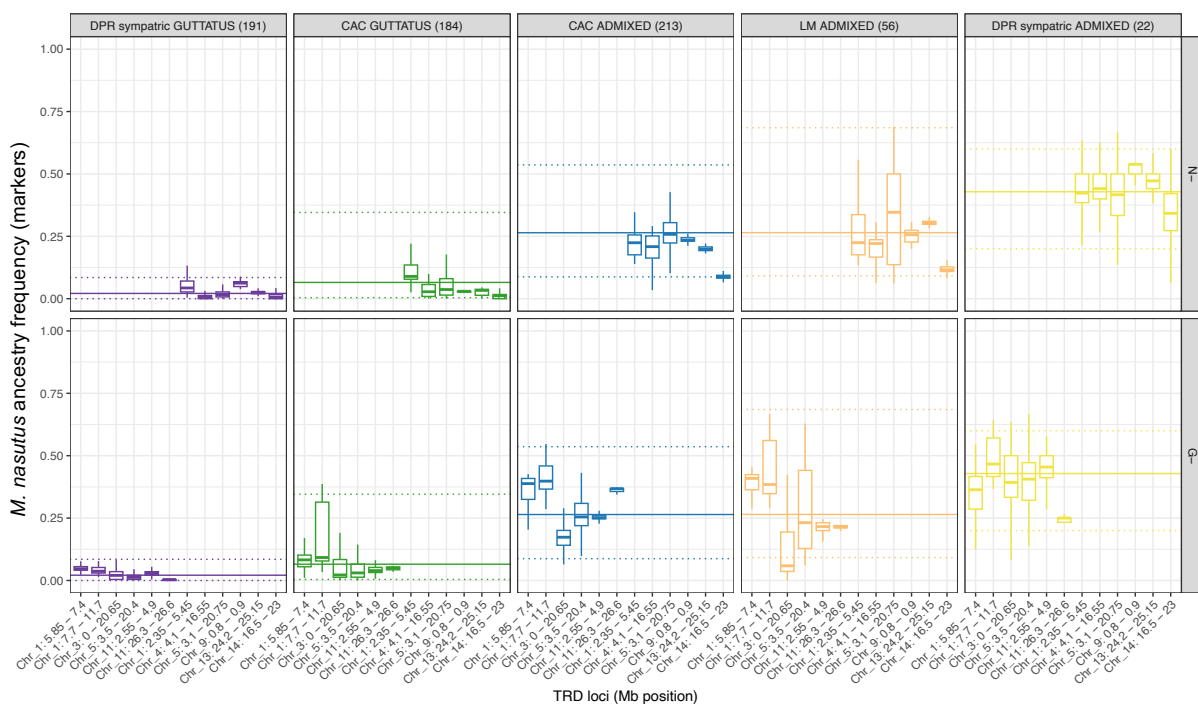


Figure 4.6. Ancestry frequencies at fertility QTL and TRD loci. A) *M. nasutus* ancestry frequencies at ancestry-informative markers within eleven fertility QTL regions identified in (Mantel & Sweigart, 2024). Frequencies shown for five sample groups with substantial genome-wide variation in admixture. Nine QTL (G- loci, lower panel) had *M. guttatus* alleles associated with lower fertility, while two QTL (N- loci, upper panel) had *M. nasutus* alleles associated with lower fertility. Boxplots indicate the median and quartiles of ancestry frequencies for ancestry loci within the range of each QTL. ‘ND’ indicates the locus contained no markers with data for that sample group. Horizontal solid lines indicate the median ancestry frequency across all markers for each sample group; horizontal dotted lines indicate the corresponding 5th and 95th percentiles of ancestry frequency. B) *M. nasutus* ancestry frequencies within twelve transmission ratio distortion (TRD) loci identified in (Mantel & Sweigart, 2024); each locus was identified as part of a pair with multi-locus distortion involving opposite ancestries. For six loci (N- loci), the *M. nasutus* allele was part of the underrepresented two-locus genotype, while for six loci (G- loci), the *M. guttatus* allele was part of the underrepresented two-locus genotype. Boxplots indicate the median and quartiles of ancestry frequencies for ancestry loci within the range of each locus. Horizontal solid lines indicate the median ancestry frequency of each sample group; horizontal dotted lines indicate the corresponding 5th and 95th percentiles of ancestry frequency.

A Fertility QTL



B Transmission Ratio Distortion loci



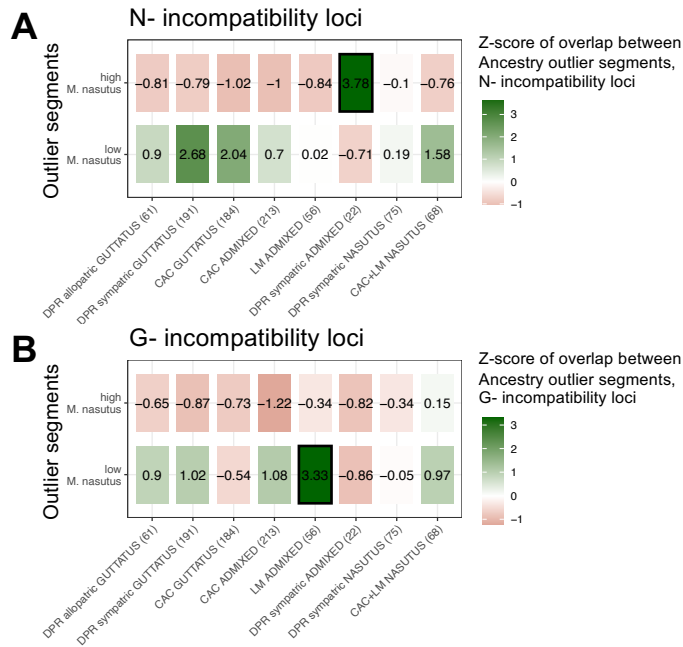


Figure 4.7. Overlap between ancestry outliers and fertility QTL or TRD loci. A) Z-scores representing the degree of length overlap between ancestry outlier segments and two N- fertility QTL or six TRD loci from (Mantel & Sweigart, 2024). Outliers were defined as segments of at least 10 ancestry loci with ancestry frequency higher than the genome-wide 95% quantile value (top row) or lower than the genome-wide 5% quantile value (bottom row) for a given sample group. Two fertility QTL have lower fertility in the *M. nasutus* genotype, and six TRD loci have *M. nasutus* alleles in an underrepresented two-locus genotype. B) Z-scores representing the degree of length overlap between ancestry outlier segments and 15 G- loci from (Mantel & Sweigart, 2024), including nine QTL with lower fertility in the *M. guttatus* genotype, and six TRD loci with *M. guttatus* alleles in an underrepresented two-locus genotype. Each value is a z-score representing the deviation of the true overlap length from the mean of the overlap between the fertility/TRD loci and 100 random permutations of the outlier segments, scaled by the standard deviation of the permuted values. Comparisons with a p-value < 0.05, after Bonferroni correction for 16 tests, are outlined in black.

CHAPTER V

CONCLUSIONS AND FUTURE DIRECTIONS

In three studies, I have characterized variation in reproductive isolation and hybridization across different scales in order to improve our understanding of the speciation process through a genomic lens. Speciation has been a subject of intense interest since Darwin (1859), and genomic studies of speciation are not new (Campbell et al., 2018; Payseur & Rieseberg, 2016). Still, both the process of speciation itself and the structure of genomes in natural populations are exceedingly complex and have eluded attempts to fully characterize them. Speciation is a process that typically takes place over thousands of years, so in our lifetimes we can only study snapshots of this process, or the signatures it leaves behind. Similarly, genomes are the product of millions of years of continuous evolution and increasing complexity, and we have only scratched the surface of how genomic processes play out across diverse organisms. Much of the genomic work on speciation to date has been either theoretical, applying mathematical models to processes unfolding in time, or primarily descriptive, sequencing organisms and identifying patterns their genomes. While we often try to draw causal links between ecological and evolutionary processes and the patterns we see in natural populations, rigorously testing these hypotheses is difficult.

Studying natural variation is an important way to address relationships between potential ecological or genomic driving forces and their evolutionary outcomes. Variation can be described at many different scales across space, across time, across species, or across loci in the genome. Here, I focus on variation in reproductive isolation and hybridization across related

species (Chapter II), across both time and small spatial scales (Chapter III), and across both wider geographic space and genomic space (Chapter IV). While much of this work is necessarily descriptive, I attempt to utilize comparisons across these scales to develop and test specific hypotheses about the nature of speciation.

*Comparative reproductive isolation across *Mimulus* taxa*

In Chapter II, I expand on our knowledge of reproductive isolation in the model genus *Mimulus* by quantifying divergence and postmating isolation in an understudied group of species from section *Eunanus*. This opens up opportunities to make broader comparisons across taxa regarding the repeatability and consistency of speciation. In contrast to model species complexes in *Mimulus*, I do not find evidence for ongoing introgression between species (Kenney & Sweigart, 2016; Streisfeld & Kohn, 2005) or the cooccurrence of distinct, incompletely isolated groups (Ivey et al., 2023; Streisfeld & Kohn, 2005). Instead, I find that these species have substantial divergence and strong postmating reproductive isolation. I do detect a cryptic lineage (Sespe Creek), but this group appears strongly isolated by postmating barriers in its own right. Consistent with a number of other *Mimulus* taxa (Coughlan et al., 2020; Oneal et al., 2016; Sandstedt et al., 2020), I find that hybrid seed inviability is an important reproductive barrier in multiple species pairs. However, my finding of parental seed size differences associated with hybrid seed inviability is novel and may have implications for the mechanisms behind this postmating barrier.

In *Mimulus* and other angiosperms, parental conflict between maternal and paternal alleles has been implicated as a selective driver of seed inviability (Coughlan et al., 2020; Lafon-Placette & Köhler, 2016; Raunsgard et al., 2018; Sandstedt & Sweigart, 2022), but despite reciprocal differences in hybrid seed size, these crosses do not usually have clear parental seed

size differences. Seed sizes in *Mimulus* section *Eunanus* might be under differential ecological selection due to the rocky scree habitats of the larger-seeded species (Baker, 1972; Cordazzo, 2002). I hypothesize that this selection could trigger developmental changes in similar endosperm-related pathways to cases of parental-conflict-mediated seed inviability, but resulting from a different selective process. Testing this hypothesis is difficult because elucidating the genetic basis of hybrid seed inviability or seed size requires hybrids, which do not readily form. However, future work could measure the strength of selection on seed survival and germination under contrasting soil and slope conditions in an experimental context or with reciprocal germination tests at natural sites. Further sampling of related species could reveal additional candidate pairs for genetic mapping, with seed size differences but weaker isolation. The *Eunanus* species complex as a whole is an untapped wealth of diversity for ecological, genomic, and speciation questions.

Another thread for future work in section *Eunanus* is the genetics of premating isolation, particularly flower color. *M. fremontii* and *M. brevipes* can be found in close proximity but are at least partly isolated by postmating barriers, implying that premating barriers could have evolved due to reinforcement to escape maladaptive hybridization (Hopkins, 2013). Indeed, they have very different flower size, shape, and color (purple vs. yellow), indicating different pollinator preferences, but there is no data on which pollinators are important for these species or whether pollinator isolation is a strong barrier. *Mimulus* section *Eunanus* has multiple other transitions between purple and yellow flowers, including polymorphism within species (*M. mephiticus*) (Baldwin et al., 2012), making it an intriguing system to study parallelism in the mechanisms of color divergence. In preliminary crosses, I have found that multigenerational hybrids show a remarkable suite of transgressive color phenotypes, including variation in pattern, hue, and

shade, reminiscent of transgressive hybrids in Chilean *Mimulus* crosses (Simmons et al., 2023). I have even found cases of different colorations on flowers within the same hybrid plant, suggestive of breakdown in the stability of gene regulatory networks. None of these questions have been properly explored as of yet, but they are intriguing points for further investigation.

Changes across space and time in hybrid populations

In Chapter III, I utilize fine-scaled spatial sampling across a decade to investigate the drivers of stability and fluctuation in ancestry within hybrid populations. While some other studies have investigated hybrid zone changes over time, they are mostly in animal systems with traditional clinal hybrid zones (Moore & Buchanan, 1985; Sullivan, 1995; Taylor et al., 2014). In contrast, *Mimulus guttatus* and *Mimulus nasutus* are closely related and typically cross-compatible plants with a mosaic geographic distribution, separated by microhabitat and mating system differences rather than landscape-level clines. We find multiple intriguing patterns in this analysis: the geographic scale of ancestry structure is incredibly small (20-50m between distinct cohorts), premating barriers are quite variable across years, and yet ancestry appears stable across time. The apparent contradiction between variable isolation and stable structure suggests a population in a dynamic equilibrium, where partial isolation is not sufficient to drive complete homogenization or complete isolation. This finding is reminiscent of within-population models of balancing selection, in which heterogeneous selection across space and time promotes increased genetic diversity by periodically favoring different suites of low-frequency variants (Abdul-Rahman et al., 2021; Delph & Kelly, 2014). The role of balancing selection in the maintenance of admixed ancestry within hybrid zones is a topic that warrants further exploration.

While many findings in this hybrid zone are intriguing, it is difficult to draw firm conclusions about long-term temporal processes from just a few years of data, especially given

gaps and inconsistencies of sampling. We intend to continue sampling these populations for additional years, generating a unique long-term dataset of admixture dynamics across time. My study has led to the formulation of clear hypotheses that we can test more rigorously in the future. Water availability across the growing season should predict the extent and overlap in flowering times between admixed cohorts, with more overlap in wetter years. The overlap in flowering times should in turn predict the extent of observed shifts in offspring ancestry. If postmating isolation is not strong, ancestry patterns in one year should reflect the offspring ancestry in the prior year. Where this is not the case, comparing allele frequencies in one year to offspring from the prior year might reveal candidate postmating barrier loci. In addition to testable hypotheses, my study highlights other forms of data that would be useful for this study area in the future. Targeted dense sampling could be used to examine the relative spatial patterns of pollen and seed flow across the landscape. Careful measurements of water availability across the season will be useful for examining the effects of year-to-year and microsite-to-microsite variation. This long-term dataset will be an important resource for understanding change over time at fine scales in hybridizing systems.

Parallelism in genomic ancestry patterns and lack of a signal for candidate loci

In Chapter IV, I use hybrid ancestry across two disparate locations to answer two major questions: 1) how similar are patterns of ancestry across the genome in replicate hybrid zones, and 2) are natural patterns of ancestry predicted by results of laboratory studies on the genetic basis of reproductive barriers.

We find substantial similarity across space in genomic patterns. At large geographic distances, similarity likely reflects selective forces acting in parallel across populations. At more proximal locations, even stronger similarity suggests connectivity via migration, including

between sympatric and allopatric locations. These patterns appear to go beyond genome-scale processes of linked selection tied to recombination rate variation, and extend to patterns at specific loci. However, there is more parallelism for putatively adaptive loci than for loci under negative selection; this might reflect a bias of the overall genomic landscape, in which adaptive loci are more easily detected, but could reflect differences in the genetic architecture of adaptive introgression vs. reproductive isolation. More complete sampling at additional locations, such as the high-elevation locations that apparently have quite distinct hybridization dynamics, would help disentangle the patterns we see from their root causes. Also, more work could be done to clean up this data and address potential biases, such as possible incorrect ancestry calls at particular sites or reasons for the dropout of missing windows. Still, this study opens up many avenues to investigate putative introgressed or introgression-resistant loci, particularly those shared across locations. Careful analysis using gene tree methods and sequence divergence, potentially with higher-coverage targeted resequencing of these areas, as well as tests for multi-locus interactions, would help to confirm their introgression status and evolutionary history, as well as point to potential reasons for their selective patterns.

A recent analysis of the classic *Mus musculus* hybrid zone suggests that natural variation is not necessarily predicted by the results of laboratory crosses for reproductive isolation (Frayer & Payseur, 2024). I find a similar result: few of the known candidate loci or QTL have expected patterns of reduced or elevated introgression in the wild. The *Mus* analysis suggests a number of possible reasons for this: incompatibilities may be effectively purged from hybrid zones prior to observation, so their selective effects are no longer visible; the incompatibilities that are noticed in the lab may not be the most important ones in nature; and a complex and polymorphic architecture of incompatibility may dilute the signal from any one analysis or dataset. In

Mimulus, the incompatibilities we tested were detected in the exact populations we studied (Mantel & Sweigart, 2024; Zuellig & Sweigart, 2018), so purging is not a likely explanation. However, we know that these incompatibilities are highly polymorphic, varying among lines from the same population, and in many cases having relatively weak quantitative effects (Mantel & Sweigart, 2024; Zuellig & Sweigart, 2018). We also tested a pair of premating isolating barriers, which was not the focus of (Frayser & Payseur, 2024), but found a similar lack of signal. We know that flowering time is an important barrier to reproduction, but it may not be under direct selection in a way that is detectable via introgression signals. Furthermore, the alleles present in northern locations may be different from southern locations where the relationship between habitat suitability and photoperiod is shifted.

Going forward in this system, there are additional candidate loci that we could test in a similar manner: minor QTL for floral size differences (Fishman et al., 2002) and for pollen precedence (Fishman et al., 2008) could provide interesting signals. More careful analysis of the specific alleles present at candidate loci in these natural populations would be useful, as well as a direct comparison of genotype with natural flowering time at Catherine Creek. As a broader speciation research field, we need a better theoretical understanding of the expectations for genomic patterns under more complex scenarios of polymorphism, linked selection, and heterogeneous environmental selection. Finally, this study highlights the need for complementary approaches to the same evolutionary questions from different data sets and different research angles, in order to get a more holistic view of the interaction between genes, genomes, and the environment during speciation.

On the importance of considering the environment in speciation

A throughline of this work is that environmental heterogeneity can produce variation in reproductive isolation and hybridization outcomes, both between and within taxa. This matches other recent work showing that environmental heterogeneity, both across space and across time, is a key factor mediating the strength of isolation (Sianta et al., 2024; Tataru et al., 2023). Climate change is expected to increase environmental heterogeneity (Rind et al., 1989), as well as pose novel selective pressures and invite new opportunities for hybridization (Bowler et al., 2015; Devictor et al., 2012; Franks & Weis, 2009; Muhlfeld et al., 2014). A better understanding of the impacts of environmental variation on hybridization outcomes will be necessary to predict and plan for ecosystem change in the 21st century.

References

- Abdul-Rahman, F., Tranchina, D., & Gresham, D. (2021). Fluctuating Environments Maintain Genetic Diversity through Neutral Fitness Effects and Balancing Selection. *Molecular Biology and Evolution*, 38(10), 4362–4375. <https://doi.org/10.1093/molbev/msab173>
- Baker, H. G. (1972). Seed Weight in Relation to Environmental Conditions in California. *Ecology*, 53(6), 997–1010. <https://doi.org/10.2307/1935413>
- Baldwin, B. G., Goldman, D., Keil, D. J., Patterson, R., Rosatti, T. J., & Wilken, D. (Eds.). (2012). *The Jepson Manual: Vascular Plants of California, Thoroughly Revised and Expanded* (2nd ed.).
- Bowler, D. E., Haase, P., Kröncke, I., Tackenberg, O., Bauer, H. G., Brendel, C., Brooker, R. W., Gerisch, M., Henle, K., Hickler, T., Hof, C., Klotz, S., Kühn, I., Matesanz, S., O'Hara, R., Russell, D., Schweiger, O., Valladares, F., Welk, E., ... Böhning-Gaese, K. (2015). A cross-taxon analysis of the impact of climate change on abundance trends in central Europe. *Biological Conservation*, 187, 41–50. <https://doi.org/10.1016/j.biocon.2015.03.034>
- Campbell, C. R., Poelstra, J. W., & Yoder, A. D. (2018). What is Speciation Genomics? The roles of ecology, gene flow, and genomic architecture in the formation of species. *Biological Journal of the Linnean Society*, 124(4), 561–583. <https://doi.org/10.1093/biolinnean/bly063>

- Cordazzo, C. V. (2002). Effect of seed mass on germination and growth in three dominant species in southern Brazilian coastal dunes. *Brazilian Journal of Biology*, 62, 427–435. <https://doi.org/10.1590/S1519-69842002000300005>
- Coughlan, J. M., Wilson Brown, M., & Willis, J. H. (2020). Patterns of Hybrid Seed Inviability in the *Mimulus guttatus* sp. Complex Reveal a Potential Role of Parental Conflict in Reproductive Isolation. *Current Biology*, 30(1), 83–93.e5. <https://doi.org/10.1016/j.cub.2019.11.023>
- Darwin, C. (1859). *On the origin of species by means of natural selection*. John Murray. http://darwin-online.org.uk/converted/pdf/1861_OriginNY_F382.pdf
- Delph, L. F., & Kelly, J. K. (2014). On the importance of balancing selection in plants. *New Phytologist*, 201(1), 45–56. <https://doi.org/10.1111/nph.12441>
- Devictor, V., van Swaay, C., Brereton, T., Brotons, L., Chamberlain, D., Heliölä, J., Herrando, S., Julliard, R., Kuussaari, M., Lindström, Å., Reif, J., Roy, D. B., Schweiger, O., Settele, J., Stefanescu, C., Van Strien, A., Van Turnhout, C., Vermouzek, Z., WallisDeVries, M., ... Jiguet, F. (2012). Differences in the climatic debts of birds and butterflies at a continental scale. *Nature Climate Change*, 2(2), 121–124. <https://doi.org/10.1038/nclimate1347>
- Fishman, L., Aagaard, J., & Tuthill, J. C. (2008). Toward the Evolutionary Genomics of Gametophytic Divergence: Patterns of Transmission Ratio Distortion in Monkeyflower (*mimulus*) Hybrids Reveal a Complex Genetic Basis for Conspecific Pollen Precedence. *Evolution*, 62(12), 2958–2970. <https://doi.org/10.1111/j.1558-5646.2008.00475.x>
- Fishman, L., Kelly, A. J., & Willis, J. H. (2002). Minor quantitative trait loci underlie floral traits associated with mating system divergence in *Mimulus*. *Evolution*, 56(11), 2138–2155. <https://doi.org/10.1111/j.0014-3820.2002.tb00139.x>
- Franks, S. J., & Weis, A. E. (2009). Climate change alters reproductive isolation and potential gene flow in an annual plant. *Evolutionary Applications*, 2(4), 481–488. <https://doi.org/10.1111/j.1752-4571.2009.00073.x>
- Frayar, M. E., & Payseur, B. A. (2024). Do genetic loci that cause reproductive isolation in the lab inhibit gene flow in nature? *Evolution*, 78(6), 1025–1038. <https://doi.org/10.1093/evolut/qpae044>
- Hopkins, R. (2013). Reinforcement in plants. *New Phytologist*, 197(4), 1095–1103. <https://doi.org/10.1111/nph.12119>
- Ivey, C. T., Habecker, N. M., Bergmann, J. P., Ewald, J., Frayer, M. E., & Coughlan, J. M. (2023). Weak reproductive isolation and extensive gene flow between *Mimulus glaucescens* and *M. guttatus* in northern California. *Evolution*, 77(5), 1245–1261. <https://doi.org/10.1093/evolut/qpae044>

- Kenney, A. M., & Sweigart, A. L. (2016). Reproductive isolation and introgression between sympatric *Mimulus* species. *Molecular Ecology*, 25(11), 2499–2517. <https://doi.org/10.1111/mec.13630>
- Lafon-Placette, C., & Köhler, C. (2016). Endosperm-based postzygotic hybridization barriers: Developmental mechanisms and evolutionary drivers. *Molecular Ecology*, 25(11), 2620–2629. <https://doi.org/10.1111/mec.13552>
- Mantel, S. J., & Sweigart, A. L. (2024). Postzygotic barriers persist despite ongoing introgression in hybridizing *Mimulus* species. *Molecular Ecology*, 33(4), e17261. <https://doi.org/10.1111/mec.17261>
- Moore, W. S., & Buchanan, E. B. (1985). Stability of the northern flicker hybrid zone in historical times: Implications for adaptive speciation theory. *Evolution*, 39(1), 135–151. <https://doi.org/10.1111/j.1558-5646.1985.tb04086.x>
- Muhlfeld, C. C., Kovach, R. P., Jones, L. A., Al-Chokhachy, R., Boyer, M. C., Leary, R. F., Lowe, W. H., Luikart, G., & Allendorf, F. W. (2014). Invasive hybridization in a threatened species is accelerated by climate change. *Nature Climate Change*, 4(7), 620–624. <https://doi.org/10.1038/nclimate2252>
- Oneal, E., Willis, J. H., & Franks, R. G. (2016). Disruption of endosperm development is a major cause of hybrid seed inviability between *Mimulus guttatus* and *Mimulus nudatus*. *New Phytologist*, 210(3), 1107–1120. <https://doi.org/10.1111/nph.13842>
- Payseur, B. A., & Rieseberg, L. H. (2016). A genomic perspective on hybridization and speciation. *Molecular Ecology*, 25(11), 2337–2360. <https://doi.org/10.1111/mec.13557>
- Raunsgard, A., Opedal, Ø. H., Ekrem, R. K., Wright, J., Bolstad, G. H., Armbruster, W. S., & Pélabon, C. (2018). Intersexual conflict over seed size is stronger in more outcrossed populations of a mixed-mating plant. *Proceedings of the National Academy of Sciences*, 115(45), 11561–11566. <https://doi.org/10.1073/pnas.1810979115>
- Rind, D., Goldberg, R., & Ruedy, R. (1989). Change in climate variability in the 21st century. *Climatic Change*, 14(1), 5–37. <https://doi.org/10.1007/BF00140173>
- Sandstedt, G. D., & Sweigart, A. L. (2022). Developmental evidence for parental conflict in driving *Mimulus* species barriers. *New Phytologist*, 236(4), 1545–1557. <https://doi.org/10.1111/nph.18438>
- Sandstedt, G. D., Wu, C. A., & Sweigart, A. L. (2020). Evolution of multiple postzygotic barriers between species of the *Mimulus tilingii* complex. *Evolution*, 75, 600–613. <https://doi.org/10.1111/evo.14105>
- Sianta, S. A., Moeller, D. A., & Brandvain, Y. (2024). The extent of introgression between incipient *Clarkia* species is determined by temporal environmental variation and mating system. *Proceedings of the National Academy of Sciences*, 121(12), e2316008121. <https://doi.org/10.1073/pnas.2316008121>

- Simmons, E. S. G., Cooley, A. M., Puzey, J. R., & ConradiSmith, G. D. (2023). A Multigenerational Turing Model Reproduces Transgressive Petal Spot Phenotypes in Hybrid *Mimulus*. *Bulletin of Mathematical Biology*, 85(12), 120. <https://doi.org/10.1007/s11538-023-01223-7>
- Streisfeld, M. A., & Kohn, J. R. (2005). Contrasting patterns of floral and molecular variation across a cline in *Mimulus aurantiacus*. *Evolution*, 59(12), 2548–2559. <https://doi.org/10.1111/j.0014-3820.2005.tb00968.x>
- Sullivan, B. K. (1995). Temporal stability in hybridization between *Bufo microscaphus* and *Bufo woodhousii* (Anura: Bufonidae): Behavior and morphology. *Journal of Evolutionary Biology*, 8(2), 233–247. <https://doi.org/10.1046/j.1420-9101.1995.8020233.x>
- Tataru, D., Wheeler, E. C., & Ferris, K. G. (2023). Spatially and temporally varying selection influence species boundaries in two sympatric *Mimulus*. *Proceedings of the Royal Society B: Biological Sciences*, 290(1992), 20222279. <https://doi.org/10.1098/rspb.2022.2279>
- Taylor, S. A., Curry, R. L., White, T. A., Ferretti, V., & Lovette, I. (2014). Spatiotemporally consistent genomic signatures of reproductive isolation in a moving hybrid zone. *Evolution*, 68(11), 3066–3081. <https://doi.org/10.1111/evo.12510>
- Zuellig, M. P., & Sweigart, A. L. (2018). A two-locus hybrid incompatibility is widespread, polymorphic, and active in natural populations of *Mimulus*. *Evolution*, 72(11), 2394–2405. <https://doi.org/10.1111/evo.13596>

APPENDIX A

SUPPLEMENTARY TABLES AND FIGURES FROM CHAPTER II

Table S2.1. Population location information and number of samples used per population.

Pop.	Species	Latitude	Longitude	Samples sequenced	Seed families	1 st gen	2 nd gen
J31	<i>M. johnstonii</i>	34.401950	-117.81580	5	4	9	7
B07	<i>M. brevipes</i>	34.358990	-118.39803	3	5	14	1
B11	<i>M. brevipes</i>	33.656155	-117.45298	6	5	13	7
B12	<i>M. brevipes</i>	32.889920	-116.57440	4	6	23	2
B19	<i>M. brevipes</i>	33.333300	-116.94200	4	6	33	20
F06	<i>M. fremontii</i>	34.693000	-119.32900	2	4	9	--
F21	<i>M. fremontii</i>	34.460560	-117.67316	3	4	11	--
F22	<i>M. fremontii</i>	34.376390	-117.59682	2	3	6	--
S25	'Sespe Creek'	34.565773	-119.26081	2	1	3	3
N01	<i>M. nanus</i>	42.638395	-118.57760	2	--	--	--
C01	<i>M. constrictus</i>	34.798960	-119.00466	1	--	--	--

† 1st gen = wild-collected, 2nd gen = greenhouse-produced from hand pollination. All sequenced individuals were grown directly from wild-collected seeds, except the two *M. nanus* samples, which were tissue collected from wild individuals.

Table S2.2. Sequencing and coverage information.

Sample info			Read pairs			Coverage per base	
Sample [†]	Pop	Batch	Raw	Aligned	%	All sites	>=4X coverage
BREV07_3_1	B07	1	47,522,461	13,277,523	27.9%	12.47	39.03%
BREV07_5_1	B07	1	49,800,600	15,174,402	30.5%	14.04	39.09%
BREV12_3_1	B12	1	49,704,577	15,840,128	31.9%	15.68	37.55%
BREV12_5_1	B12	1	49,785,466	14,465,999	29.1%	13.52	39.79%
BREV12_6_1	B19	1	49,555,797	15,518,193	31.3%	14.54	39.49%
BREV19_1_1	B19	1	49,652,489	16,050,274	32.3%	15.62	37.83%
FREM22_6_1	F22	1	49,332,251	8,451,160	17.1%	7.30	33.81%
JOHN31_5_1	J31	1	49,714,033	13,727,590	27.6%	12.82	37.51%
BREV07_6_1	B07	2	25,682,314	5,672,576	22.1%	5.67	32.30%
BREV11_5_3	B11	2	28,911,480	5,804,484	20.1%	5.35	31.82%
BREV12_4_1	B12	2	28,052,293	6,443,496	23.0%	6.28	33.38%
BREV19_2_1	B19	2	29,339,237	7,002,395	23.9%	6.87	33.83%
BREV19_3_1	B19	2	36,190,791	3,248,019	9.0%	3.10	21.10%
BREV19_4_1	B19	2	14,858,533	2,429,889	16.4%	2.19	16.18%
FREM06_3_2	F06	2	30,075,452	5,408,560	18.0%	4.51	26.45%
FREM06_4_3	F06	2	29,784,495	6,422,159	21.6%	5.68	29.74%
FREM21_2_1	F21	2	33,601,408	5,451,858	16.2%	4.60	27.10%
FREM22_5_1	F22	2	24,172,306	2,679,095	11.1%	2.35	19.37%
JOHN31_1_1	J31	2	47,347,007	6,441,313	13.6%	6.20	31.22%
JOHN31_2_1	J31	2	64,312,829	10,195,170	15.9%	9.88	36.43%
JOHN31_4_1	J31	2	23,147,570	3,938,270	17.0%	3.82	24.04%
JOHN31_5_2	J31	2	32,146,268	5,382,624	16.7%	4.86	25.48%
BREV11_1_11	B11	3	12,194,294	3,311,065	27.2%	3.16	18.64%
BREV11_3_4	B11	3	16,149,111	4,165,590	25.8%	3.88	24.46%
BREV11_5_1	B11	3	19,609,920	5,018,757	25.6%	4.72	26.70%
BREV11_6_11	B11	3	17,240,947	4,349,611	25.2%	4.02	26.63%
FREM21_3_11	F21	3	31,307,714	5,636,788	18.0%	4.86	30.25%
FREM21_4_11	F21	3	23,442,361	4,186,587	17.9%	3.67	24.90%
SESP25_1_1	S25	3	37,881,366	8,233,912	21.7%	7.70	34.70%
SESP25_1_2	S25	3	22,013,442	5,056,136	23.0%	4.58	29.93%
NANU01_1	N01	3	15,545,612	2,612,978	16.8%	2.26	14.32%
NANU01_2	N01	3	14,729,812	2,603,428	17.7%	2.38	17.61%
CONS01_1_1	C01	4	16,909,453	4,072,713	24.1%	4.10	16.6%
AURA_ari	SRX6077155		15,314,328	5,630,656	36.8%	7.14	64.06%
AURA_aur	SRX6077153		18,453,062	7,584,963	41.1%	10.02	75.92%
AURA_gra	SRX6077242		12,454,896	6,061,103	48.7%	7.84	66.98%
AURA_pun	SRX6077404		14,056,153	5,076,335	36.1%	6.53	60.90%
CLEV_cle	SRX6077240		12,810,520	4,567,552	35.7%	5.58	53.01%

[†] 1st number indicates population, 2nd number indicates maternal family, 3rd number refers to the individual plant. *M. nanus* samples (NANU01_1 and NANU01_2) are distinct wild plants.

Table S2.3. Pedigree information for *M. brevipes* and *M. fremontii* F1 families.

Line name	Line type	Maternal	Paternal	Number of individuals used		
				Pollen viability	Maternal crosses	Paternal crosses
BF1	BxF F1 hybrid	B19_3_2	F6_4_3	0	1	1
BF2	BxF F1 hybrid	B12_5_2	F21_3_1	2	1	1
BF3	BxF F1 hybrid	B19_1_1	F6_3_1	7	7	6
BF5	BxF F1 hybrid	B19_4_14	F21_4_11	4	2	0

Table S2.4. Nucleotide diversity and divergence values at synonymous sites. Values were calculated using pixy for all individual pairwise comparisons, then averaged across comparisons within a group.

Species	Heterozygosity	Diversity
AUR	0.0025	0.0163
B	0.0198	0.0229
C	0.0135	--
F	0.0209	0.0316
J	0.0170	0.0174
N	0.0154	0.0324
S	0.0180	0.0178

Species 1	Species 2	Divergence
AUR	C	0.0940
AUR	N	0.0945
AUR	S	0.0991
AUR	B	0.1000
AUR	J	0.1005
AUR	F	0.1027
B	J	0.0525
B	F	0.0599
B	S	0.0697
B	C	0.0707
B	N	0.0830
C	S	0.0605
C	J	0.0702
C	F	0.0730
C	N	0.0753
F	J	0.0598
F	S	0.0715
F	N	0.0860
J	S	0.0692
J	N	0.0828
N	S	0.0791

Table S2.5. ABBA-BABA test results for the ‘complete’ SNP dataset.

P1	P2	P3	D	Z-score	Significance	<i>f4</i>-ratio	BBA	ABBA	BABA
B	J	F	0.025	4.1	***	0.017	140199	101734	96787
B	J	S	0.077	14.5	***	0.023	225194	81493	69780
J	F	S	0.012	2.5	n.s.	0.004	193112	89515	87397
B	F	S	0.080	16.0	***	0.027	192818	93743	79920
B	J	C	0.060	12.4	***	0.015	217482	60440	53601
J	F	C	0.001	0.3	n.s.	0.000	190562	65810	65623
B	F	C	0.054	11.7	***	0.016	189827	68457	61448
C	S	J	0.193	38.6	***	0.072	128359	98268	66479
C	S	F	0.199	39.8	***	0.093	129273	101021	67419
C	S	B	0.186	35.9	***	0.077	133849	94966	65145
B	J	N	0.058	13.7	***	0.011	310955	59579	53098
F	J	N	0.027	8.0	***	0.006	281234	65746	62337
B	F	N	0.024	7.2	***	0.005	280189	64934	61873
J	S	N	0.091	24.8	***	0.027	184994	88233	73578
B	S	N	0.127	28.9	***	0.037	179236	92602	71726
F	S	N	0.110	28.9	***	0.032	190288	90971	73014
J	C	N	0.099	23.1	***	0.032	126808	82206	67332
B	C	N	0.131	25.7	***	0.042	124234	85599	65732
F	C	N	0.114	23.8	***	0.037	129720	84553	67229
S	C	N	0.036	7.9	***	0.010	166409	65159	60632
*** $p_{\text{corr}} < 0.001$, n.s.=not significant, p_{corr} is Bonferroni-corrected for 20 tests.									

Table S2.6. ABBA-BABA test results from four randomly downsampled datasets.

P1	P2	P3	D			<i>f4</i> -ratio			Significant iterations
			min	median	max	min	median	max	
C	S	F	0.149	0.172	0.178	0.059	0.064	0.068	4
C	S	J	0.142	0.163	0.169	0.049	0.055	0.058	4
C	S	B	0.135	0.155	0.162	0.048	0.053	0.055	4
F	C	N	0.16	0.179	0.18	0.048	0.053	0.055	4
B	C	N	0.138	0.164	0.191	0.042	0.05	0.056	4
J	C	N	0.14	0.154	0.16	0.042	0.045	0.049	4
F	S	N	0.124	0.14	0.153	0.034	0.037	0.041	4
B	S	N	0.106	0.124	0.156	0.029	0.034	0.041	4
J	S	N	0.102	0.114	0.124	0.028	0.03	0.033	4
S	C	N	0.057	0.065	0.087	0.015	0.017	0.024	4
B	F	S	-0.006	0.044	0.068	-0.002	0.014	0.021	3
B	J	S	-0.001	0.043	0.076	0	0.012	0.021	3
J	B	F	-0.014	0.018	0.064	-0.007	0.009	0.033	1
B	J	C	-0.01	0.034	0.064	-0.002	0.008	0.015	3
F	J	N	0.03	0.035	0.039	0.007	0.007	0.008	4
B	F	C	-0.027	0.022	0.043	-0.007	0.006	0.011	2
B	J	N	-0.02	0.027	0.062	-0.004	0.005	0.011	3
F	J	C	0.005	0.011	0.018	0.001	0.003	0.005	0
F	B	N	-0.018	0.007	0.057	-0.004	0.001	0.012	1
J	F	S	-0.005	0.003	0.007	-0.002	0.001	0.002	0

Significance for each iteration determined via Z-score from an ABBA-BABA test, Bonferroni-corrected for 20 tests, $p_{corr} < 0.001$. Significant tests support introgression from P3 into P2.

Bolded values were significant in the same direction for all four downsampled iterations.

Table S2.7. Window-based *df* and diversity metrics for introgression tests.

Quartet C,S;F,A					
Window group	Mean <i>df</i>	Median <i>df</i>	Mean CONS:pi	Mean SESP:pi	Mean FREM:pi
All windows	0.164	0.147	0.0069	0.0115	0.0143
All windows, filtered for pi>0†	0.159	0.142	0.0106	0.0143	0.0169
Top <i>df</i> windows‡	0.715	0.699	0.0059	0.0087	0.0117
pi decile§	Mean <i>df</i>	pi decile§	Mean <i>df</i>	pi decile§	Mean <i>df</i>
CONS:pi 1 st	0.181	SESP:pi 1 st	0.185	FREM:pi 1 st	0.187
CONS:pi 2nd	0.186	SESP:pi 2nd	0.181	FREM:pi 2nd	0.195
CONS:pi 3rd	0.189	SESP:pi 3rd	0.197	FREM:pi 3rd	0.185
CONS:pi 4th	0.184	SESP:pi 4th	0.182	FREM:pi 4th	0.175
CONS:pi 5th	0.174	SESP:pi 5th	0.165	FREM:pi 5th	0.189
CONS:pi 6th	0.165	SESP:pi 6th	0.170	FREM:pi 6th	0.173
CONS:pi 7th	0.150	SESP:pi 7th	0.160	FREM:pi 7th	0.140
CONS:pi 8th	0.153	SESP:pi 8th	0.138	FREM:pi 8th	0.141
CONS:pi 9th	0.121	SESP:pi 9th	0.125	FREM:pi 9th	0.113
CONS:pi 10th	0.090	SESP:pi 10th	0.088	FREM:pi 10th	0.094
Quartet B,C;N,A					
Window group	Mean <i>df</i>	Median <i>df</i>	Mean BREV:pi	Mean CONS:pi	Mean NANU:pi
All windows	0.112	0.108	0.0123	0.0065	0.0112
All windows, filtered for pi>0†	0.108	0.105	0.0145	0.0100	0.0132
Top <i>df</i> windows‡	0.564	0.536	0.0110	0.0065	0.0112
pi decile§	Mean <i>df</i>	pi decile§	Mean <i>df</i>	pi decile§	Mean <i>df</i>
BREV:pi 1st	0.118	CONS:pi 1 st	0.137	NANU:pi 1st	0.105
BREV:pi 2nd	0.110	CONS:pi 2nd	0.124	NANU:pi 2nd	0.101
BREV:pi 3rd	0.130	CONS:pi 3rd	0.111	NANU:pi 3rd	0.118
BREV:pi 4th	0.117	CONS:pi 4th	0.122	NANU:pi 4th	0.125
BREV:pi 5th	0.134	CONS:pi 5th	0.111	NANU:pi 5th	0.136
BREV:pi 6th	0.115	CONS:pi 6th	0.108	NANU:pi 6th	0.115
BREV:pi 7th	0.118	CONS:pi 7th	0.092	NANU:pi 7th	0.104
BREV:pi 8th	0.098	CONS:pi 8th	0.108	NANU:pi 8th	0.109
BREV:pi 9th	0.085	CONS:pi 9th	0.094	NANU:pi 9th	0.092
BREV:pi 10th	0.052	CONS:pi 10th	0.070	NANU:pi 10th	0.074

† Windows where pi=0 for any of the three species were removed.

‡ Positive *df* values from the 100 windows with the highest absolute value of *df*.

§ 1st decile includes the 10% of windows with the lowest pi, 10th decile includes the 10% of windows with the highest pi, etc. after removal of windows with pi=0.

Table S2.8. Crossing success and seed viability for each cross type.

Maternal species	Paternal species	Cross type	Crossing success (total crosses)	Seed viability (total seeds)
<i>M. johnstonii</i>	<i>M. johnstonii</i>	JxJ	0.538 (39)	0.745 (306)
<i>M. johnstonii</i>	<i>M. brevipes</i>	JxB	0.440 (25)	0.000 (98)
<i>M. johnstonii</i>	<i>M. fremontii</i>	JxF	0.600 (10)	0.000 (78)
<i>M. johnstonii</i>	Sespe Creek	JxS	0.600 (5)	0.015 (66)
<i>M. brevipes</i>	<i>M. johnstonii</i>	BxJ	0.484 (31)	0.006 (630)
<i>M. brevipes</i>	<i>M. brevipes</i>	BxB	0.902 (51)	0.981 (7292)
<i>M. brevipes</i>	<i>M. fremontii</i>	BxF	0.694 (36)	0.821 (2226)
<i>M. brevipes</i>	Sespe Creek	BxS	0.500 (6)	0.018 (217)
<i>M. fremontii</i>	<i>M. johnstonii</i>	FxJ	0.500 (20)	0.012 (246)
<i>M. fremontii</i>	<i>M. brevipes</i>	FxB	0.017 (58)	0.952 (21)
<i>M. fremontii</i>	<i>M. fremontii</i>	FxF	0.382 (55)	0.902 (911)
<i>M. fremontii</i>	Sespe Creek	FxS	0.091 (11)	0.000 (8)
Sespe Creek	<i>M. johnstonii</i>	SxJ	0.125 (8)	0.000 (4)
Sespe Creek	<i>M. brevipes</i>	SxB	0.333 (6)	0.000 (38)
Sespe Creek	<i>M. fremontii</i>	SxF	0.000 (4)	n/a (0)
Sespe Creek	Sespe Creek	SxS	0.409 (22)	0.909 (88)

Bolded values are significant in Tukey tests compared to both parental cross types.

Table S2.9. Tetrazolium staining results.

Maternal	Paternal	Red (Viable)	Pink (or mixed)	White (Inviably)	Proportion[†]	Expected viability
B	B	60	10	7	0.909	YES
B	F	48	3	1	0.981	YES
B	J	0	0	73	0	NO
B	S	0	2	62	0.031	NO
F	B	-	-	-	-	-
F	F	45	8	1	0.981	YES
F	J	0	0	48	0	NO
F	S	0	1	9	0.100	NO
J	B	0	0	27	0	NO
J	F	0	2	24	0.077	NO
J	J	12	2	8	0.636	YES
J	S	0	0	18	0	NO
S	B	0	0	13	0	NO
S	F	-	-	-	-	-
S	J	0	0	5	0	NO
S	S	15	0	5	0.750	YES

[†]Proportion = (red + pink)/total. FxB and SxF were not tested with tetrazolium due to low seed availability. Expected viability is based on visual inspection of seeds prior to staining.

Table S2.10. Average seed measurements and sample sizes by cross type.

Maternal species	Paternal species	Cross type	Seeds measured (number of fruits)	Average seed length (mm)	Average seed length/width ratio
<i>M. johnstonii</i>	<i>M. johnstonii</i>	JxJ	67 (7)	0.858	2.070
<i>M. johnstonii</i>	<i>M. brevipes</i>	JxB	50 (7)	0.865	4.729
<i>M. johnstonii</i>	<i>M. fremontii</i>	JxF	26 (3)	0.687	4.580
<i>M. johnstonii</i>	Sespe Creek	JxS	12 (1)	0.726	3.038
<i>M. brevipes</i>	<i>M. johnstonii</i>	BxJ	66 (5)	0.447	2.060
<i>M. brevipes</i>	<i>M. brevipes</i>	BxB	60 (4)	0.597	1.813
<i>M. brevipes</i>	<i>M. fremontii</i>	BxF	45 (3)	0.470	1.677
<i>M. brevipes</i>	Sespe Creek	BxS	56 (2)	0.523	2.155
<i>M. fremontii</i>	<i>M. johnstonii</i>	FxJ	9 (2)	0.579	2.901
<i>M. fremontii</i>	<i>M. brevipes</i>	FxB	15 (1)	0.584	1.989
<i>M. fremontii</i>	<i>M. fremontii</i>	FxF	27 (3)	0.641	2.149
<i>M. fremontii</i>	Sespe Creek	FxS	8 (1)	0.377	3.134
Sespe Creek	<i>M. johnstonii</i>	SxJ	22 (1)	0.415	3.274
Sespe Creek	<i>M. brevipes</i>	SxB	38 (2)	0.409	3.384
Sespe Creek	<i>M. fremontii</i>	SxF	15 (1)†	0.224†	2.965†
Sespe Creek	Sespe Creek	SxS	61 (6)	0.757	1.757

Bolded values are significant in Tukey tests compared to both parental cross types.

†Measured SxF ‘seeds’ were indistinguishable from unfertilized ovules.

Table S2.11. Pollen viability and counts with sample sizes.

Maternal species	Paternal species	Cross type	Average pollen viability (number of individuals) †	Pollen per unit area (number of individuals)
<i>M. brevipes</i>	<i>M. brevipes</i>	BxB	0.839 (13)	55.6 (13)
<i>M. brevipes</i>	<i>M. fremontii</i>	BxF	0.235 (12)	16.6 (13)
<i>M. fremontii</i>	<i>M. fremontii</i>	FxF	0.922 (8)	5.6 (10)

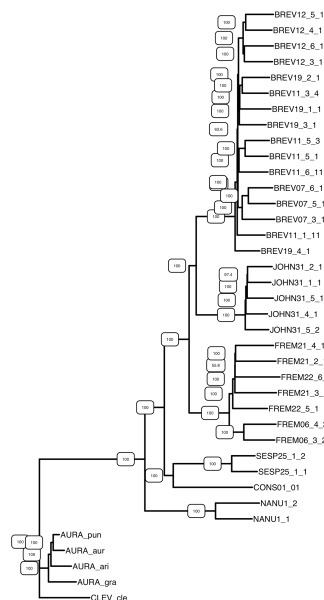
†excluding flowers with <10 counted pollen grains

Table S2.12. Statistical model structures and AIC values compared to null model equivalents.

Trait of interest	Model structure	Model type	AIC	AIC of null model †
Crossing success	Fruit_produced ~ Cross_type	BRGLM (Firth binomial)	467.4	576.2
Crossing success of BxF hybrids	Fruit_produced ~ Cross_type + (1 Maternal family)	GLMER (binomial)	253.3	291.7
Seed viability	cbind(viable_seeds, inviable_seeds) ~ Cross_type	BRGLM (Firth binomial)	1566.2	10533.6
Seed length (pure species only)	Length ~ Cross_type + (1 Fruit)	LMER	-600.8	-598.5
Seed length (all cross types)	Length ~ Cross_type + (1 Fruit)	LMER	-1131.6	-1141.3
Seed length/width ratio	Length/width ~ Cross_type + (1 Fruit)	LMER	1041.0	1122.1
Pollen counts	round(Pollen_per_square) ~ Cross_type + (1 Population)	GLMER (binomial)	435.2	438.2
Pollen viability	cbind(n_viable, n_inviable) ~ Cross_type + (1 Population)	GLMER (binomial)	491.0	501.2
Female fertility	Seed_count ~ Cross_type	LM	900.8	913.5
†null model is the equivalent model excluding Cross_type as a model term.				

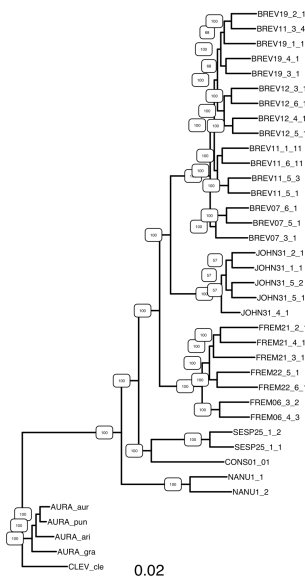
Figure S2.1. Comparison of phylogenetic methods. (A) Neighbor-joining tree produced from ‘complete’ SNP dataset, with heterozygous sites randomly assigned to one allele. (B) Maximum-likelihood tree produced using RAxML from the ‘complete’ SNP dataset using the GTR+Gamma model, with heterozygous sites randomly assigned to one allele, and an ascertainment bias correction with Felsenstein’s method. (C) ASTRAL tree produced from 17,524 gene trees, each created using RAxML with the GTR+Gamma model using SNPs from the ‘genic’ SNP dataset within a given gene’s coordinates, with heterozygous sites randomly assigned to one allele. Each node is labelled with the quartet support scores for the consensus, 1st alternative, and 2nd alternative topologies. (D) The same ASTRAL consensus tree shown in (C) but with posterior probabilities for each node.

A



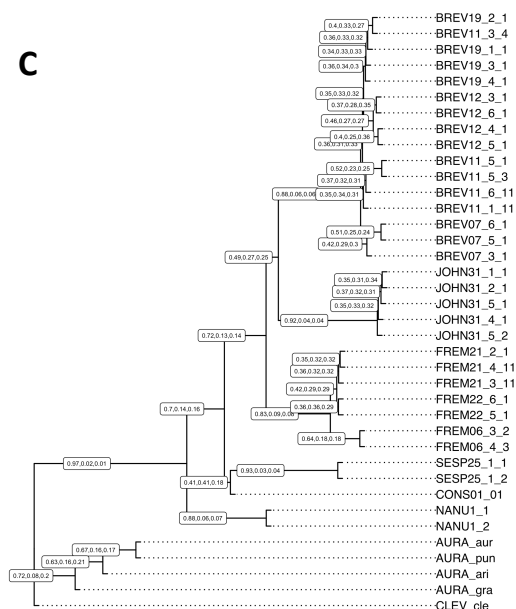
Neighbor-joining tree

B



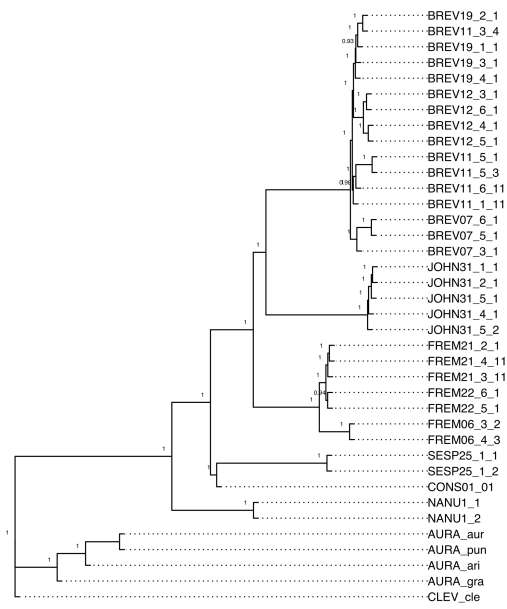
Maximum-likelihood tree

C



ASTRAL tree with quartet scores

D



ASTRAL tree with posterior probabilities

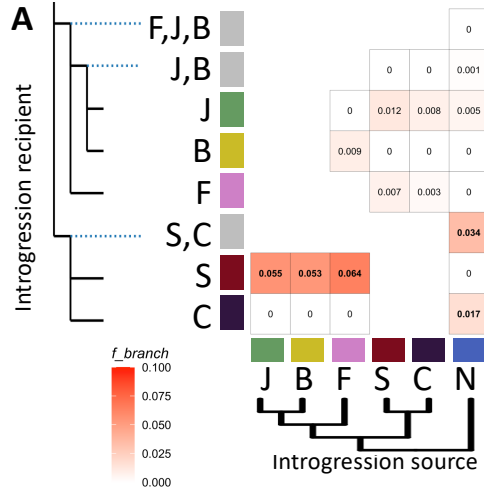





















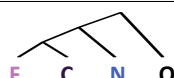







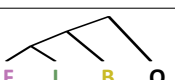


Figure S2.2. F_{branch} results for ‘downsampled’ iterations. Each value is the median F_{branch} value from among four independent downsampled datasets. Bolded values have nonzero F_{branch} values in all four downsampled datasets and were supported by significant ABBA-BABA tests ($p_{corr} < 0.001$); non-bolded values had $F_{branch} = 0$ for at least one dataset. Results are qualitatively consistent with the ‘complete’ dataset (Fig 3) except that the full dataset indicates introgression from *M. fremontii* into *M. johnstonii* whereas downsampled iterations have partial support for introgression into *M. brevipes* instead. Estimated F_{branch} values tend to be higher for the full dataset than the downsampled datasets. J = *M. johnstonii*, B = *M. brevipes*, F = *M. fremontii*, S = Sespe creek population, C = *M. constrictus*, N = *M. nanus*.

Figure S2.3. TWISST gene tree discordance summaries. Twenty trios were tested, using *M. aurantiacus* as the outgroup in all tests. Shown are the ten trios with significant overrepresentation of one alternate topology over another (binomial test, expected ratio 1:1, Bonferroni-corrected $p_{\text{corr}} < 0.05$). For each trio, the number of genes supporting each topology are given, along with the percentage of total trees. A total of 17,524 genes were used for all trios. J = *M. johnstonii*, B = *M. brevipes*, F = *M. fremontii*, S = Sespe creek population, C = *M. constrictus*, N = *M. nanus*, O = *M. aurantiacus* complex.

Consensus topology	Alternate topology 1	Alternate topology 2	Binomial Test p_{corr}
 7207 (41.1%)	 7330 (41.8%)	 2987 (17.0%)	<0.001
 7,253 (41.4%)	 7,299 (41.6%)	 2,972 (17.0%)	<0.001
 7,183 (41.4%)	 7,287 (41.6%)	 3,054 (17.0%)	<0.001
 12,874 (73.5%)	 2,667 (15.2%)	 1,983 (11.3%)	<0.001
 12,869 (73.4%)	 2,613 (14.9%)	 2,042 (11.7%)	<0.001
 12,828 (73.2%)	 2,638 (15.1%)	 2,059 (11.7%)	<0.001
 12,312 (70.3%)	 2,773 (15.8%)	 2,439 (13.9%)	<0.001
 11,114 (63.4%)	 3,350 (19.1%)	 3,061 (17.5%)	0.006
 12,328 (70.3%)	 2,744 (15.7%)	 2,452 (14.0%)	0.001
 8,528 (48.7%)	 4,673 (26.7%)	 4,323 (24.7%)	0.004

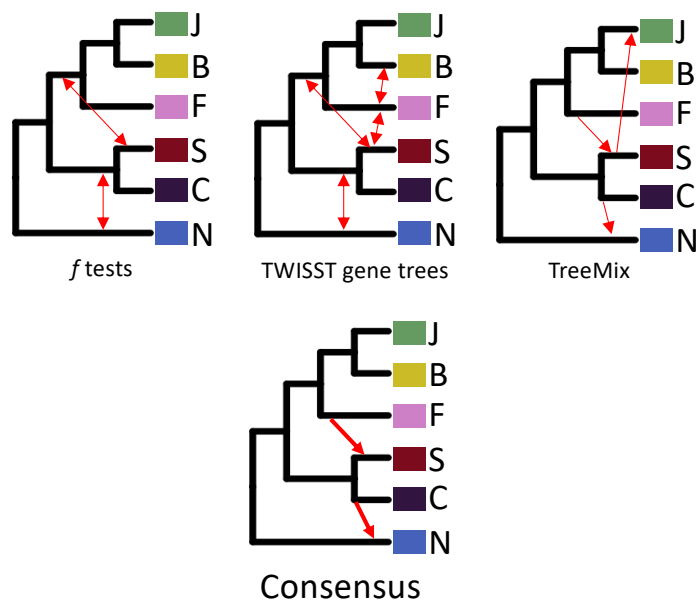


Figure S2.4. Summary of best-supported introgression events from multiple methods. Two introgression events (F → S and C → N) were supported by *f* tests, TWISST, and TreeMix results, with additional events supported by some but not all methods. For *f* tests, only results consistent across the ‘complete’ dataset and all downsampled iterations are included. Note that *f* tests and TWISST cannot determine directionality, which is inferred only using TreeMix. The exact identity and direction of individual introgression events is difficult to determine due to limited sampling and shared evolutionary history; for example, signatures of introgression from *M. fremontii*, *M. johnstonii*, and *M. brevipes* into Sespe Creek may be caused by a single event in a common ancestor, an unsampled relative, or just one of the sampled species. J = *M. johnstonii*, B = *M. brevipes*, F = *M. fremontii*, S = Sespe creek population, C = *M. constrictus*, N = *M. nanus*.

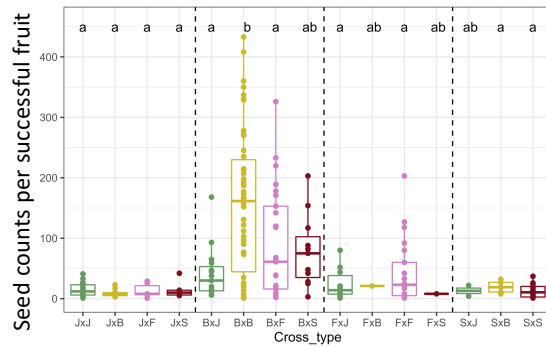


Figure S2.5. Seed counts for intra- and interspecific crosses. Seed counts per fruit for fruits producing at least one seed. Letters indicate significance in post-hoc Tukey tests from a linear model; cross types sharing a letter are not significantly different. *M. brevipes* produces more seeds per fruit than the other species. Interspecific crosses BxJ and BxF produced fewer seeds per fruit than the maternal parent (BxB), but were not different from the paternal parent (JxJ or FxF).

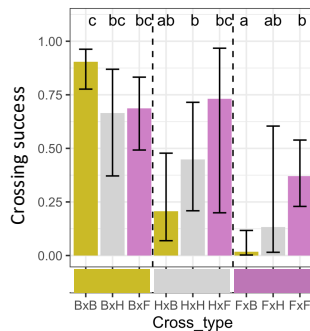


Figure S2.6. Crossing success in BxF hybrids is additive on maternal and paternal sides. Crossing success (probability of at least one seed produced from a cross) from crosses involving *M. brevipes* x *M. fremontii* F1 hybrids and their parental species. Shown are least-square means and asymptotic confidence intervals from a binomial GLMM. Letters indicate significance in post-hoc Tukey tests; cross types sharing a letter are not significantly different. HxB crosses have intermediate crossing success compared to BxB and FxB crosses (gold bars), indicating an additive effect of maternal species. Similarly, FxH crosses have intermediate crossing success compared to FxB and FxF crosses (right three bars), indicating an additive effect of paternal species. HxH have intermediate success compared to BxH and FxH (gray bars), as well as compared to HxB and HxF (middle three bars), further supporting additive patterns. B = *M. brevipes*, F = *M. fremontii*, H = F1 hybrids between *M. brevipes* and *M. fremontii*.

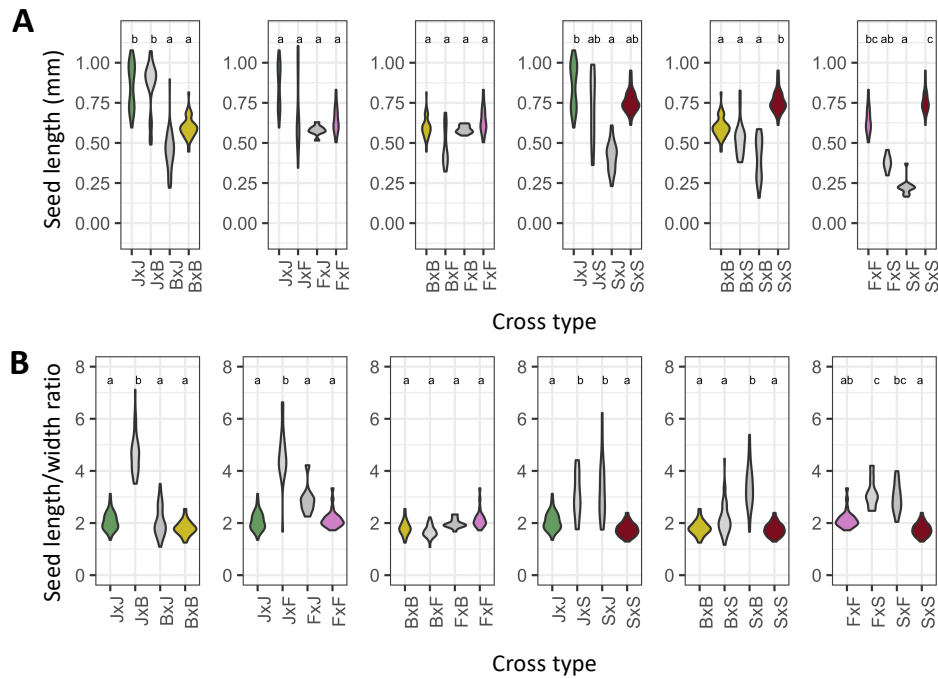


Figure S2.7. Seed measurements for all intra- and interspecific cross types. (A) Seed length and (B) length/width ratios. Violin plots show distribution of individual seeds. Letters indicate significance from post-hoc Tukey tests after running six independent species-pair models. One species pair had a significant reciprocal difference in hybrid seed size, while three pairs had a significant reciprocal difference in seed length/width ratio; length/width ratio was elevated compared to both parents for six cross types.

APPENDIX B

SUPPLEMENTARY TABLES AND FIGURES FROM CHAPTER III

Table S3.1. Summary of maternal and offspring samples by year, stream and cohort.

Breakdown of wild-collected (maternal) plants sequenced for this study.								
Year	Stream	Total maternal	Maternal <i>M. nasutus</i> (HI>0.95)	Maternal <i>M. guttatus</i> (HI<0.15)	Maternal <i>admixed</i> (HI 0.15-0.8)	Maternal <i>M. sookensis</i>		
2019	CAC_S1	126	29	38	59			
2021	CAC_S1	17	0	13	4			
2021	CAC_S2	11	0	7	4			
2021	LM	61	3	2	56			
2022	CAC_S1	122	19	48	55			
2022	CAC_S2	122	10	34	67	11		
2012	CAC_S1	23	3	4	16			
2012	CAC_S2	52	4	40	8			
Samples for 2012 are taken from a previous study (see Methods). Samples with <25,000 called ancestry-informative sites were excluded and are not counted here.								
Breakdown of maternal families and offspring sequenced for this study.*								
Year	Stream	Families with offspring data	Total Fruits	Mixed fruits (selfed and outcrossed offspring)	Total offspring	Selfed	Outcrossed	Ambiguous†
2019	CAC_S1	66	82	42	430	111	289	30
2021	CAC_S1	10	22	5	157	12	131	14
2021	CAC_S2	8	17	7	95	19	68	8
2021	LM**	35	72	-	456	-	-	-
2022	CAC_S1	35	36	16	220	101	116	3
2022	CAC_S2	31	31	16	207	54	149	4
*Offspring with genotypes incompatible with their assigned maternal sample were excluded and are not counted here (Methods).								
**BORICE selfing vs. outcrossing analysis was not run for LM samples								
***Note that seven 2019 fruits and two 2022 fruits (as well as all 2021 data) had no flowering date data and were excluded from analyses involving flowering date								
†Posterior probability <0.9 for mating status								

Table S3.2. Reference panel lines used for SNP panel creation.

Line	Hybrid index	Ancestry heterozygosity	Group	SRA
CAC110	0.3086	0.007947	CAC high-coverage	SRX21547080
CAC112	0.3432	0.005230	CAC high-coverage	SRX21547081
CAC134	0.1322	0.008466	CAC high-coverage	SRX21547082
CAC141	0.1195	0.003133	CAC high-coverage	SRX21547083
CAC162	0.1327	0.004123	CAC high-coverage	SRX21547084
CAC262	0.2703	0.001954	CAC high-coverage	SRX21547085
CAC415	0.1116	0.007585	CAC high-coverage	SRX21547086
CACG6	0.1816	0.015306	CAC high-coverage	SRX525044
MAR3	0.0003	0.000691	<i>M. guttatus</i> panel	SRX030542
AHQT1G	0.0001	0.000158	<i>M. guttatus</i> panel	SRX142379
ALK24	0.0003	0.000548	<i>M. guttatus</i> panel	SRX6914884
ATTU	0.0001	0.000278	<i>M. guttatus</i> panel	SRX10011990
BOG10	0.0012	0.001538	<i>M. guttatus</i> panel	SRX030570
CSS4	0.0049	0.001454	<i>M. guttatus</i> panel	SRX6435296
DUN	0.0002	0.000366	<i>M. guttatus</i> panel	SRR072711, SRR072713
GUT5	0.0102	0.005636	<i>M. guttatus</i> panel	SRX10011991
IM62	0.0000	0.000000	<i>M. guttatus</i> panel	SRP010318
IM767	0.0002	0.000413	<i>M. guttatus</i> panel	SRX487581
INV	0.0014	0.001873	<i>M. guttatus</i> panel	SRX6914899
LMC24	0.0113	0.006520	<i>M. guttatus</i> panel	SRX030680
MEX	0.0140	0.026151	<i>M. guttatus</i> panel	SRX6914887
Odell	0.0002	0.000336	<i>M. guttatus</i> panel	SRR10194640
PED5	0.0068	0.011444	<i>M. guttatus</i> panel	SRR071969
REM8G	0.0049	0.009701	<i>M. guttatus</i> panel	SRX030546
SCH	0.0222	0.044261	<i>M. guttatus</i> panel	SRX371892
SHG	0.0248	0.040797	<i>M. guttatus</i> panel	SRX10011994
SLP9	0.0016	0.002627	<i>M. guttatus</i> panel	SRX142377
SOL	0.0006	0.001255	<i>M. guttatus</i> panel	SRX6914890
SWB	0.0011	0.000664	<i>M. guttatus</i> panel	SRX030679
TSG3	0.0005	0.000952	<i>M. guttatus</i> panel	SRX2019854
YJS6	0.0013	0.001824	<i>M. guttatus</i> panel	SRX030545
YVO6	0.0442	0.068937	<i>M. guttatus</i> panel	SRX6914891
CACN9	1.0000	0.000000	<i>M. nasutus</i> panel (CAC)	SRR1259271
Koot	1.0000	0.000000	<i>M. nasutus</i> panel	SRR1259272
NHN26	0.9999	0.000235	<i>M. nasutus</i> panel	SRX525051
SF	0.9954	0.001040	<i>M. nasutus</i> panel	SRX116529

Table S3.3. Distribution of samples from each plot in each PCA cluster.

Stream	Plot	Number of samples in PCA Cluster				
		NAS	LM	CAC-A	CAC-B	CAC-C
CAC_S1	S1_1	20	<i>NA</i>	24	<i>NA</i>	1
CAC_S1	S1_2	1	<i>NA</i>	48	<i>NA</i>	<i>NA</i>
CAC_S1	S1_3	7	<i>NA</i>	1	1	35
CAC_S1	S1_4	<i>NA</i>	<i>NA</i>	<i>NA</i>	53	1
CAC_S1	S1_5A	20	<i>NA</i>	<i>NA</i>	<i>NA</i>	1
CAC_S1	S1_5B	<i>NA</i>	<i>NA</i>	<i>NA</i>	1	16
CAC_S1	S1_6	<i>NA</i>	1	<i>NA</i>	1	33
CAC_S2	S2_1	7	<i>NA</i>	<i>NA</i>	<i>NA</i>	3
CAC_S2	S2_2	3	<i>NA</i>	<i>NA</i>	<i>NA</i>	1
CAC_S2	S2_4A	<i>NA</i>	<i>NA</i>	<i>NA</i>	<i>NA</i>	26
CAC_S2	S2_4B	<i>NA</i>	<i>NA</i>	<i>NA</i>	<i>NA</i>	24
CAC_S2	S2_5	<i>NA</i>	<i>NA</i>	1	<i>NA</i>	30
CAC_S2	S2_6	<i>NA</i>	<i>NA</i>	1	<i>NA</i>	26
Little Maui (LM)	LM1	1	5	<i>NA</i>	<i>NA</i>	<i>NA</i>
Little Maui (LM)	LM2	1	8	<i>NA</i>	<i>NA</i>	<i>NA</i>
Little Maui (LM)	LM3	1	7	<i>NA</i>	<i>NA</i>	<i>NA</i>
Little Maui (LM)	LM4	<i>NA</i>	7	<i>NA</i>	<i>NA</i>	<i>NA</i>
Little Maui (LM)	LM5	<i>NA</i>	7	<i>NA</i>	<i>NA</i>	<i>NA</i>
Little Maui (LM)	LM6	<i>NA</i>	8	<i>NA</i>	<i>NA</i>	<i>NA</i>
Little Maui (LM)	LM7	<i>NA</i>	8	<i>NA</i>	<i>NA</i>	<i>NA</i>
Little Maui (LM)	LM8	<i>NA</i>	8	<i>NA</i>	<i>NA</i>	<i>NA</i>

Table S3.4. Statistical model output summaries.

Model	Model	Formula	n, df	Estimate	LRT χ^2	p-value
Plot median HI x flower date	beta	Median HI ~ Median flower date	25, 3	pseudo.r2=0.6073	24.832	<0.0001***
Individual HI x flower date	beta	Maternal HI ~ Flower date	141, 3	pseudo.r2=0.2745	49.081	<0.0001***
		Maternal HI ~ Flower date * Year	141, 5	pseudo.r2=0.3255	13.398	0.0012**
Selfing rate (including <i>M. nasutus</i>)	GLM (binomial -- logit)	(Selfed v. Outcrossed offspring) ~ Maternal HI	151, 2	pseudo.r2=0.1487	92.395	<0.0001***
		(Selfed v. Outcrossed offspring) ~ Maternal HI*Year	151, 6	pseudo.r2=0.2137	40.322	<0.0001***
Selfing rate (excluding <i>M. nasutus</i>)	GLM (binomial -- logit)	(Selfed v. Outcrossed offspring) ~ Maternal HI	140, 2	pseudo.r2=0.0238	12.184	0.0004***
		(Selfed v. Outcrossed offspring) ~ Maternal HI * Year	140, 6	pseudo.r2=0.1214	49.947	<0.0001***
Selfing rate (excluding HI>0.5)	GLM (binomial -- logit)	(Selfed v. Outcrossed offspring) ~ Maternal HI	134, 2	pseudo.r2=0.0000	0.014	0.9058
		(Selfed v. Outcrossed offspring) ~ Year	134, 3	pseudo.r2=0.0674	28.701	<0.0001***
Offspring HI (selfs only)	LM	(Offspring mean HI - Maternal HI) ~ Maternal HI	105, 3	adj.r2=0.0624	7.783	0.0053**
		(Offspring mean HI - Maternal HI) ~ Maternal HI * Year	105, 7	adj.r2=0.1956	20.249	0.0004***
Offspring HI (outcrosses only)	LM	(Offspring mean HI - Maternal HI) ~ Maternal HI	158, 3	adj.r2=0.5017	111.051	<0.0001***
		(Offspring mean HI - Maternal HI) ~ Maternal HI * Year	158, 7	adj.r2=0.6026	39.849	<0.0001***
Offspring HI (selfs + outcrosses)	LM	(Offspring mean HI - Maternal HI) ~ Maternal HI	263, 3	adj.r2=0.2898	90.991	<0.0001***
		(Offspring mean HI - Maternal HI) ~ Maternal HI * Mate history	263, 5	adj.r2=0.3809	38.163	<0.0001***
		(Offspring mean HI - Maternal HI) ~ Maternal HI * Mate history * Year	263, 13	adj.r2=0.4897	59.058	<0.0001***

ANOVA	Predictor	Df	Sum sq	% variance	F	p
Offspring HI (selfs only)	Maternal HI	1	0.023721	7.144	9.237	0.0030**
	Year	2	0.044001	13.252	8.5671	0.0004***
	Maternal HI*Year	2	0.010074	3.034	1.9615	0.1461
	Residuals	99	0.254235	76.570		
Offspring HI (outcrosses only)	Maternal HI	1	0.37125	50.483	199.419	<0.0001***
	Year	2	0.05797	7.883	15.5705	<0.0001***
	Maternal HI*Year	2	0.0232	3.155	6.2314	0.0025**
	Residuals	152	0.28298	38.480		
Offspring HI (selfs + outcrosses)	Maternal HI	1	0.32138	29.247	150.1584	<0.0001***
	Mate type	1	0.03775	3.435	17.6399	<0.0001***
	Year	2	0.08889	8.089	20.7657	<0.0001***
	Maternal HI * Mate type	1	0.07188	6.541	33.5824	<0.0001***
	Maternal HI * Year	2	0.02356	2.144	5.5037	0.0046**
	Mate type * Year	2	0.0097	0.883	2.2649	0.106
	Maternal HI * Mate type * Year	2	0.00849	0.773	1.984	0.1397
	Residuals	251	0.53721	48.888		

Table S3.5. Summary of selfing estimation from BORICE.

Year	Cohort	Fruits w/ ≥ 2 called offspring			Proportion mixed	Mixed fruits only Mean proportion selfed
		100% outcrossed	Mixed	100% selfed		
2019	<i>M. guttatus</i> (HI<0.15)	12	11	2	0.44	0.3799
2021	<i>M. guttatus</i> (HI<0.15)	20	7	0	0.26	0.2940
2022	<i>M. guttatus</i> (HI<0.15)	10	17	7	0.50	0.2987
2019	Admixed (HI>0.15, HI<0.8)	12	31	1	0.70	0.3015
2021	Admixed (HI>0.15, HI<0.8)	6	5	1	0.42	0.4595
2022	Admixed (HI>0.15, HI<0.8)	8	15	5	0.54	0.4215
2019	<i>M. nasutus</i> (HI>0.95)	0	1	3	0.25	0.8571
2021	<i>M. nasutus</i> (HI>0.95)	-	-	-	-	-
2022	<i>M. nasutus</i> (HI>0.95)	0	0	3	0.00	-
TOTAL		68	87	22	0.49	0.3523
<i>Admixed subgroups</i>						
2019	Admixed main cohort (HI>0.15, HI<0.5)	10	30	1	0.73	0.2865
2021	Admixed main cohort (HI>0.15, HI<0.5)	4	5	0	0.56	0.4595
2022	Admixed main cohort (HI>0.15, HI<0.5)	8	13	4	0.52	0.3673
2019	N-backcrosses (HI>0.5, HI<0.8)	2	1	0	0.33	0.7500
2021	N-backcrosses (HI>0.5, HI<0.8)	2	0	1	0.00	-
2022	N-backcrosses (HI>0.5, HI<0.8)	0	2	1	0.67	0.7740

Table S3.6. Summary of sibship estimation from BORICE.

Year	Cohort	Fruits w/ ≥ 2 outcrossed offspring			Outcrossed sibship pairs	
		Total fruits	Fruits w/ half-siblings	Proportion w/ half-siblings	Total pairs	Proportion half-sibs
2019	<i>M. guttatus</i> (HI<0.15)	18	17	0.9444	357	0.5966
2021	<i>M. guttatus</i> (HI<0.15)	13	13	1.0000	1131	0.9240
2022	<i>M. guttatus</i> (HI<0.15)	26	21	0.8077	412	0.6383
2019	Admixed (HI>0.15, HI<0.8)	32	30	0.9375	533	0.8218
2021	Admixed (HI>0.15, HI<0.8)	5	5	1.0000	300	0.7767
2022	Admixed (HI>0.15, HI<0.8)	21	14	0.6667	387	0.5891
2019	<i>M. nasutus</i> (HI>0.95)	-	-	-	-	-
2021	<i>M. nasutus</i> (HI>0.95)	-	-	-	-	-
2022	<i>M. nasutus</i> (HI>0.95)	-	-	-	-	-
TOTAL		115	100	0.8696	3120	0.7756
<i>Admixed subgroups</i>						
2019	Admixed main cohort (HI>0.15, HI<0.5)	30	28	0.9333	529	0.8204
2021	Admixed main cohort (HI>0.15, HI<0.5)	4	4	1.0000	225	0.7867
2022	Admixed main cohort (HI>0.15, HI<0.5)	20	14	0.7000	384	0.5938
2019	N-backcrosses (HI>0.5, HI<0.8)	2	2	1.0000	4	1.0000
2021	N-backcrosses (HI>0.5, HI<0.8)	1	1	1.0000	75	0.7467
2022	N-backcrosses (HI>0.5, HI<0.8)	1	0	0.0000	3	0.0000

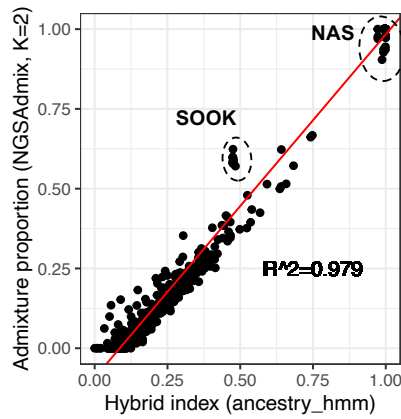


Figure S3.1. Correlation between hybrid index and NGSAdmix structure results.

NGSAdmix value is the proportion assignment to one of two clusters by NGSAdmix with $K=2$; hybrid index is the proportion of sites with *M. nasutus* ancestry from local ancestry inference using AncestryHMM. The circled ‘SOOK’ cluster that deviates from the 1-1 line is composed of *M. sookensis* polyploid individuals (see Methods). ‘NAS’ = *M. nasutus*.

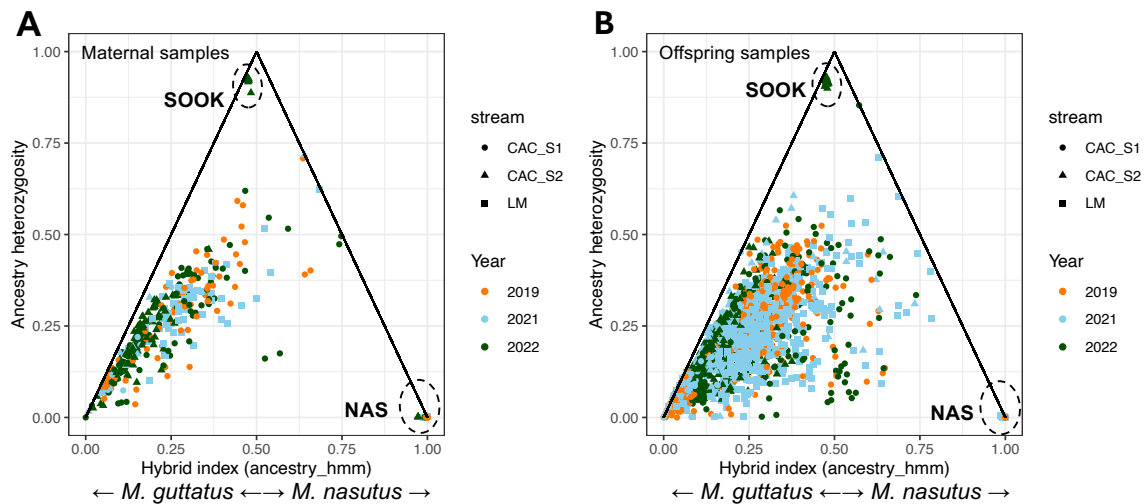


Figure S3.2. Ancestry heterozygosity vs. hybrid index for maternal and offspring plants.

AncestryHMM hybrid index and ancestry heterozygosity outputs for A) maternal and B) offspring individuals. Hybrid index is the proportion of ancestry-informative sites called with *M. nasutus* ancestry from AncestryHMM. Ancestry heterozygosity is the proportion of ancestry-informative sites called has heterozygous (out of the total number of called ancestry-informative sites). First-generation hybrids between 100% *M. guttatus* and 100% *M. nasutus* would have an expected $HI=0.5$ and $AH=1.0$; their offspring would be expected to have lower ancestry heterozygosity (~ 0.5 if selfed). The circled ‘SOOK’ cluster highlights a group of polyploid *M. sookensis* maternal plants; their offspring have similarly high AH values, which indicates fixed heterozygosity (polyploidy) rather than diploid F1 status.

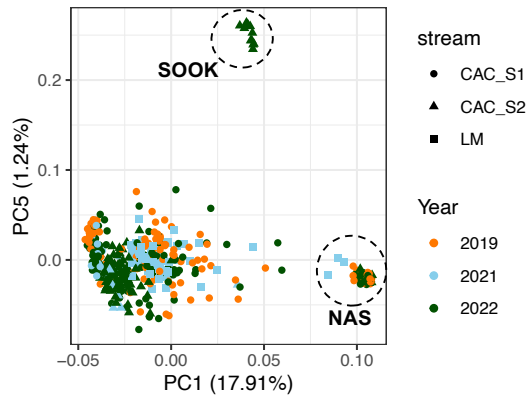


Figure S3.3. PC5 identification of *M. sookensis*. Principal components 1 and 5 from PCAngsd analysis (PCs 1-4 shown in Figure 2A). PC1 correlates strongly with *M. guttatus* vs. *M. nasutus* ancestry (Figure 2B). PC5 clearly delineates a group of individuals identified as the polyploid species *M. sookensis*.

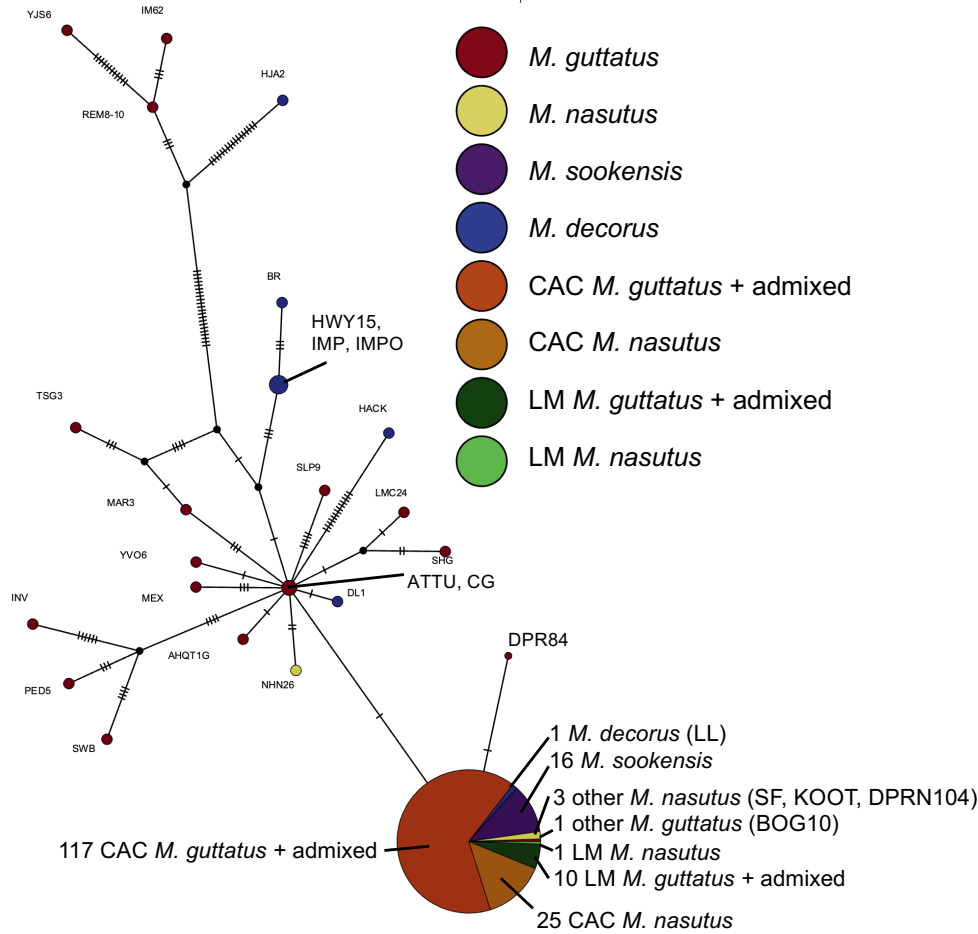


Figure S3.4. Complete mitochondrial capture of a *M. nasutus* haplotype in sympatric *M. guttatus*. NJ-Net haplotype network of mitochondrial variation built from 147 CAC maternal samples, 11 LM maternal samples, and 41 additional samples (98) from the *M. guttatus* species complex, using 120 total variant sites (39 parsimony-informative). The mitochondrial network largely agrees with the chloroplast network (Figure 3), with a single haplotype present in all maternal samples from Catherine Creek and Little Maui, including *M. guttatus*, admixed, and *M. nasutus* samples. All *M. nasutus* samples from across the range share this haplotype or a close derivative, as do samples from *M. sookensis* (a polyploid with *M. nasutus* as maternal parent). *M. guttatus* haplotypes are more variable, with only one non-CAC-area sample sharing the *M. nasutus* haplotype. *M. decorus* is another member of the *M. guttatus* species complex with variable mitochondrial haplotypes.

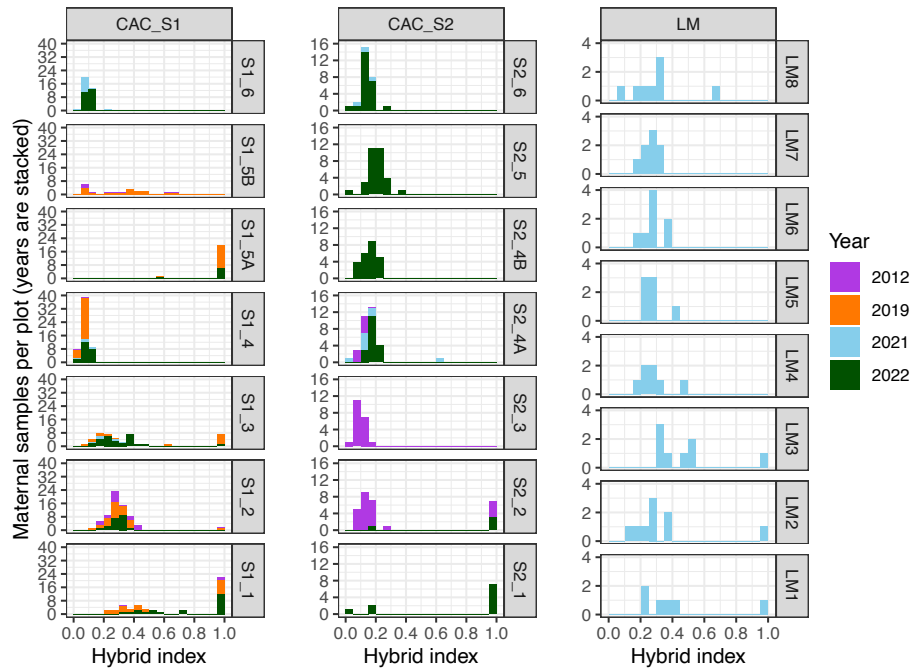


Figure S3.5. Hybrid index at plot resolution for all three streams. A) Histogram of maternal samples sequenced from each plot within each stream, binned by hybrid index, with sampling years in stacked bars. Hybrid index of 0.0 indicates *M. guttatus*, 1.0 indicates *M. nasutus*.

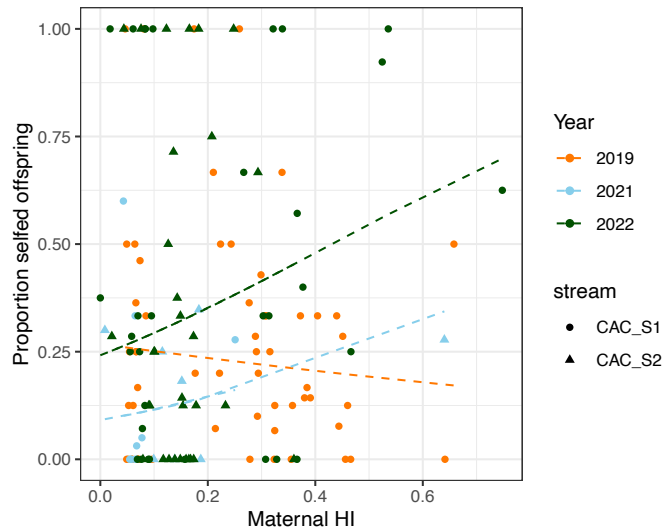


Figure S3.6. Selfing rate vs. hybrid index. Proportion of selfed vs. outcrossed offspring per maternal family is slightly correlated with maternal hybrid index (proportion *M. nasutus* ancestry of the maternal plant). *M. nasutus* families are excluded from this plot and the associated model fit (44 of 45 offspring of *M. nasutus* individuals were inferred to be selfed). Self vs. outcross determined by the BORICE Bayesian model. Offspring without $\geq 90\%$ posterior probability of either state were removed. Dashed lines indicate model fits from a logistic regression with formula $\text{Maternal HI} \sim \text{Proportion selfed} * \text{Year}$ (Table S5).

APPENDIX C

SUPPLEMENTARY TABLES AND FIGURES FROM CHAPTER IV

Table S4.1. Summary of collection locations used in this study.

Location	Latitude	Longitude	Area	Subgroup	Collection type	Date(s) sampled
CAC	45.71133	-121.3637	Northern	CAC	Wild-growing individuals	2019, 2021, 2022
LM	45.70391	-121.3947	Northern	LM	Wild-growing individuals	2021
BFR	37.70078	-120.401	Southern	DPR area allopatric	Wild-collected seeds	4/27/21
COP	37.9759	-120.637	Southern	DPR area allopatric	Wild-collected seeds	6/2/21
GCH	37.75303	-120.248	Southern	DPR area allopatric	Wild-collected seeds	6/3/21
MOC	37.82066	-120.319	Southern	DPR area allopatric	Wild-collected seeds	5/24/21
RCF	37.67907	-120.352	Southern	DPR area allopatric	Wild-collected seeds	6/3/21
RHI	37.82725	-120.467	Southern	DPR area allopatric	Wild-collected seeds	5/7/21, 5/23/21
CST	37.98647	-120.394	Southern	DPR area sympatric	Wild-collected seeds	5/13/21, 6/2/21
GVP	37.84051	-120.227	Southern	DPR area sympatric	Wild-collected seeds	5/25/21, 6/9/21
JAC	37.8987	-120.407	Southern	DPR area sympatric	Wild-collected seeds	5/13/21, 6/3/21
MFA	37.78383	-120.319	Southern	DPR area sympatric	Wild-collected seeds	5/20/21
MFB	37.76285	-120.33	Southern	DPR area sympatric	Wild-collected seeds	6/10/21
MFR	37.74843	-120.333	Southern	DPR area sympatric	Wild-collected seeds	5/14/21, 5/20/21, 6/10/21
NBR	38.05214	-120.471	Southern	DPR area sympatric	Wild-collected seeds	4/28/21, 6/1/2021
NDP	37.98542	-120.415	Southern	DPR area sympatric	Wild-collected seeds	5/12/21, 6/2/21
OPG	37.81361	-120.272	Southern	DPR area sympatric	Wild-collected seeds	6/9/21
RCR	38.01841	-120.749	Southern	DPR area sympatric	Wild-collected seeds	5/19/21
DPI	37.83037	-120.337	Southern	DPR proper	Wild-collected seeds	5/24/21
DPR	37.82925	-120.339	Southern	DPR proper	Wild-growing individuals	April-June 2021
HHR	37.90113	-119.837	Southern	High-elevation	Wild-collected seeds	5/25/21
HHT	37.95786	-119.785	Southern	High-elevation	Wild-collected seeds	5/25/21
TUO	37.81411	-119.862	Southern	High-elevation	Wild-collected seeds	6/9/21

Table S4.2. Summary of sample sizes by location and admixture cohort.

Location	Subgroup	Sequenced	Passed filter	<i>M. guttatus</i>	<i>admixed</i>	<i>M. nasutus</i>
CAC	CAC	510	462	184	213	65
LM	LM	64	61	2	56	3
BFR	DPR area allopatric	10	10	10	0	0
COP	DPR area allopatric	9	8	8	0	0
GCH	DPR area allopatric	10	10	10	0	0
MOC	DPR area allopatric	10	10	10	0	0
RCF	DPR area allopatric	12	12	12	0	0
RHI	DPR area allopatric	11	11	11	0	0
CST	DPR area sympatric	19	18	2	0	16
GVP	DPR area sympatric	8	7	4	1	2
JAC	DPR area sympatric	30	30	18	1	11
MFA	DPR area sympatric	12	12	4	0	8
MFB	DPR area sympatric	8	8	4	0	4
MFR	DPR area sympatric	15	15	14	0	1
NBR	DPR area sympatric	25	25	11	0	14
NDP	DPR area sympatric	10	9	2	0	7
OPG	DPR area sympatric	9	6	5	0	1
RCR	DPR area sympatric	16	16	12	0	4
DPI	DPR proper	7	6	6	0	0
DPR	DPR proper	117	108	100	1	7
HHR	High-elevation	5	5	1	4	0
HHT	High-elevation	19	19	6	8	5
TUO	High-elevation	13	9	2	7	0

Table S4.3. Outlier window statistics for each sample group.

Low-<i>M. nasutus</i> outliers	Total Nonmissing windows	Outlier windows	Proportion of windows
DPR area allopatric <i>M. guttatus</i>	3625	1358	0.3746
DPR area sympatric <i>M. guttatus</i>	3392	170	0.0501
CAC <i>M. guttatus</i>	3589	180	0.0502
CAC <i>admixed</i>	3642	183	0.0502
LM <i>admixed</i>	3585	180	0.0502
DPR area sympatric <i>admixed</i>	2999	150	0.0500
DPR area sympatric <i>M. nasutus</i>	3714	186	0.0501
CAC+LM <i>M. nasutus</i>	3709	199	0.0537
High-<i>M. nasutus</i> outliers	Total Nonmissing windows	Outlier windows	Proportion of windows
DPR area allopatric <i>M. guttatus</i>	3625	182	0.0502
DPR area sympatric <i>M. guttatus</i>	3392	170	0.0501
CAC <i>M. guttatus</i>	3589	180	0.0502
CAC <i>admixed</i>	3642	183	0.0502
LM <i>admixed</i>	3585	180	0.0502
DPR area sympatric <i>admixed</i>	2999	150	0.0500
DPR area sympatric <i>M. nasutus</i>	3714	2423	0.6524
CAC+LM <i>M. nasutus</i>	3709	3258	0.8784
Outliers in shaded rows are excluded from comparisons.			

Table S4.4. Outlier segment statistics for each sample group.

Low-<i>M. nasutus</i> outliers	Number of outlier segments	Total Nonmissing markers	Markers in outlier segments	Proportion of markers	Length of segments (bp)	Prop. of reference (bp)
DPR area allopatric <i>M. guttatus</i>	897	206607	95336	0.4571	172823135	0.5097
DPR area sympatric <i>M. guttatus</i>	184	200455	9504	0.0456	27590455	0.0814
CAC <i>M. guttatus</i>	71	205588	10269	0.0492	12652851	0.0373
CAC <i>admixed</i>	56	206932	10101	0.0484	20940380	0.0618
LM <i>admixed</i>	59	205781	10060	0.0482	34344932	0.1013
DPR area sympatric <i>admixed</i>	109	184585	8639	0.0414	30823782	0.0909
DPR area sympatric <i>M. nasutus</i>	117	208009	9791	0.0469	9089617	0.0268
CAC+LM <i>M. nasutus</i>	51	207913	10344	0.0496	20308953	0.0599
High-<i>M. nasutus</i> outliers	Number of outlier segments	Total Nonmissing markers	Markers in outlier segments	Proportion of markers	Length of segments (bp)	Prop. of reference (bp)
DPR area allopatric <i>M. guttatus</i>	144	206607	9732	0.0467	6300492	0.0186
DPR area sympatric <i>M. guttatus</i>	107	200455	9328	0.0447	6907752	0.0204
CAC <i>M. guttatus</i>	54	205588	10121	0.0485	8365957	0.0247
CAC <i>admixed</i>	53	206932	10204	0.0489	9704619	0.0286
LM <i>admixed</i>	44	205781	10141	0.0486	6905826	0.0204
DPR area sympatric <i>admixed</i>	151	184585	9019	0.0432	10304908	0.0304
DPR area sympatric <i>M. nasutus</i>	603	208009	156613	0.7509	233257597	0.6880
CAC+LM <i>M. nasutus</i>	157	207913	192583	0.9234	288285556	0.8503
Outliers in shaded rows are excluded from comparisons.						

Table S4.5. Comparison of outlier windows and outlier segments for each sample group.

	Total windows containing outlier segments	Excluding missing windows	Overlap: Outlier windows containing outlier segments	Proportion of outlier windows in overlap	Proportion of segment- containing windows in overlap
Low-<i>M. nasutus</i> outliers					
DPR area allopatric <i>M. guttatus</i>	3694	2051	1343	0.9890	0.6548
DPR area sympatric <i>M. guttatus</i>	612	289	163	0.9588	0.5640
CAC <i>M. guttatus</i>	281	201	156	0.8667	0.7761
CAC <i>admixed</i>	442	184	169	0.9235	0.9185
LM <i>admixed</i>	706	293	179	0.9944	0.6109
DPR area sympatric <i>admixed</i>	667	254	144	0.9600	0.5669
DPR area sympatric <i>M. nasutus</i>	237	202	168	0.9032	0.8317
CAC+LM <i>M. nasutus</i>	429	259	195	0.9799	0.7529
High-<i>M. nasutus</i> outliers					
DPR area allopatric <i>M. guttatus</i>	203	198	153	0.8407	0.7727
DPR area sympatric <i>M. guttatus</i>	194	156	137	0.8059	0.8782
CAC <i>M. guttatus</i>	189	163	158	0.8778	0.9693
CAC <i>admixed</i>	217	191	175	0.9563	0.9162
LM <i>admixed</i>	156	146	145	0.8056	0.9932
DPR area sympatric <i>admixed</i>	263	179	121	0.8067	0.6760
DPR area sympatric <i>M. nasutus</i>	4854	2961	2409	0.9942	0.8136
CAC+LM <i>M. nasutus</i>	5813	3369	3250	0.9975	0.9647
Outliers in shaded rows are excluded from comparisons.					

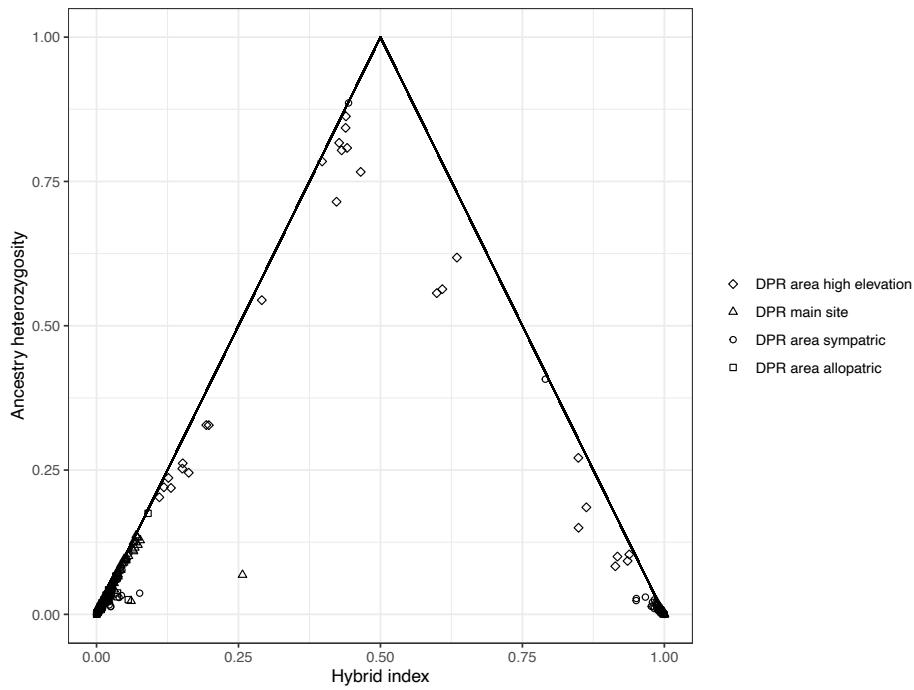


Figure S4.1. Ancestry heterozygosity in southern samples. Ancestry heterozygosity is the proportion of called ancestry-informative markers with a heterozygous call for an individual. Hybrid index is the proportion of *M. nasutus* calls out of called ancestry-informative markers for an individual. Allowing for some uncertainty in heterozygous calling, we classified any sample with ancestry heterozygosity > 0.75 as a putative first-generation hybrid: seven samples from high-elevation sympatric locations, and one sample from another sympatric location.

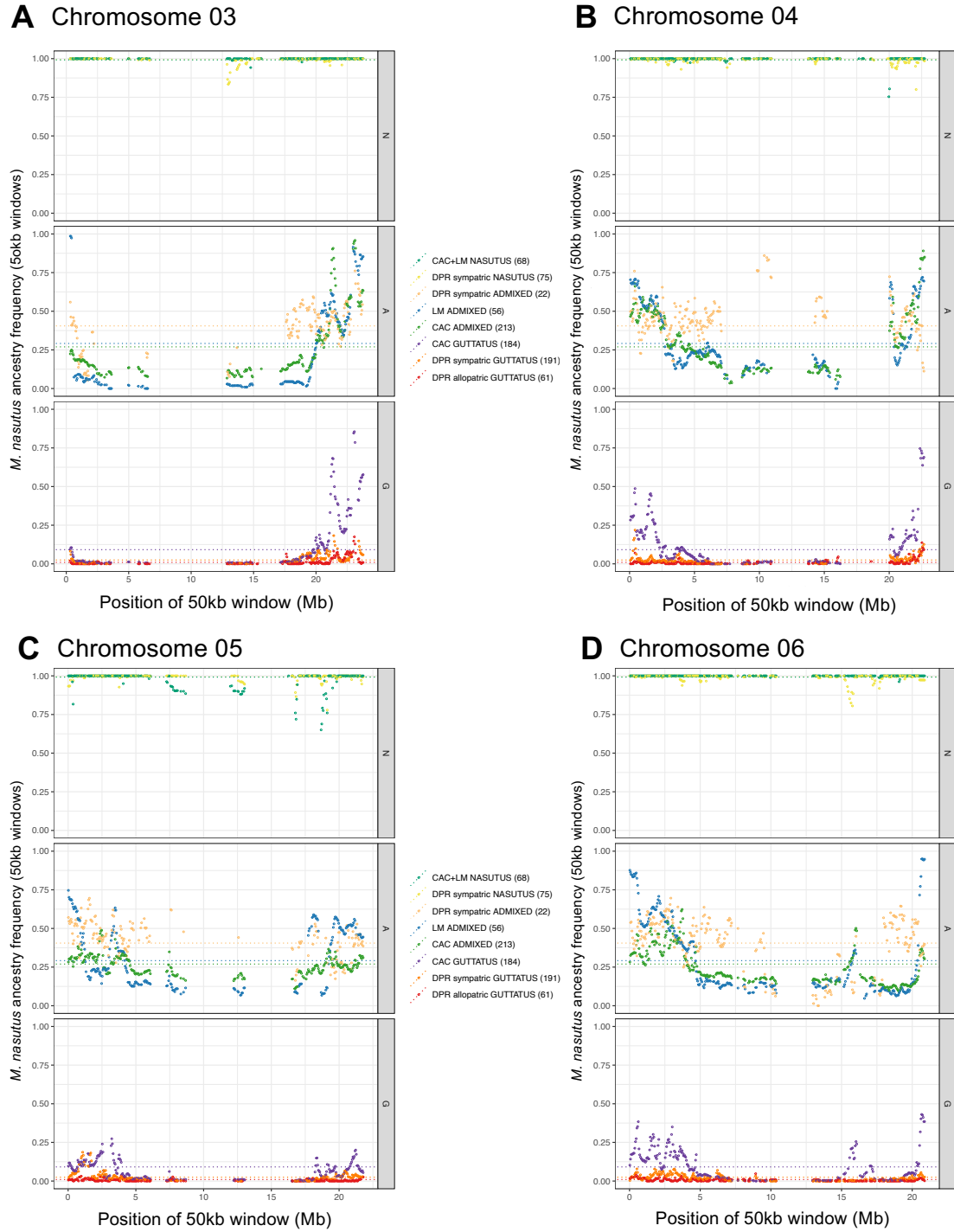


Figure S4.2. Distributions of ancestry frequencies across chromosomes 03, 04, 05, and 06. *M. nasutus* ancestry frequencies in 50kb windows along chromosome 03 (panel A), chromosome 04 (panel B), chromosome 05 (panel C), and chromosome 06 (panel D) for eight sample groups, organized into *M. guttatus* (G), *admixed* (A), and *M. nasutus* (N) cohorts. Horizontal dotted lines indicate the mean ancestry frequency for each sample group.

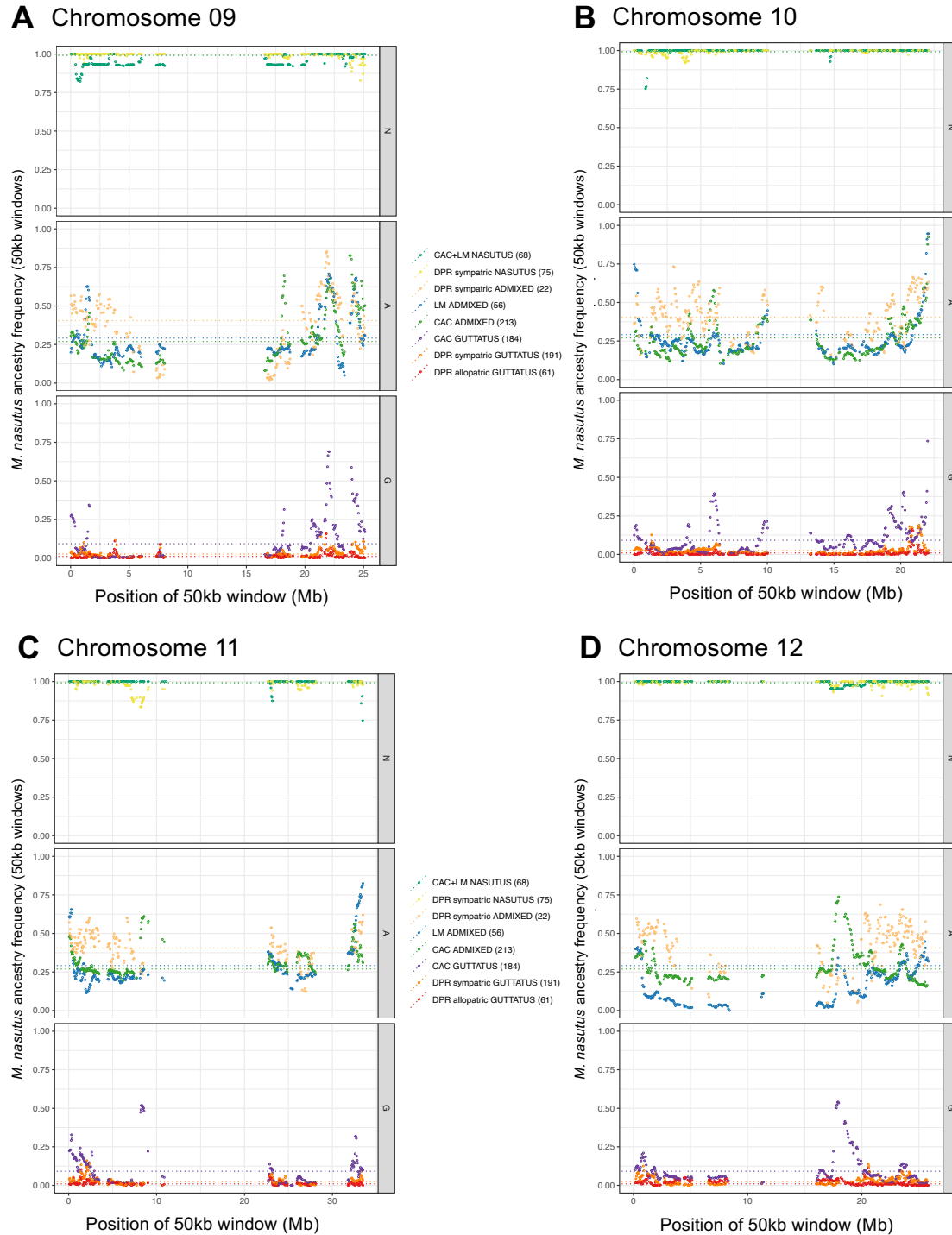


Figure S4.3. Distributions of ancestry frequencies across chromosomes 09, 10, 11, and 12. *M. nasutus* ancestry frequencies in 50kb windows along chromosome 09 (panel A), chromosome 10 (panel B), chromosome 11 (panel C), and chromosome 12 (panel D) for eight sample groups, organized into *M. guttatus* (G), *admixed* (A), and *M. nasutus* (N) cohorts. Horizontal dotted lines indicate the mean ancestry frequency for each sample group.

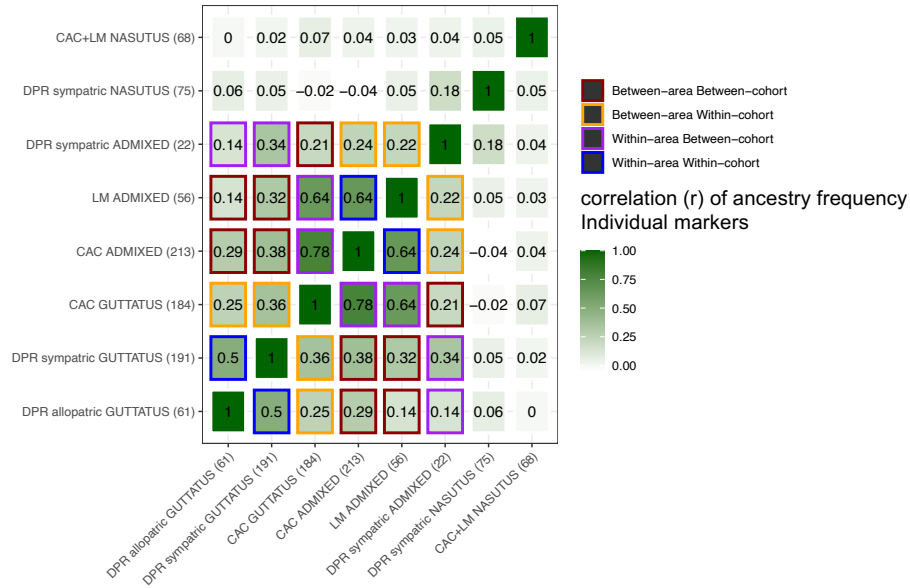
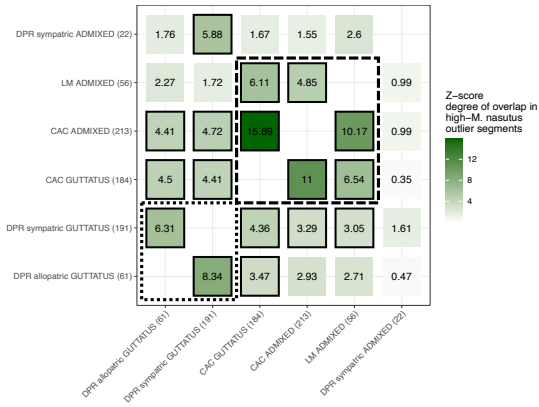


Figure S4.4. Marker-by-marker ancestry correlations. Correlation (r) values between ancestry frequencies of 179,983 ancestry-informative markers across the genome for each pair of sample groups. All included markers have genotype calls in at least 50% of samples within every sample group. Colored outlines indicate whether a comparison is being made within or between areas (northern vs. southern) and within or between cohorts (*M. guttatus* vs. *admixed*); comparisons involving *M. nasutus* groups are not outlined. Numbers in parentheses indicate sample size of each group.

A High-*M. nasutus* outlier segments



B Low-*M. nasutus* outlier segments

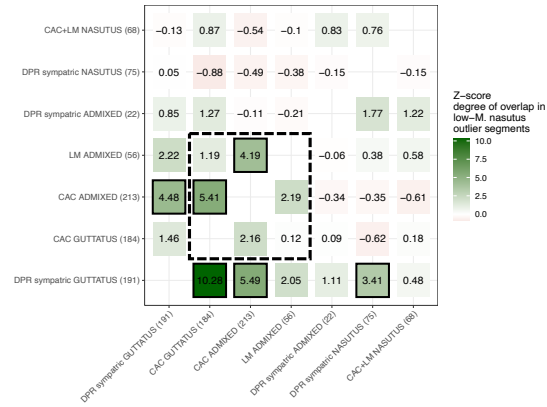


Figure S4.5. Overlap in ancestry outlier segments across sample groups. A) Overlap in high-*M. nasutus* outlier segments. Outlier are defined as segments of at least 10 ancestry markers with ancestry frequency higher than the genome-wide 95% quantile value for a given sample group, with overlap scored as the total shared length in bp of two sets of segments. Each value is a z-score representing to what degree two sets of outlier segments overlap more often than expected by chance. Z-scores are calculated as the deviation of the true overlap length from the mean of the overlap between one sample group (x-axis) and 100 random permutations of the second group (y-axis), scaled by the standard deviation of the permuted values. Comparisons with a p-value < 0.05, after Bonferroni correction for 30 tests, are outlined in solid black. B) Overlap in low-*M. nasutus* outliers, defined as segments of at least 10 ancestry markers with ancestry frequency lower than the genome-wide 5% quantile value for a given sample group. Each value is a z-score as in (A). Comparisons with a p-value < 0.05, after Bonferroni correction for 42 tests, are outlined in black. As a reference point, the dashed box indicates northern – northern group comparisons; the dotted box indicates *M. guttatus* – *M. guttatus* comparisons.