

EVOLUTION AND GENETICS OF HYBRID LETHALITY IN *MIMULUS*

by

MATTHEW P. ZUELLIG

(Under the Direction of Andrea Sweigart)

ABSTRACT

Understanding the evolutionary forces and genetic mechanisms that give rise to new species is a major goal among biologists. The evolution of reproductive barriers is a key component of speciation, as these barriers prevent interspecific gene flow, allowing for species to diverge along distinct evolutionary trajectories. While numerous studies have inferred the importance of particular reproductive barriers in the evolution of species, no study has identified the genetic basis and evolutionary dynamics of reproductive barriers in species that co-occur and hybridize in nature. My dissertation is the first to resolve this problem, by studying the evolutionary genetics of hybrid lethality that occurs between two closely related species of yellow monkeyflower, *Mimulus guttatus* and *M. nasutus*, which co-occur throughout western North America and exhibit patterns of ongoing gene flow. In the first chapter I take a genetic mapping approach to describe the genetic basis of hybrid lethality in *Mimulus*, demonstrating that hybrid lethality is caused by two recessive nuclear loci (*h13* and *h14*) consistent with the Dobzhansky-Muller incompatibility model. In chapter two, I perform a large crossing experiment, to

determine the distribution, frequency, and origin of hybrid lethality alleles in *M. guttatus* and *M. nasutus* populations throughout their common range. I demonstrate that *hl13* and *hl14* are widespread and polymorphic throughout the *Mimulus* range, where polymorphism is likely caused by interspecific gene flow and incompatibility alleles may be maintained in migration-selection balance. Finally, chapter three describes a series of transcriptomic experiments that were used to identify the genes that cause hybrid lethality. Our data demonstrate that the gene *plastid transcriptionally active chromosome 14 (pTAC14)* was duplicated in *M. guttatus* followed by nonfunctionalization of the ancestral copy, whereas no duplication occurred in *M. nasutus*. Hybrid lethality occurs when the chromosomal region associated with the ancestral copy of *pTAC14* is homozygous for *M. guttatus* alleles and the region where the functional *pTAC14* copy transposed, but is lacking in *M. nasutus*, is homozygous for *M. nasutus* alleles. This work addresses long-standing questions regarding the origin of reproductive barriers and the mechanisms that allow them to persist in systems where gene flow is ongoing.

INDEX WORDS: Speciation, Dobzhansky-Muller Incompatibility, *Mimulus*, Hybridization

EVOLUTION AND GENETICS OF HYBRID LETHALITY IN *MIMULUS*

By

MATTHEW P. ZUELLIG

BS, Indiana Wesleyan University, 2008

MS, Grand Valley State University, 2011

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

2017

© 2017

Matthew P. Zuellig

All Rights Reserved

EVOLUTION AND GENETICS OF HYBRID LETHALITY IN *MIMULUS*

By

MATTHEW P. ZUELLIG

Major Professor:	Andrea Sweigart
Committee:	David Hall
	Michael Arnold
	John Burke
	Chung-Jui Tsai
	Kelly Dyer

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2017

ACKNOWLEDGEMENTS

Over the past few years I have been very fortunate to be surrounded by many kind, caring, and supportive people. These people have helped me grow as a scientist and have served as an intellectual and emotional support network throughout grad school. First, I want to thank Andrea Sweigart, who supported me throughout my dissertation and has been a truly excellent advisor. Her enthusiasm for science and general positivity has generated a comfortable and exciting lab environment. Beyond being a great scientist and academic advisor, Andrea has also been a great friend and our discussions have changed my perspective on important issues well beyond the realm of science.

A number of people have helped me develop and execute my project. My committee: Dave, Mike, CJ, John, and Kelly, helped me develop my projects, offered lots of feedback, and always stayed positive about my findings! A number of fellow lab members helped me out intellectually and logistically: Amanda Kenney, Rachel Kerwin, Taylor Harrell, Austin Garner, Ananya Moorthy, Sam Mantel, Fernando Fernandez, and Rachel Hughes. Finally, I want to thank my family and friends, who supported me throughout this process and have helped make graduate school one of the best experiences of my life!

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.....	iv
LIST OF TABLES.....	ix
LIST OF FIGURES.....	x
CHAPTER	
1 INTRODUCTION TO DISSERTATION.....	1
References.....	7
2 THE GENETIC BASIS OF HYBRID LETHALITY IN <i>MIMULUS</i>	10
Abstract.....	11
Introduction.....	11
Methods.....	15
Results.....	17
Discussion.....	24
References.....	29
Tables.....	33
Figures.....	36

3	DYNAMICS OF HYBRID INCOMPATIBILITY IN NATURE.....	42
	Abstract.....	43
	Introduction.....	43
	Methods.....	48
	Results.....	51
	Discussion.....	54
	References.....	62
	Tables.....	66
	Figures.....	68
4	A CRITICAL PHOTOSYNTHESIS GENE UNDERLIES HYBRID LETHALITY IN <i>MIMULUS</i>	74
	Abstract.....	75
	Introduction.....	75
	Methods.....	77
	Results.....	79
	Discussion.....	82
	References.....	86
	Tables.....	89
	Figures.....	90
5	CONCLUSION.....	95

LIST OF TABLES

	Page
Table 1: Table 1.1.....	33
Table 2: Table 1.2.....	34
Table 3: Table 2.1.....	66
Table 4: Table 3.1.....	89

LIST OF FIGURES

	Page
Figure 1: Figure 1.1.....	36
Figure 2: Figure 1.2.....	37
Figure 3: Figure 1.3.....	38
Figure 4: Figure 1.4.....	39
Figure 5: Figure 1.5.....	40
Figure 6: Figure 1.6.....	41
Figure 7: Figure 2.1.....	68
Figure 8: Figure 2.2.....	70
Figure 9: Figure 2.3.....	71
Figure 10: Figure 2.4.....	72
Figure 11: Figure 2.5.....	73
Figure 12: Figure 3.1.....	90
Figure 13: Figure 3.2.....	91
Figure 14: Figure 3.3.....	92
Figure 15: Figure 3.4.....	93
Figure 16: Figure 3.5.....	94

CHAPTER 1

INTRODUCTION TO THESIS

Speciation occurs when previously interfertile populations become genetically isolated through the evolution of reproductive barriers (biological species concept; Dobzhansky 1937, Mayr 1942). Identifying the traits and genes underlying these barriers, as well as the mechanisms driving their evolution, can shed light on the speciation process. Reproductive barriers are generally broken down into two categories: prezygotic barriers, which prevent mating between populations, and postzygotic barriers, which reduce the fitness of hybrid offspring (Coyne and Orr, 2004). Prezygotic barriers, such as differences in flowering time or mate choice, often appear to be a simple byproduct of adaptation to variable environmental conditions in diverging populations (Schluter, 2009). However, the evolution of postzygotic barriers is less straightforward, because selection cannot favor the evolution of maladaptive phenotypes (Orr and Presgraves 2004). To overcome this problem, Dobzhansky (1937) and Muller (1942) proposed that alleles causing reduced hybrid fitness may function properly in the genetic background of pure species, but fail to function properly in a hybrid background. Their key insight was that epistatically interacting genes often coevolve within lineages, but hybridization combines genes that have not coevolved together into a common genetic background where they fail to interact properly. Thus,

selection favoring the production of unfit hybrids need not be invoked and postzygotic barriers can be explained by diverging populations accumulating genetic differences that interact poorly in the hybrid background upon secondary contact.

The hybrid incompatibility (HI) model proposed by Dobzhansky (1937) and Muller (1942) provides a genetic framework for understanding how postzygotic isolation evolves, but makes no predictions about the evolutionary forces acting on the genes involved. Several HI genes have now been identified in diverse taxa, providing clues about the evolutionary forces that act on HI alleles in nature (Presgraves 2010, Maheshwari and Barbash 2011, Sweigart and Willis 2012). In *Drosophila*, many of the HI genes that have been identified harbor signatures of positive selection, where rapid evolutionary change appears to be driven largely by internal genetic conflicts imposed by rapidly evolving heterochromatic regions, selfish transposable elements, and meiotic drive elements, rather than ecological adaptations (Sawamura and Yamamoto 1997, Brideau et al. 2006, Phadnis and Orr 2009, Ferree and Barbash 2009, Tang and Presgraves 2009). Differential suppression of selfish elements on sex-linked chromosomes is also thought to be an important mechanism for driving the evolution of hybrid sterility in mice (Mihola et al. 2009, Larson et al. 2016) and differential resolution of parental conflicts over maternal investment in progeny may be a common cause of HI in plants and animals (Haig and Westoby 1989, Vrana et al. 2000, Josefsson et al. 2006, Burkart-Waco et al. 2012, Rebernig et al. 2015, Garner et al. 2016). Alternatively, rapid evolution of immunity genes driven by local pathogen regimes appears to be a common ecologically driven route to the evolution of HIs that cause autoimmune defects

in plant hybrids (Bomblies et al. 2007, Bomblies 2010, Chae et al. 2014). These studies indicate that internal genetic conflicts and strong ecological selection, involving rapid evolutionary change, may play a disproportionate role in driving the evolution of HIs between species/populations.

Gene duplication and transposition have also been proposed as major factors in the evolution of hybrid incompatibilities, where the reciprocal loss of duplicated genes or transposition of a gene, places critical genes in distinct genomic locations causing deleterious phenotypes in a fraction of hybrids (Werth and Windham 1991, Lynch and Force 2000). The first example of this came from a cross between *Drosophila melanogaster* and *D. simulans*, where a gene critical for male fertility, *JYAlpha*, transposed to a new position in the *D. simulans* lineage. A fraction of the hybrids between these species completely lack *JYAlpha*, causing male sterility (Masly et al. 2006). Three similar examples come from plants and all involve the reciprocal loss of duplicated genes, where neo- or non-functionalization of alternative duplicate gene copies occurred in divergent lineages of *Arabidopsis thaliana* (Bikard et al. 2009) and *Oryza* (Mizuta et al. 2010, Yamagata et al. 2010). Interestingly, all of the instances of reciprocal loss in plants stem from gene-specific duplications, rather than sorting of duplicate genes that occurred during ancestral whole genome duplications – implicating a general role for the mutagenic effects of transposable elements in the evolution of species. Furthermore, HIs caused by transposition/gene loss are likely examples of the evolution of HIs by genetic drift, as opposed to the examples mentioned above that likely involve selection.

In theory, genetic drift or selection can drive the evolution of HI alleles, leading to the fixation of multiple incompatibilities between species that leave them completely reproductively isolated from one another upon secondary contact. However, many closely related species exhibit patterns of incomplete reproductive isolation where interspecific hybridization is ongoing (Grant and Grant 1992, Presgraves 2002, Kenney and Sweigart 2016) and hybrid zones are common (Barton and Hewitt 1985, Harrison 1993, Arnold 1997). Theoretical studies have therefore focused on the evolutionary forces and genetic mechanisms required for the maintenance of incompatibility alleles among naturally hybridizing species, where HIs may actually contribute to reproductive isolation in nature. When gene flow is ongoing, negative selection against HI alleles occurs whenever unfit hybrids are formed, which can result in the loss of incompatibility alleles in favor of compatible alternatives that exist as segregating variation within species or introgress from heterospecifics. This negative selection may be counterbalanced by selection favoring incompatibility alleles *within* species (Gavrilets 1997, Bank et al. 2012), strong linkage between incompatibility alleles and selected alleles (Wright et al. 2013), or by locking incompatibility alleles within species-specific inversions (Noor et al. 2001, Navarro and Barton 2003). These studies indicate that in hybridizing species the genetic context and evolutionary forces acting on HI alleles are important variables for determining whether particular HIs play a permanent or transitory role in reproductive isolation. Therefore, the efficacy of incompatibility as a reproductive barrier should be studied in species pairs that naturally hybridize in nature, where the fitness effects of incompatibility alleles *within* and *between* species can be diagnosed.

Here we investigate the genetic basis and evolutionary dynamics of a HI that causes lethality between two closely related species of yellow monkeyflower, *Mimulus guttatus* and *M. nasutus*. These recently diverged species (200-500 kya; Brandvain et al. 2014) co-occur throughout much of their ranges in the western United States, where reproductive isolation between populations occurs through a number of prezygotic (Martin and Willis 2007, Kenney and Sweigart 2016) and postzygotic barriers (Vickery 1978, Sweigart et al. 2006, 2007, Case and Willis 2008). Despite numerous reproductive barriers, interspecific hybridization has been observed in sympatric populations (Martin and Willis 2007, Kenney and Sweigart 2016) and genome-wide estimates of hybridization indicate historical and ongoing gene flow in the system (Brandvain et al. 2014). This system of incomplete isolation with recurrent hybridization is ideal for investigating the evolutionary forces that drive the evolution of HIs within population and affect their eventual maintenance or degradation.

Our investigation provides one of the first investigations of a HI between naturally hybridizing species, providing important details about the genetic and evolutionary mechanisms that drive the evolution of HIs while also shedding light on the dynamics of HIs in natural populations. In chapters 2 and 4, we use a series of genetic mapping and transcriptomic experiments to demonstrate that hybrid lethality is caused by two unlinked recessive nuclear alleles. Furthermore, these experiments provide strong evidence that duplication of a critical photosynthetic gene, pTAC14, followed by loss of the ancestral copy in *M. guttatus* causes lethality in this system. In chapter 3, we demonstrate the hybrid lethality alleles are widespread and polymorphic in both species.

We also show that patterns of genetic variation associated with hybrid incompatibility alleles are consistent with selection acting against introgression of incompatibility alleles across species barriers and provide evidence that hybrid lethality may be maintained at equilibrium within these species. Taken together, these data provide the first detailed investigation of a HI in a naturally hybridizing populations.

References

- Arnold, M. L. 1997. Natural hybridization and evolution. Oxford University Press.
- Bank, C., R. Bürger, and J. Hermisson. 2012. The limits to parapatric speciation: Dobzhansky–Muller incompatibilities in a continent–island model. *Genetics* 191:845-863.
- Barton, N. H. and G. M. Hewitt. 1985. Analysis of hybrid zones. *Annual review of Ecology and Systematics* 16:113-148.
- Bikard, D., D. Patel, C. Le Metté, V. Giorgi, C. Camilleri, M. J. Bennett, and O. Loudet. 2009. Divergent Evolution of Duplicate Genes Leads to Genetic Incompatibilities Within *A. thaliana*. *Science* 323:623-626.
- Bomblies, K. 2010. Doomed lovers: mechanisms of isolation and incompatibility in plants. *Annual review of plant biology* 61:109-124.
- Bomblies, K., J. Lempe, P. Eppe, N. Warthmann, C. Lanz, J. L. Dangl, and D. Weigel. 2007. Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biol* 5:e236.
- Brandvain, Y., A. M. Kenney, L. Flagel, G. Coop, and A. L. Sweigart. 2014. Speciation and introgression between *Mimulus nasutus* and *Mimulus guttatus*. *PLoS Genet* 10:e1004410.
- Brideau, N. J., H. A. Flores, J. Wang, S. Maheshwari, X. Wang, and D. A. Barbash. 2006. Two Dobzhansky-Muller Genes Interact to Cause Hybrid Lethality in *Drosophila*. *Science* 314:1292-1295.
- Burkart-Waco, D., C. Josefsson, B. Dilkes, N. Kozloff, O. Torjek, R. Meyer, T. Altmann, and L. Comai. 2012. Hybrid incompatibility in *Arabidopsis* is determined by a multiple-locus genetic network. *Plant physiology* 158:801-812.
- Case, A. L. and J. H. Willis. 2008. Hybrid male sterility in *Mimulus* (Phrymaceae) is associated with a geographically restricted mitochondrial rearrangement. *Evolution* 62:1026-1039.
- Chae, E., K. Bomblies, S.-T. Kim, D. Karelina, M. Zaidem, S. Ossowski, C. Martín-Pizarro, R. A. Laitinen, B. A. Rowan, and H. Tenenboim. 2014. Species-wide genetic incompatibility analysis identifies immune genes as hot spots of deleterious epistasis. *Cell* 159:1341-1351.
- Coyne, J. A. and H. A. Orr. 2004. Speciation. Sunderland, MA. Sinauer Associates, Inc.
- Dobzhansky, T. and T. G. Dobzhansky. 1937. Genetics and the Origin of Species. Columbia University Press.
- Ferree, P. M. and D. A. Barbash. 2009. Species-specific heterochromatin prevents mitotic chromosome segregation to cause hybrid lethality in *Drosophila*. *PLoS Biol* 7:e1000234.
- Garner, A. G., A. M. Kenney, L. Fishman, and A. L. Sweigart. 2016. Genetic loci with parent-of-origin effects cause hybrid seed lethality in crosses between *Mimulus* species. *New Phytologist*.
- Gavrilets, S. 1997. Hybrid Zones With Dobzhansky-Type Epistatic Selection. *Evolution* 51:1027-1035.
- Grant, P. R. and R. Grant. 1992. Hybridization of bird species. *Science* 256:193.
- Haig, D. and M. Westoby. 1989. Parent-specific gene expression and the triploid endosperm. *The American Naturalist* 134:147-155.
- Harrison, R. G. 1993. Hybrid zones and the evolutionary process. Oxford University Press on

Demand.

- Josefsson, C., B. Dilkes, and L. Comai. 2006. Parent-dependent loss of gene silencing during interspecies hybridization. *Current Biology* 16:1322-1328.
- Kenney, A. M. and A. L. Sweigart. 2016. Reproductive isolation and introgression between sympatric *Mimulus* species. *Molecular ecology* 25:2499-2517.
- Larson, E. L., S. Keeble, D. Vanderpool, M. D. Dean, and J. M. Good. 2016. The composite regulatory basis of the large X-effect in mouse speciation. *Molecular Biology and Evolution*:msw243.
- Lynch, M. and A. G. Force. 2000. The Origin of Interspecific Genomic Incompatibility via Gene Duplication. *The American Naturalist* 156:590-605.
- Maheshwari, S. and D. A. Barbash. 2011. The Genetics of Hybrid Incompatibilities. *Annual Review of Genetics* 45:331-355.
- Martin, N. H. and J. H. Willis. 2007. Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric *Mimulus* species. *Evolution* 61:68-82.
- Masly, J. P., C. D. Jones, M. A. Noor, J. Locke, and H. A. Orr. 2006. Gene transposition as a cause of hybrid sterility in *Drosophila*. *Science* 313:1448-1450.
- Mayr, E. 1942. *Systematics and the origin of species, from the viewpoint of a zoologist*. Harvard University Press.
- Mihola, O., Z. Trachtulec, C. Vlcek, J. C. Schimenti, and J. Forejt. 2009. A mouse speciation gene encodes a meiotic histone H3 methyltransferase. *Science* 323:373-375.
- Mizuta, Y., Y. Harushima, and N. Kurata. 2010. Rice pollen hybrid incompatibility caused by reciprocal gene loss of duplicated genes. *Proceedings of the National Academy of Sciences* 107:20417-20422.
- Muller, H. J. 1942. Isolating mechanisms, evolution and temperature. Pp. 71-125. *Biol. Symp.*
- Navarro, A. and N. H. Barton. 2003. Chromosomal Speciation and Molecular Divergence--Accelerated Evolution in Rearranged Chromosomes. *Science* 300:321-324.
- Noor, M. A., K. L. Grams, L. A. Bertucci, and J. Reiland. 2001. Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences* 98:12084-12088.
- Orr, H. A. and D. C. Presgraves. 2000. Speciation by postzygotic isolation: forces, genes and molecules. *BioEssays* 22:1085-1094.
- Phadnis, N. and H. A. Orr. 2009. A single gene causes both male sterility and segregation distortion in *Drosophila* hybrids. *science* 323:376-379.
- Presgraves, D. C. 2010. The molecular evolutionary basis of species formation. *Nature Reviews Genetics* 11:175-180.
- Presgraves, D. C. and M. Noor. 2002. Patterns of postzygotic isolation in *Lepidoptera*. *Evolution* 56:1168-1183.
- Sawamura, K. and M.-T. Yamamoto. 1997. Characterization of a reproductive isolation gene, zygotic hybrid rescue, of *Drosophila melanogaster* by using minichromosomes. *Heredity* 79.
- Schluter, D. 2009. Evidence for ecological speciation and its alternative. *Science* 323:737-741.
- Sweigart, A. L., L. Fishman, and J. H. Willis. 2006. A Simple Genetic Incompatibility Causes Hybrid Male Sterility in *Mimulus*. *Genetics* 172:2465-2479.

- Sweigart, A. L., A. R. Mason, and J. H. Willis. 2007. Natural variation for a hybrid incompatibility between two species of *Mimulus*. *Evolution* 61:141-151.
- Sweigart, A. L. and J. H. Willis. 2012. Molecular evolution and genetics of postzygotic reproductive isolation in plants. *F1000 biology reports* 4:23.
- Tang, S. and D. C. Presgraves. 2009. Evolution of the *Drosophila* nuclear pore complex results in multiple hybrid incompatibilities. *Science* 323:779-782.
- Vickery Jr, R. K. 1978. Case studies in the evolution of species complexes in *Mimulus*. Pp. 405-507. *Evolutionary biology*. Springer.
- Vrana, P. B., J. A. Fossella, P. Matteson, T. del Rio, M. J. O'Neill, and S. M. Tilghman. 2000. Genetic and epigenetic incompatibilities underlie hybrid dysgenesis in *Peromyscus*. *Nature genetics* 25:120-124.
- Werth, C. R. and M. D. Windham. 1991. A Model for Divergent, Allopatric Speciation of Polyploid Pteridophytes Resulting from Silencing of Duplicate-Gene Expression. *The American Naturalist* 137:515-526.
- Wright, K. M., D. Lloyd, D. B. Lowry, M. R. Macnair, and J. H. Willis. 2013. Indirect evolution of hybrid lethality due to linkage with selected locus in *Mimulus guttatus*. *PLoS Biol* 11:e1001497.
- Yamagata, Y., E. Yamamoto, K. Aya, K. T. Win, K. Doi, Sobrizal, T. Ito, H. Kanamori, J. Wu, T. Matsumoto, M. Matsuoka, M. Ashikari, and A. Yoshimura. 2010. Mitochondrial gene in the nuclear genome induces reproductive barrier in rice. *Proceedings of the National Academy of Sciences* 107:1494-1499.

CHAPTER 2

THE GENETIC BASIS OF HYBRID LETHALITY IN *MIMULUS*¹

¹ Zuellig, M.P. and A. L. Sweigart, To be submitted to *Nature*.

Abstract

The evolution of reproductive barriers is an essential process underlying the evolution and maintenance of species. Hybrid incompatibilities represent a common class of reproductive barriers that render hybrids unfit, thereby inhibiting gene flow between species. However, in systems that experience ongoing hybridization, selection may act against incompatibility alleles in favor of introgression of compatible alternatives from heterospecifics, resulting in the degradation of hybrid incompatibilities from the system. The loss of incompatibility alleles is largely contingent on the level of gene flow in the system, the genetic architecture of the incompatibility, and the evolutionary forces acting on incompatibility alleles. Here we explore the genetics of a hybrid incompatibility that causes lethality in a fraction of F2 progeny between *Mimulus guttatus* and *M. nasutus*, two closely related wildflowers that live in sympatry and exchange genes at low rates. We genetically map all loci involved in hybrid lethality and identify strong candidate genes that very likely underlie this incompatibility. Our results suggest that hybrid lethality has a simple genetic architecture and is caused by neutral mutational processes involving gene duplication.

Introduction

Identifying the evolutionary and genetic mechanisms that contribute to species diversification and maintenance remains a challenge in evolutionary biology. The evolution of reproductive barriers is a critical aspect of the speciation process, which results in the reduction of interspecific gene exchange and allows for incipient species

to evolve along distinct evolutionary trajectories. Interspecific hybrids often exhibit intrinsic postzygotic defects, including hybrid inviability and sterility, which presumably act as reproductive barriers in secondary contact zones (Coyne and Orr 2004).

Numerous studies in plants (e.g., Rieseberg 2000, Moyle and Graham 2005, Fishman et al. 2006, Sweigart et al. 2006, Bomblies et al. 2007) and animals (e.g., Price et al. 2002, Shuker et al. 2005, Brideau et al. 2006, Stelkens et al. 2015) attribute these hybrid defects to the presence of interspecific hybrid incompatibilities (HIs), where divergent alleles from each species exhibit negative epistasis when brought together in a hybrid genome (Bateson 1909, Dobzhansky 1937, Muller 1942). Alternatively, hybrid dysfunction can result from divergent evolution at paralogous genes (e.g., Bikard et al., 2009, Mizuta et al., 2010, Yamagata et al. 2010), which implicates an important role for gene duplication and passive evolutionary processes (e.g., drift, relaxed selection) in the evolution of HIs (Werth and Windham 1991, Lynch and Force 2000, Lynch and Conery 2000). Identifying the genes involved in HIs can help differentiate between these mechanisms, providing valuable data regarding the evolution of species.

While the Dobzhansky-Muller model provides valuable insight into the genetic basis of postzygotic isolation, it makes no predictions regarding the evolutionary forces acting on incompatibility alleles. In theory, incompatibilities should be purged from species pairs that experience ongoing gene flow, due to negative selection acting against incompatibility alleles when unfit hybrids are formed (Gavrilets 1997, Bank et al. 2012). This negative selection can be counterbalanced by selection favoring incompatibility alleles *within* species (Gavrilets 1997, Bank et al. 2012), strong linkage of

incompatibility alleles to selected alleles, or by locking incompatibility alleles within species-specific inversions (Noor et al. 2001, Navarro and Barton 2003). Indeed, several incompatibility genes that have been identified harbor signatures of positive selection (Ting et al. 1998, Presgraves et al. 2003, Brideau et al. 2006, Oliver et al. 2009, Tang and Presgraves 2009), though these signatures can be caused by a wide spectrum of evolutionary forces (e.g., ecological selection, sexual selection, genomic conflict, mutation pressure), not all of which are expected to increase fitness within species. Additionally, several incompatibility genes appear to have evolved by genetic drift (e.g., Mizuta et al. 2010, Bikard et al. 2009, Yamagata et al. 2010), indicating that they might be rapidly purged upon secondary contact. Thus, the genetic dissection of incompatibilities alone is not always sufficient for determining whether they will serve as effective barriers to gene flow in natural populations. Rather, the efficacy of incompatibility as a reproductive barrier should be studied in species pairs that naturally hybridize in nature, where the fitness effects of incompatibility alleles *within* and *between* species can be diagnosed.

Here we take the first step toward understanding the evolutionary genetic basis of a HI causing seedling lethality between two closely related species of yellow monkeyflower, *Mimulus guttatus* and *M. nasutus*. These recently diverged species (200-500kya; Brandvain et al. 2014) co-occur throughout much of their ranges in the western United States, where reproductive isolation between sympatric populations occurs through a number of prezygotic [habitat isolation (Kiang and Hamrick 1978), reproductive asynchrony (Martin and Willis 2007, Kenney and Sweigart 2016), mating

system isolation (Martin and Willis 2007), pollen tube competition (Diaz and MacNair 1999, Martin and Willis 2007), pollen production (Martin and Willis 2007)] and postzygotic isolating barriers [pollen sterility (Vickery 1978, Sweigart et al. 2006, 2007, Case and Willis 2008, Martin and Willis 2010), anther sterility (Fishman and Willis 2006, Barr and Fishman 2009,), seed inviability (Vickery 1978, Martin and Willis 2010), seedling inviability (Kiang and Libby 1972, Vickery 1973, Willis 1992, MacNair et al., 1993]. HIs in the *Mimulus* system exhibit a large amount of polymorphism (Vickery 1973, Case and Willis 2006, Martin and Willis 2010) with polymorphism in some cases being attributed to variation at particular HI alleles (Christie and MacNair 1987, Sweigart et al. 2007). Despite numerous reproductive barriers, sympatric populations of these species can be found throughout their shared range, where interspecific hybridization rates observed in natural populations ranged from ~0-2% (Martin and Willis 2007) and genome-wide estimates of hybridization indicate historical and ongoing gene flow between the two species (Brandvain et al. 2014, Kenney and Sweigart 2016). Introgression is largely asymmetric, with alleles from *M. nasutus* moving into the *M. guttatus* genome (Sweigart and Willis 2003, Martin and Willis 2007, Brandvain et al. 2014, Kenney and Sweigart 2016). This system of incomplete isolation with recurrent hybridization is ideal for investigating the evolutionary forces that drive the evolution of hybrid incompatibilities within populations and affect their eventual maintenance or degradation.

In this study, we perform high-resolution genetic mapping to identify the loci that cause hybrid lethality between sympatric populations of *M. guttatus* and *M. nasutus*.

Our study provides a complete characterization of the genetic basis and architecture of hybrid lethality in this system, including all interacting DMI loci. Furthermore, we provide evidence that duplication of a single gene that is critical for transcribing the chloroplast genome underlies hybrid lethality in this system. Our study contributes to our understanding of how DMIs evolve within species and sets the stage for future studies to investigate the evolutionary dynamics of DMIs that occur between species that experience gene flow, where the action of reproductive barriers may actually contribute speciation.

Methods

***Mimulus* lines and genetic crosses**

To study the genetics of hybrid lethality, we first generated inbred lines of *M. guttatus* and *M. nasutus* derived from wild individuals collected from the Don Pedro Reservoir (DPR) population in central California, where both species coexist within centimeters of one another (Martin and Willis 2007). Wild-collected seed was sown on moist Fafard 3-B potting soil in 2.5" pots, cold-stratified in the dark at 4C for seven days, and moved the UGA greenhouses to germinate. Upon germination, a single seedling was transplanted to its own 2.5" pot, where upon flowering it was self-fertilized. After three generations of inbreeding with single-seed descent, each line (DPRG102, *M. guttatus* and DPRN104, *M. nasutus*) was intercrossed to generate reciprocal F1 and F2 families (maternal parent always listed first in crosses).

Assessment of hybrid lethality phenotype

Hybrid lethality occurs in the F₂ progeny of reciprocal crosses of DPRG102 and DPRN104 and is easily characterized by seedlings completely lacking chlorophyll ('white seedlings'; see Figure 1.1). We planted the selfed progeny of DPRG102, DPR104, and reciprocal F₁ and F₂ seeds and calculated the ratio of green to white seedlings in each cross-type, to infer the putative genetic basis of hybrid lethality. We planted approximately 600 seeds from each cross and tallied green and white seedlings for the period of a week, at which point they stopped emerging.

Molecular analysis and genetic mapping

DNA was extracted from seedlings and adult leaf tissue using a standard CTAB-chloroform protocol (Doyle and Doyle 1987) modified for use in a 96-well format. Genotyping was performed using a combination of exon-primed intron-spanning size polymorphic markers containing 5' fluorescent tags (6-FAM or HEX) and SNP-containing gene fragments that were analyzed through Sanger sequencing. We designed size-polymorphic and SNP markers from polymorphisms observed in whole genome re-sequence data of multiple lines of *M. guttatus* and *M. nasutus* and confirmed polymorphisms by genotyping parental lines used in our study. A standard touchdown PCR protocol was used in all amplifications and Sanger sequencing reactions were prepared using BigDye v3.1 mastermix (Applied Biosystems, Foster City, USA). Genotyping and Sanger reactions were run on an ABI3730XL automated DNA sequencer at the Georgia Genomics Facility and analyzed using GENEMARKER (Hulce

et al. 2011) and Sequencher (Gene Codes Corporation, Ann Arbor, USA) software, respectively.

To investigate the genetics of hybrid lethality, we performed two rounds of bulked segregant analysis (BSA). In the first round, we genotyped separate pools of green and white DPRG102 x DPRN104 F2 seedlings using size-polymorphic markers that span the *Mimulus* genome, which were ordered based on mapped positions from Fishman et al. (2001) (see Results). In the second round of BSA, we obtained genome-wide SNP genotypes for two pools of DPRG102 x DPRN104 F3 seedlings. Equimolar amounts of green and white seedling DNA were separately pooled and sent to the Duke University sequencing facility, where Illumina libraries with unique barcodes were prepared and subsequently sequenced on the Illumina Hi-seq (100bp single-end reads). Reads from both pools were aligned to the *M. guttatus* (IM62) reference genome (www.phytozome.net), along with previously generated whole genome resequence DPRN104 and MED84. Reads were aligned using Burrows-Wheeler Aligner (bwa, Li and Durbin 2009) with a minimum alignment quality threshold of Q29 (filtering done with samtools, Li et al. 2009). We identified SNPs that differentiated the IM62 lines from the DPRN104 *M. nasutus* parental line using the samtools mpileup function.

Results

Phenotypic segregation of hybrid lethality

As a first step toward investigating the genetic basis of hybrid lethality between inbred lines of *M. guttatus* (DPRG102) and *M. nasutus* (DPRN104) from the sympatric

Don Pedro Reservoir (DPR) site, we characterized the segregation of white and green seedling in parental and reciprocal F1 and F2 hybrids (Table 1.1). White seedlings never segregated among the selfed progeny of parental lines or among the F1 hybrids, but were observed in both F2 hybrid populations (DPRG102xDPRN104, N=516, 7.36% white seedlings; DPRN104xDPRG102, N=661, 5.75% white seedlings). Reciprocity in the segregation of white seedlings suggests a nuclear, rather than cyto-nuclear, genetic incompatibility. Chi-squared tests reject several genetic models that could potentially explain the observed segregation ratios, but could not reject a two-locus model involving two recessive alleles in either F2 population, or when their ratios were combined (Table 1.1). Taken together, these results suggest that *Mimulus* hybrid lethality at DPR is caused by a two-locus recessive-recessive hybrid incompatibility.

Identification of loci involved in hybrid lethality

To genetically map loci underlying hybrid lethality, we performed two rounds of BSA. For the first BSA, we formed eight sets of pooled DNA from F2 seedlings (6 individuals per pool; white seedlings = 4 sets, green seedlings = 4 sets) from the DPRG102 x DPRN104 cross. Because the crossing results suggested that incompatibility alleles act recessively, our expectation was that white seedlings should be homozygous for genetic markers linked to the causal loci. In contrast, genetic markers that are unlinked to a lethality-causing locus should have a 50% chance of being heterozygous and a 50% chance of being homozygous for *M. nasutus* alleles. We genotyped 126 size-polymorphic markers spanning the *Mimulus* genome and identified five that appeared associated with

hybrid lethality (i.e., markers that were homozygous in at least two sets of bulked white seedlings). By genotyping these markers in a DPRN104 x DPR102 F2 mapping population, we determined that all five markers mapped to a single locus near the center of chromosome 14, hereafter referred to as *hybrid lethal 14 (hl14)*. Marker M241 was almost completely fixed for DPRN104 alleles in white F2 seedlings ($p < 0.0005$), but segregated near Mendelian genotypic ratios in green F2s ($p = 0.467$). Additionally, we observed that F3 families produced from green F2s always segregated green:white ratios of 3:1, 15:1, or 1:0 and no white seedlings were ever produced when *hl14* was homozygous for DPRG102 alleles (data not shown), consistent with a model involving a single additional recessive locus that interacts with *hl14*.

To identify the interacting locus, we performed an additional round of BSA, this time controlling for genotype at *hl14*. We generated F3 families from green DPRN104 x DPRG102 F2s that were homozygous for DPRN104 alleles at *hl14*, which resulted exclusively in families that segregate green:white ratios of either 3:1 or 1:0. We reasoned that if hybrid lethality is caused by *hl14* and a single interacting locus, white F3s should be homozygous for DPRG102 alleles at the interactor, whereas green F3s from families that do not segregate white seedlings should be homozygous for DPRN104 alleles. All additional genomic regions are expected to segregate Mendelian SNP frequencies and are not expected to differ in allele frequency. We tested this prediction by examine allele frequency differences throughout the genome in a pool of 34 white F3 seedlings and a pool of 26 F3 green, where each F3 seedling was derived from a different maternal F2 parent. We identified 235,922 high-quality SNPs in our

analysis and calculated the average allele frequency difference between green and white pools in 200-SNP sliding windows across the 14 *Mimulus* chromosomes, where the most divergent 0.5% windows were considered strong candidates for the second lethality locus (Figure 1.2). This approach revealed a single locus at the distal end of chromosome 13, hereafter referred to as hybrid lethal 13 (*hl13*), which contained all of the windows that exhibit extreme allele frequency differences.

To fine-map hybrid lethality loci identified in our BSA, we generated a large F2 mapping population (N=3210) and genotyped each individual at markers linked to *hl13* and *hl14* (*hl13* markers: M208 & M236; *hl14* markers: M241 & M132). We found a significant association between marker genotype and hybrid lethality: nearly all white seedlings are homozygous for DPRG102 alleles at *hl13* and DPRN104 alleles at *hl14* ($p < 0.0005$) and green seedlings are never homozygous for this genotype ($p < 0.0005$) (Figure 1.3). Additionally, we self-fertilized 402 green F2s with known genotypes at *hl13:hl14* to Mendelize each locus and found that this two-locus genotype perfectly predicts the green:white ratios of their F3 progeny (e.g., green F2s that are heterozygous at both loci segregate 15:1 white seedlings in the F3 generation; data not shown).

We initially scored recombinants as mismatches between marker genotypes and phenotypes (e.g., white seedlings heterozygous at both *hl13* markers suggests that *hl13* is outside the bounds of those two markers). In this way, we discovered that the two markers used to screen our mapping population at *hl13* did not flank the causal locus and recombinants were subsequently screened with a more distal flanking marker

(M255); whereas the two markers at *h/14* were shown to flank that locus. Screening recombinants in this fashion identified 5 informative recombinants at *h/13* and 13 informative recombinants at *h/14*. Recombinants were genotyped at additional size-polymorphic and SNP-based markers throughout the interval to identify recombination breakpoints and resolve the mapped intervals (Figure 1.4). *h/13* was mapped to a 72.2kb region at the distal end of chromosome 13 between markers seq43 and M256, which contains 24 genes. *h/14* was mapped to a 51.6kb region in center of chromosome 14, which contains 6 genes and a gap of unknown size in the *M. guttatus* IM62 reference genome. A list of candidate genes identified within each interval, their top blast hit(s) in *Arabidopsis thaliana*, gene ontology terms, known mutant phenotypes, and predicted functions is provided in Table 1.2.

Identification of Strong Candidate Genes

To investigate whether genes within our mapped intervals are misexpressed in white seedlings, we performed quantitative PCR on DPRG102, DPRN104, green F2, and white F2 seedlings. RNA was extracted from pools of 10 seedlings for each genotypic class using a Zymo MicroRNA Kit (Zymo Research, Irvine, USA) followed by cDNA synthesis with GOscript Reverse Transcriptase (Promega, Madison, USA). We designed exon-specific primers to amplify fragments of strong candidate genes (indicated in Table 1.2) and amplified fragments using our standard touchdown PCR protocol. Visualization of gene fragments on a 1% agarose gel revealed that a single gene, Migut.M02023, is expressed in all green seedlings, but never in white seedlings

(Figure 1.5). The *Arabidopsis thaliana* homolog of this gene, pTAC14, is a critical component of the transcriptionally active chromosome (TAC) complex, which plays a critical role in the transcription of the chloroplast genome and causes complete seedling lethality via chlorosis when knocked out (Gao et al. 2011). Interestingly, this gene located in the *hl13* interval has a close paralog, Migut.L00467, on an unmapped scaffold in the *M. guttatus* IM62 reference genome. Visual inspection of several re-sequenced *M. guttatus* and *M. nasutus* lines (from Brandvain et al. 2014) revealed excessive heterozygosity at both Migut.M02023 and Migut.L00467 in *M. guttatus* re-sequence data, indicative of two copies being present in the genome. However, no heterozygosity was present at either paralog in re-sequenced *M. nasutus* lines, indicating the presence of a single copy of this gene. Note that illumina re-sequenced data was aligned to the reference IM62 genome, which allowed reads to map to both paralogs (Brandvain et al. 2014). Analysis of these aligned reads revealed that they mapped equally well to both paralogs in the genome and patterns of SNP variation within the reads strongly supports our conclusion that *M. guttatus* contains two closely related paralogs, while *M. nasutus* contains a single copy of Migut.M02023.

We hypothesized that the *M. guttatus* Migut.L00467 paralog is located in the reference genome gap identified in the *hl14* interval. To test this hypothesis, we first attempted to map Migut.L00467 to the *hl14* region. We designed primers to several conserved exon regions within Migut.M02023/Migut.L00467. We sequenced these gene fragments from genomic DNA from DPRG102, DPRN104, and ten individuals each of the nine two-locus (*hl13;hl14*) F2 genotype classes, which were identified by genotyping

F2 individuals with the flanking markers used in our mapping experiment. Analysis of SNP variation in these data confirmed the presence of two distinct copies of Migut.M02023/Migut.L00467 in DPRG102 (haplotypes G1 and G2), but only a single copy in DPRN104 (haplotype N1, Figure 1.6). Furthermore, we found a perfect association between haplotypes and two-locus genotype: G1 is present in all individuals carrying DPRG102 alleles at *h13*, G2 is present in all individuals carrying DPRG102 alleles at *h14*, and N1 is present in all individuals carrying DPRN104 alleles at *h14*. This perfect association supports our hypothesis that the Migut.L00467 paralog of Migut.M02023 maps to the *h14* interval.

To add additional support for our hypothesis, we obtained bacterial artificial chromosomes (BACs) that were generated from the IM62 *M. guttatus* reference genome that span *h13* and *h14*, respectively. We used our conserved primers to amplify fragments of Migut.M02023/Migut.L00467 in these BACs, successfully amplifying paralogs in both regions (data not shown). These data provide an additional layer of evidence that Migut.M02023/Migut.L00467 paralogs exists at both *h13* and *h14* in *M. guttatus*.

Taken together, these data strongly implicate a role for Migut.M02023/Migut.L00467 in hybrid lethality, presumably due to inactivation or non-functionalization of the G1 haplotype in *M. guttatus*. Indeed, we only observe the G2 haplotype in cDNA generated from DPRG102, indicating that G1 is also inactive in the *M. guttatus* background. Hybrid lethality occurs in individuals that are homozygous for *M. guttatus* alleles at *h13* (i.e., homozygous for G1) and homozygous for *M. nasutus*

alleles at *hl14* (i.e., homozygous for a region lacking a copy of Migut.M02023/Migut.L00467).

Discussion

The biological species concept posits that speciation occurs through the accrual of reproductive isolating barriers that collectively prevent interspecific gene flow (Mayr 1963). Characterization of the genes and evolutionary forces underlying these barriers provides a powerful framework for understanding how species evolve and persist through time. Hybrid incompatibilities represent a broad class of reproductive barriers that presumably play an important role in speciation by reducing hybrid fitness. Despite their ubiquity in nature, we know surprisingly little about the genetic and evolutionary mechanisms that underlie the evolution of HIs. Here, we provide a detailed genetic study of a HI that causes lethality in hybrids of *M. guttatus* and *M. nasutus*. We demonstrate that two recessive nuclear loci, *hl13* and *hl14*, completely account for hybrid lethality in this system. Furthermore, we provide compelling evidence that a gene duplication in *M. guttatus* underlies this incompatibility, supporting a role for neutral mutational processes in the evolution of reproductive barriers.

The Genetic Architecture of Hybrid Incompatibilities

The genetic architecture of HIs has been characterized in several plant and animal systems. Interestingly, there appear to be some differences in architecture between kingdoms, with plants harboring relatively simple HIs (Moyle and Graham

2005, Fishman and Willis 2006, Sweigart et al. 2006, Bikard et al. 2006, Mizuta et al. 2010, Yamagata et al. 2010) and animals harboring more complex incompatibilities (Palopoli and Wu 1994, Presgraves et al. 2003, Tao et al. 2003, Masly and Presgraves 2007, Good et al. 2008). In the *Mimulus* species studied here, additional two-locus HIs have been identified (Sweigart et al. 2006, Fishman et al. 2006) and intraspecific *M. guttatus* HIs are also genetically simple (Willis 1993, Christie and Macnair 1994, Wright et al. 2013). Whether or not the differences in genetic architecture observed in plants and animals are due to biological differences is difficult to resolve and are complicated by a number of factors that may affect genetic architecture including differences in divergence times, life histories, presence of sex chromosomes, and time in sympatry (see Maheshwari and Barbash 2011).

An interesting factor that might affect the genetic architecture of HIs is the presence of ongoing gene flow in sympatry. Orr (1995) points out that complex HIs may be ‘easier’ to evolve, because it is easier for evolution to arrive at complex genotypic combinations without passing through maladaptive intermediates (i.e., the unfit hybrid incompatibility genotype). Thus, complex HIs may commonly occur between allopatric species pairs. However, upon hybridization, complex HIs may also be ‘easier’ to degrade for the same reasons that they were easy to generate *within* species. The presence of multiple ‘fit’ genotypic combinations in recombinant hybrids should greatly facilitate the degradation of incompatible genotypes, as it is easier for evolution to arrive at these fit genotypic combinations without passing through the maladaptive HI genotype. In essence, two-locus HIs may have steeper fitness valleys than genetically

complex HIs and therefore might be more difficult to degrade upon secondary contact than their genetically complex counterparts. Of course, additional factors including the dominance of HI alleles and linkage disequilibrium also play a critical factor in the ease of HI degradation upon secondary contact (Bank et al. 2012). Nonetheless, it seems possible that systems experiencing low levels of gene flow may exhibit a bias toward genetically simple HIs and that some of the differences in genetic architecture observed between plants and animals could be due to bias in the relative contributions of gene flow in the systems that have been studied.

Candidate Genes Underlying Hybrid Lethality

We provide compelling evidence for the role of gene duplication in the evolution of hybrid lethality. Genetic mapping narrowed our list of candidate genes to 25 and 6 in *hl13* and *hl14*, respectively. Investigations of expression of our strongest candidate genes in both of these intervals revealed that our strongest candidate gene in *hl13*, Migut.M02023, is expressed in all green seedling backgrounds, but never in white seedlings. Additionally, we mapped a duplicate copy of this gene, Migut.L00467, located on the unmapped scaffold_193 in the *M. guttatus* IM62 reference genome, to *hl14* and confirmed that duplicate copies of this gene are located at both loci by amplifying fragments of this gene in *M. guttatus* IM62 BACs that span both mapped loci. Furthermore, sequence variation reveals the presence of two distinct haplotypes of this gene in DPRG102, but only a single haplotype in DPRN104, indicating that this duplication is specific to *M. guttatus*.

Migut.M02023 and Migut.L00467 are closely related paralogs, sharing 90.2% sequence similarity in the IM62 reference genome and a common homolog with the *Arabidopsis thaliana* gene *plastid transcriptionally active 14* (*pTAC14*, AT4G20130). *pTAC14* is a highly conserved chloroplast-located protein that is a member of the transcriptionally active chromosome (TAC) and plays an important role in regulating chloroplast development and plastid gene expression (Gao et al. 2011). In *A. thaliana*, *ptac14* knockout plants exhibit disrupted chloroplast development and plastid gene expression, resulting in the production of white seedlings that die shortly after germination (Gao et al. 2011). In our study, we provide compelling evidence that a homolog of this critical gene is not expressed in hybrid lethal individuals, which bare very close resemblance to *A. thaliana ptac14* knockouts. Furthermore, *Mimulus* hybrid lethals harbor only one copy of this critical gene (G1), which also appears to be misexpressed in parental DPRG102 (*M. guttatus*) seedlings. Together, these data indicate that hybrid lethality in *Mimulus* occurs when the apparently nonfunctional G1 is the only copy of *pTAC14* present in the hybrid background.

Gene duplicates are common in plant and animal taxa and whole genome duplication events are associated with increases in biological complexity, evolutionary novelties, and consequent species diversification (Van de Peer et al. 2009). The evolutionary fate of the vast majority of gene duplicates is loss (Lynch and Conery 2002), though contemporary plant genomes retain an average of 65% paralogous genes (Panchy et al. 2016). Given that loss of duplicate genes is so common, several have hypothesized that reciprocal loss of duplicated genes may result in the rapid

evolution of HIs (Werth and Wndham 1991, Lynch and Force 2000). Our data indicate that this process may underlie the evolution of hybrid lethality in *Mimulus*, where loss of the ancestral copy of pTAC14 at *hl13* in *M. guttatus* results in hybrid lethality in a fraction of F2 progeny. Interestingly, the transposition of pTAC14 to *hl14* in *M. guttatus* may be the result of a single gene transposition, as opposed to sorting of an ancestral duplication. This is supported by the lack of amplifiable pTAC14 fragments at *hl14* in *M. nasutus* and the presence of MULE-like elements, which are the most common cause of TE-mediated gene transposition in plants (Bennetzen 2005), flanking the copy of pTAC14 located at *hl14* in *M. guttatus*. The context of gene duplication is important, because transpositional duplications are often associated with changes in promotor sequence and dosage imbalances, which may account for the observation that genes duplicated by different mechanisms are enriched for different gene families (Wang et al. 2012). Interestingly, of the three additional studies that have identified duplicate genes underlying HIs in plants, all implicate a role for recent small scale duplications, rather than sorting of duplicates from ancient whole genome duplications (Bikard et al. 2006, Yamagata et al. 2010, and Mizuta et al. 2010). Whether these findings are coincidental or reflect a biologically relevant mechanism for the generation of HIs is unknown and future studies should investigate this matter further.

References

- Bank, C., R. Bürger, and J. Hermisson. 2012. The limits to parapatric speciation: Dobzhansky–Muller incompatibilities in a continent–island model. *Genetics* 191:845-863.
- Barr, C. M. and L. Fishman. 2010. The nuclear component of a cytonuclear hybrid incompatibility in *Mimulus* maps to a cluster of pentatricopeptide repeat genes. *Genetics* 184:455-465.
- Bateson, W. 1909. Heredity and variation in modern lights. *Darwin and modern science* 85:101.
- Bennetzen, J. L. 2005. Transposable elements, gene creation and genome rearrangement in flowering plants. *Current opinion in genetics & development* 15:621-627.
- Bikard, D., D. Patel, C. Le Metté, V. Giorgi, C. Camilleri, M. J. Bennett, and O. Loudet. 2009. Divergent Evolution of Duplicate Genes Leads to Genetic Incompatibilities Within *A. thaliana*. *Science* 323:623-626.
- Bomblies, K., J. Lempe, P. Eppe, N. Warthmann, C. Lanz, J. L. Dangl, and D. Weigel. 2007. Autoimmune Response as a Mechanism for a Dobzhansky-Muller-Type Incompatibility Syndrome in Plants. *PLOS Biology* 5:e236.
- Brandvain, Y., A. M. Kenney, L. Flagel, G. Coop, and A. L. Sweigart. 2014. Speciation and introgression between *Mimulus nasutus* and *Mimulus guttatus*. *PLoS Genet* 10:e1004410.
- Brideau, N. J., H. A. Flores, J. Wang, S. Maheshwari, X. Wang, and D. A. Barbash. 2006. Two Dobzhansky-Muller Genes Interact to Cause Hybrid Lethality in *Drosophila*. *Science* 314:1292-1295.
- Case, A. L. and J. H. Willis. 2008. Hybrid male sterility in *Mimulus* (Phrymaceae) is associated with a geographically restricted mitochondrial rearrangement. *Evolution* 62:1026-1039.
- Chou, J.-Y., Y.-S. Hung, K.-H. Lin, H.-Y. Lee, and J.-Y. Leu. 2010. Multiple molecular mechanisms cause reproductive isolation between three yeast species. *PLoS Biol* 8:e1000432.
- Christie, P. and M. R. Macnair. 1987. The distribution of postmating reproductive isolating genes in populations of the yellow monkey flower, *Mimulus guttatus*. *Evolution*:571-578.
- Codes, G. 2000. Sequencher: version 4.1. 2. gene codes corporation. Ann Arbor.
- Coyne, J. A. and H. A. Orr. 2004. Speciation. Sunderland, MA. Sinauer Associates, Inc.
- Diaz, A. and M. Macnair. 1999. Pollen tube competition as a mechanism of prezygotic reproductive isolation between *Mimulus nasutus* and its presumed progenitor *M. guttatus*. *New Phytologist* 144:471-478.
- Dobzhansky, T. and T. G. Dobzhansky. 1937. *Genetics and the Origin of Species*. Columbia University Press.
- Doyle, J. and J. Doyle. 1987. Genomic plant DNA preparation from fresh tissue-CTAB method. *Phytochem Bull* 19:11-15.
- Fishman, L., A. J. Kelly, E. Morgan, and J. H. Willis. 2001. A genetic map in the *Mimulus guttatus* species complex reveals transmission ratio distortion due to heterospecific interactions. *Genetics* 159:1701-1716.
- Fishman, L., J. H. Willis, and J. Kohn. 2006. A cytonuclear incompatibility causes anther sterility in *Mimulus* hybrids. *Evolution* 60:1372-1381.

- Gao, Z.-P., Q.-B. Yu, T.-T. Zhao, Q. Ma, G.-X. Chen, and Z.-N. Yang. 2011. A functional component of the transcriptionally active chromosome complex, Arabidopsis pTAC14, interacts with pTAC12/HEMERA and regulates plastid gene expression. *Plant physiology* 157:1733-1745.
- Gavrilets, S. 1997. Hybrid Zones With Dobzhansky-Type Epistatic Selection. *Evolution* 51:1027-1035.
- Good, J. M., M. A. Handel, and M. W. Nachman. 2008. ASYMMETRY AND POLYMORPHISM OF HYBRID MALE STERILITY DURING THE EARLY STAGES OF SPECIATION IN HOUSE MICE. *Evolution; international journal of organic evolution* 62:50-65.
- Hulce, D., X. Li, T. Snyder-Leiby, and C. S. Johathan Liu. 2011. GeneMarker® Genotyping Software: Tools to Increase the Statistical Power of DNA Fragment Analysis. *Journal of Biomolecular Techniques : JBT* 22:S35-S36.
- Kenney, A. M. and A. L. Sweigart. 2016. Reproductive isolation and introgression between sympatric *Mimulus* species. *Molecular ecology* 25:2499-2517.
- Kiang, Y. 1973. Floral structure, hybridization and evolutionary relationship of two species of *Mimulus*. *Rhodora* 75:225-238.
- Kiang, Y. and J. Hamrick. 1978. Reproductive isolation in the *Mimulus guttatus* M. *nasutus* complex. *American Midland Naturalist*:269-276.
- Kiang, Y. and W. Libby. 1972. Maintenance of a lethal in a natural population of *Mimulus guttatus*. *The American Naturalist* 106:351-367.
- Li, H. and R. Durbin. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25:1754-1760.
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and S. Genome Project Data Processing. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078-2079.
- Lynch, M. and J. S. Conery. 2000. The Evolutionary Fate and Consequences of Duplicate Genes. *Science* 290:1151-1155.
- Lynch, M. and A. G. Force. 2000. The Origin of Interspecific Genomic Incompatibility via Gene Duplication. *The American Naturalist* 156:590-605.
- MacNair, M. R., S. E. Smith, and Q. J. Cumbes. 1993. Heritability and distribution of variation in degree of copper tolerance in *Mimulus guttatus* at Copperopolis, California. *HEREDITY-LONDON*- 71:445-445.
- Maheshwari, S. and D. A. Barbash. 2011. The Genetics of Hybrid Incompatibilities. *Annual Review of Genetics* 45:331-355.
- Martin, N. H. and J. H. Willis. 2007. Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric *Mimulus* species. *Evolution* 61:68-82.
- Martin, N. H. and J. H. Willis. 2010. Geographical variation in postzygotic isolation and its genetic basis within and between two *Mimulus* species. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365:2469-2478.
- Masly, J. P. and D. C. Presgraves. 2007. High-resolution genome-wide dissection of the two rules of speciation in *Drosophila*. *PLoS Biol* 5:e243.
- Mayr, E. 1963. Animal species and evolution. *Animal species and evolution*.

- Mizuta, Y., Y. Harushima, and N. Kurata. 2010. Rice pollen hybrid incompatibility caused by reciprocal gene loss of duplicated genes. *Proceedings of the National Academy of Sciences* 107:20417-20422.
- Moyle, L. C. and E. B. Graham. 2005. Genetics of Hybrid Incompatibility Between *Lycopersicon esculentum* and *L. hirsutum*. *Genetics* 169:355-373.
- Muller, H. J. 1942. Isolating mechanisms, evolution and temperature. Pp. 71-125. *Biol. Symp.*
- Navarro, A. and N. H. Barton. 2003. Chromosomal Speciation and Molecular Divergence-- Accelerated Evolution in Rearranged Chromosomes. *Science* 300:321-324.
- Noor, M. A., K. L. Grams, L. A. Bertucci, and J. Reiland. 2001. Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences* 98:12084-12088.
- Oliver, P. L., L. Goodstadt, J. J. Bayes, Z. Birtle, K. C. Roach, N. Phadnis, S. A. Beatson, G. Lunter, H. S. Malik, and C. P. Ponting. 2009. Accelerated evolution of the Prdm9 speciation gene across diverse metazoan taxa. *PLoS Genet* 5:e1000753.
- Orr, H. A. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics* 139:1805-1813.
- Panchy, N., M. Lehti-Shiu, and S.-H. Shiu. 2016. Evolution of Gene Duplication in Plants. *Plant Physiology* 171:2294-2316.
- Presgraves, D. C., L. Balagopalan, S. M. Abmayr, and H. A. Orr. 2003. Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. *Nature* 423:715-719.
- Price, T. D., M. M. Bouvier, and S. Edwards. 2002. THE EVOLUTION OF F1 POSTZYGOTIC INCOMPATIBILITIES IN BIRDS. *Evolution* 56:2083-2089.
- Rieseberg, L. H. 2000. CROSSING RELATIONSHIPS AMONG ANCIENT AND EXPERIMENTAL SUNFLOWER HYBRID LINEAGES. *Evolution* 54:859-865.
- Seidel, H. S., M. V. Rockman, and L. Kruglyak. 2008. Widespread genetic incompatibility in *C. elegans* maintained by balancing selection. *science* 319:589-594.
- Shuker, D. M., K. Underwood, T. M. King, and R. K. Butlin. 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* 272:2491-2497.
- Stelkens, R. B., C. Schmid, and O. Seehausen. 2015. Hybrid Breakdown in Cichlid Fish. *PLOS ONE* 10:e0127207.
- Sweigart, A. and J. Willis. 2003. Patterns of nucleotide diversity are affected by mating system and asymmetric introgression in two species of *Mimulus*. *Evolution* 57:2490-2506.
- Sweigart, A. L., L. Fishman, and J. H. Willis. 2006. A Simple Genetic Incompatibility Causes Hybrid Male Sterility in *Mimulus*. *Genetics* 172:2465-2479.
- Sweigart, A. L., A. R. Mason, and J. H. Willis. 2007. Natural variation for a hybrid incompatibility between two species of *Mimulus*. *Evolution* 61:141-151.
- Tang, S. and D. C. Presgraves. 2009. Evolution of the *Drosophila* nuclear pore complex results in multiple hybrid incompatibilities. *Science* 323:779-782.
- Ting, C.-T., S.-C. Tsaur, M.-L. Wu, and C.-I. Wu. 1998. A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science* 282:1501-1504.
- Van de Peer, Y., S. Maere, and A. Meyer. 2009. The evolutionary significance of ancient

- genome duplications. *Nature Reviews Genetics* 10:725-732.
- Vickery Jr, R. K. 1978. Case studies in the evolution of species complexes in *Mimulus*. Pp. 405-507. *Evolutionary biology*. Springer.
- Wang, Y., X. Wang, and A. H. Paterson. 2012. Genome and gene duplications and gene expression divergence: a view from plants. *Annals of the New York Academy of Sciences* 1256:1-14.
- Werth, C. R. and M. D. Windham. 1991. A Model for Divergent, Allopatric Speciation of Polyploid Pteridophytes Resulting from Silencing of Duplicate-Gene Expression. *The American Naturalist* 137:515-526.
- Willis, J. H. 1992. Genetic analysis of inbreeding depression caused by chlorophyll-deficient lethals in *Mimulus guttatus*. *HEREDITY-LONDON*- 69:562-562.
- Wright, K. M., D. Lloyd, D. B. Lowry, M. R. Macnair, and J. H. Willis. 2013. Indirect evolution of hybrid lethality due to linkage with selected locus in *Mimulus guttatus*. *PLoS Biol* 11:e1001497.
- Yamagata, Y., E. Yamamoto, K. Aya, K. T. Win, K. Doi, Sobrizal, T. Ito, H. Kanamori, J. Wu, T. Matsumoto, M. Matsuoka, M. Ashikari, and A. Yoshimura. 2010. Mitochondrial gene in the nuclear genome induces reproductive barrier in rice. *Proceedings of the National Academy of Sciences* 107:1494-1499.

Table 1.1: Segregation of white seedlings in parental and reciprocal hybrid crosses.

Maternal parents are listed first in crosses. 'G' is DPRG102 (*M. guttatus*) and 'N' is DPRN104 (*M. nasutus*). χ^2 tests with one d.f. used to assess green:white ratios expected under various genetic models: 2 recessive loci (15:1), dominant-recessive (13:3), 3 recessive alleles (63:1), and 2 recessive and 1 dominant (61:3). Non-significant values (bold) indicate models consistent with observed ratios.

Genotype	# Green	# White	2-loci (15:1)	2-loci (13:3)	3-loci (63:1)	3-loci (61:3)
DPRG102	534	0	35.6***	123.23***	8.46**	26.26***
DPRN104	355	0	23.66***	81.92***	5.63*	17.46***
GxN F1	634	0	42.26***	146.31***	10.06**	31.18***
NxG F1	577	0	38.47***	133.15***	9.16**	28.38***
GxN F2	478	38	1.09	43.91***	112.92***	8.28**
NxG F2	623	38	0.28	73.34***	75.32***	1.67

*=P<0.05; **=P<0.01; ***=P<0.001

Table 1.2: Candidate genes at *hl13* and *hl14*

Gene	Annotation	Predicted Function	Chlorosis/lethality phenotype in <i>Arabidopsis</i>
<i>hl13</i>			
Migut.M02005	Allantoate deiminase	Purine turnover	None
Migut.M02006	None	ATP binding, carbohydrate binding, kinase activity	None
Migut.M02007	Purine biosynthesis protein 6/7, PUR6/7	Catalyzes aspartate addition to alpha-amino group to purine backbone	Localized to chloroplast, chloroplast stroma
Migut.M02008	Zinc finger protein 207	Encodes SUF4, a putative transcription factor required for delayed flowering in winter-annual	None
Migut.M02009	Protein unusual floral organs	Required for proper identity of floral meristem.	None
Migut.M02010	GTP-binding protein ERG	RNA binding, GTP binding	None
Migut.M02011	DNA polymerase ETA	Catalyses translesion synthesis in response to UV damage.	None
Migut.M02012	Phosphoglycerate mutase family protein	Catalytic activity	None
Migut.M02013	Proteasome subunit alpha type-4	Encodes the alpha-3 subunit of 20S proteasome	None
Migut.M02014	Haloacid dehalogenase-like hydrolase	Hydrolase, catalytic, and phosphoglycolate phosphatase activity	None
Migut.M02015	Mitochondrial translocator assembly and maintenance protein 41	None	Localized to chloroplast
Migut.M02016 & Migut.M02019	Cationic amino acid transporter 5	Amino acid transporter	None
Migut.M02017 & Migut.M02020 & Migut.M02022	40S ribosomal protein S28	Structural constituent of ribosome, mRNA binding	None
Migut.M02018	Mitochondrial amidoxime-reducing component 1	Molybdenum ion binding, catalytic activity	Localized to chloroplast
Migut.M02021	Myb-like DNA binding domain	Telomeric repeat binding protein	None
Migut.M02023	N-lysine methyltransferase EFM1; pTAC14	Regulatory role in chloroplast development and gene expression	Localized to chloroplast. Chloroplast fails to develop in knockout mutants, leading to premature lethality at the cotyledon stage.
Migut.M02024	Vacuolar-sorting receptor 6	Calcium ion binding, protein targeting to vacuole	None

Migut.M02025	Protein trigalactosyldiacylglycerol 2, chloroplastic	Involved in lipid transfer from ER to chloroplast.	Mutants consistently smaller and slightly pale, with reduced chlorophyll content.
Migut.M02026	PPR repeat family	Involved in chloroplast mRNA editing.	Localized to chloroplast.
Migut.M02027	Myb-like DNA-binding protein	None	None
Migut.M02028	Transducin/WD40 domain-containing protein	None	None
Migut.M02029	LRR receptor-like serine/threonine protein kinase GS01-related	Required for formation of normal epidermal surface during embryogenesis	None
<i>hl14</i>			
Migut.N01485	Clustered mitochondria protein	Regulates mitochondrial association time and involved in mitochondrial fusion	Numerous phenotypes including lower biomass, shorter roots, short etiolated hypocotyls, large number of dead root cells, and overall transcriptome reprogramming.
Migut.N01486	DNAJ homolog subfamily C member	Involved in female gametophyte development	None
Migut.N01487	C2 calcium/lipid binding plant phosphoribosyltransferase family protein	Involved in transport of florigen FT from companion cells to sieve elements	None
Migut.N01488	Sister chromatid cohesion 1 protein 3	Essential for megagametogenesis and plays important role in pollen development, a key component of meiotic and mitotic cohesion complexes	Gametophytic lethality phenotypes in male and female gametophytes – no evidence for seed development defects
Migut.N01489	Thioredoxin superfamily protein	Oxidation-reduction process	None
Migut.N01490	AXI 1 protein-like protein	Transferase activity	None

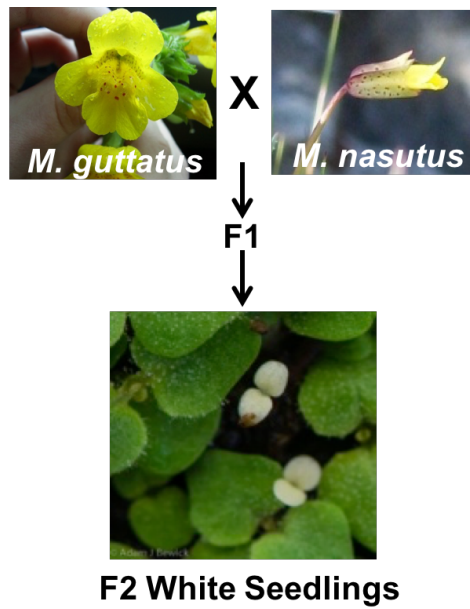


Figure 1.1: *Mimulus guttatus*, *M. nasutus*, and white and green F2 progeny. Photo taken by Adam Bewick.

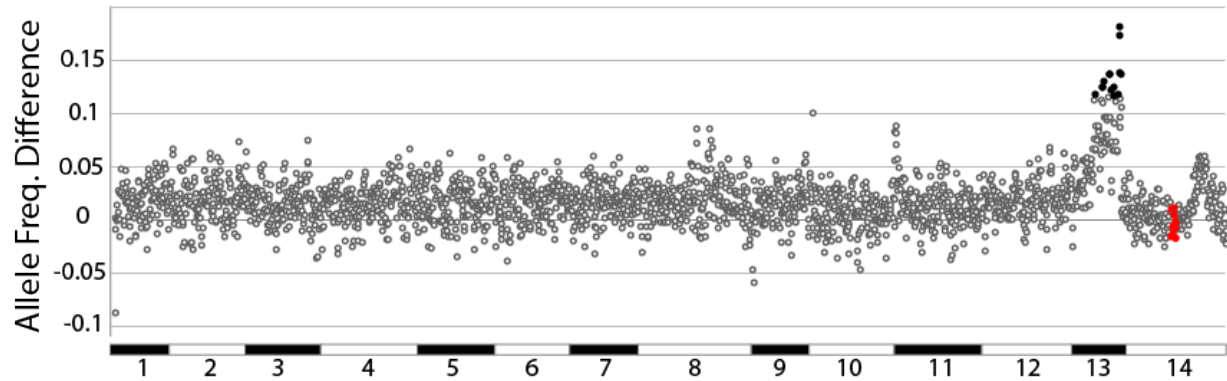


Figure 1.2: Bulk Segregant Analysis. Difference in in allele frequency between green and white F3 bulked pools are plotted along the fourteen *Mimulus* chromosomes.

Average allele frequency was calculated in windows containing 200 SNPs and plotted in sliding windows with 100 SNP overlap between windows. The 0.5% most divergent windows are highlighted as black dots, which are all located at the distal end of chromosome 13 in contiguous windows and represent the candidate *h/13* region. Red dots, which overlap the mapped interval for *h/14*, show the 0.5% least divergent windows (calculated as absolute difference in allele frequency).

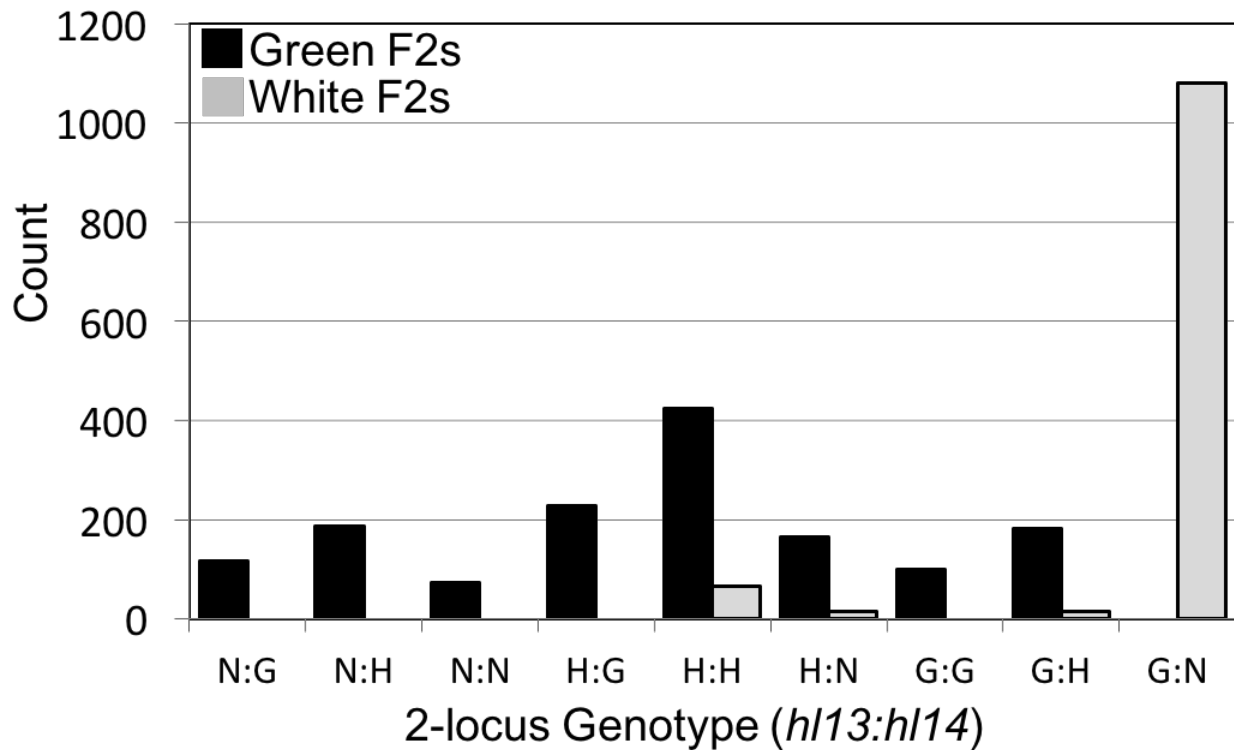


Figure 1.3: Two-locus genotypes of green and white seedlings at markers linked to *hl13* and *hl14*. Count represents the number of F2s with a given genotype. Genotypes are homozygous *M. guttatus* ('G'), homozygous *M. nasutus* ('N'), and heterozygous ('H'). Note that green F2s carry all genotypes except G:N, whereas the majority of white F2s carry this genotype.

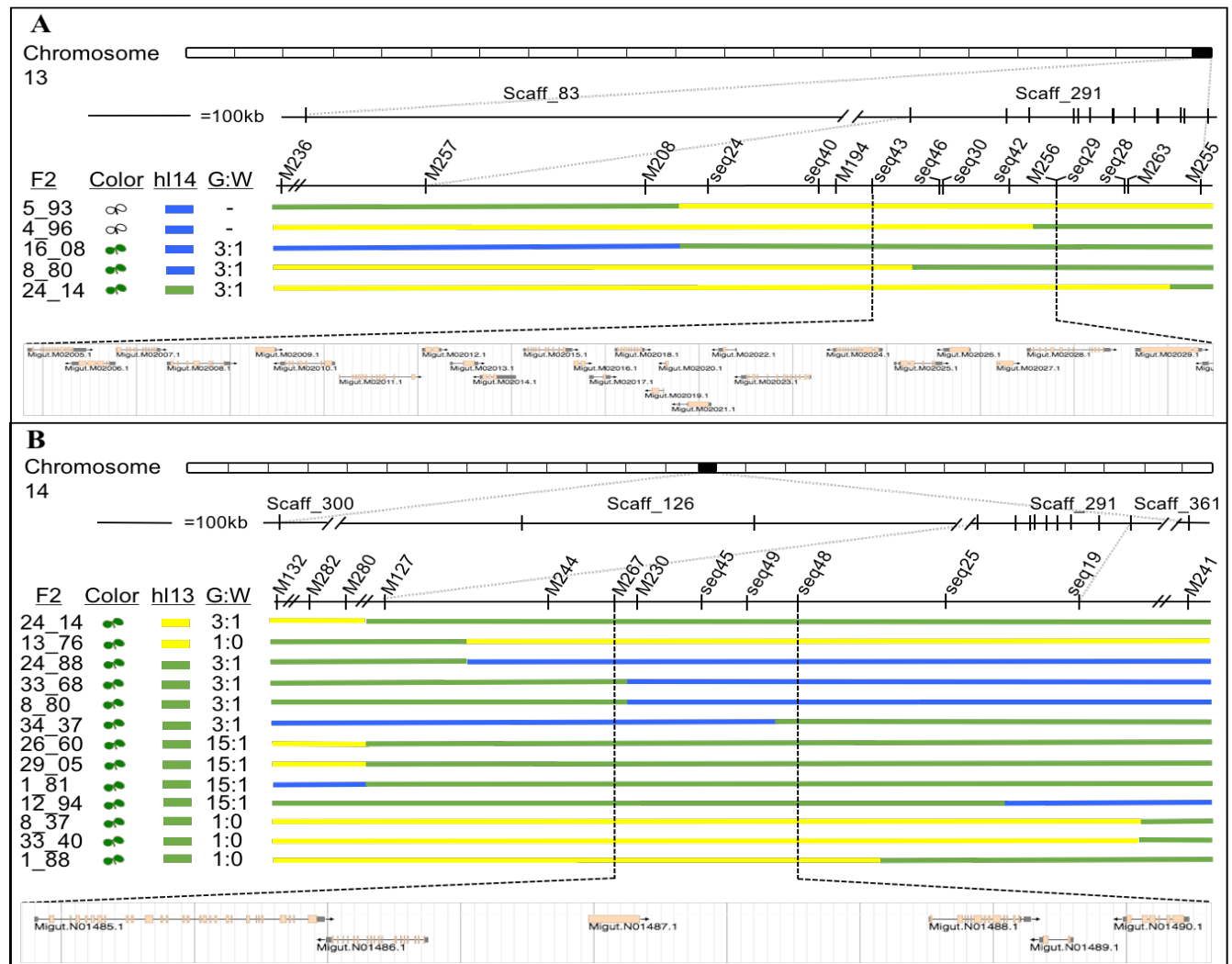


Figure 1.4: Fine map of *hl13* (A) and *hl14* (B). Chromosomal position and scaffold order (v1.1 *M. guttatus* IM62 reference genome) are shown at top, with marker positions indicated along each scaffold. Marker genotypes are yellow (homozygous DPRG102), blue (homozygous DPRN104), and green (heterozygous). The F2 individual used for mapping, color of F2 (green or white), genotype at opposite locus (*M. guttatus*, *M. nasutus*, or heterozygous), and green:white ratio of F3 progeny are given for each recombinant.

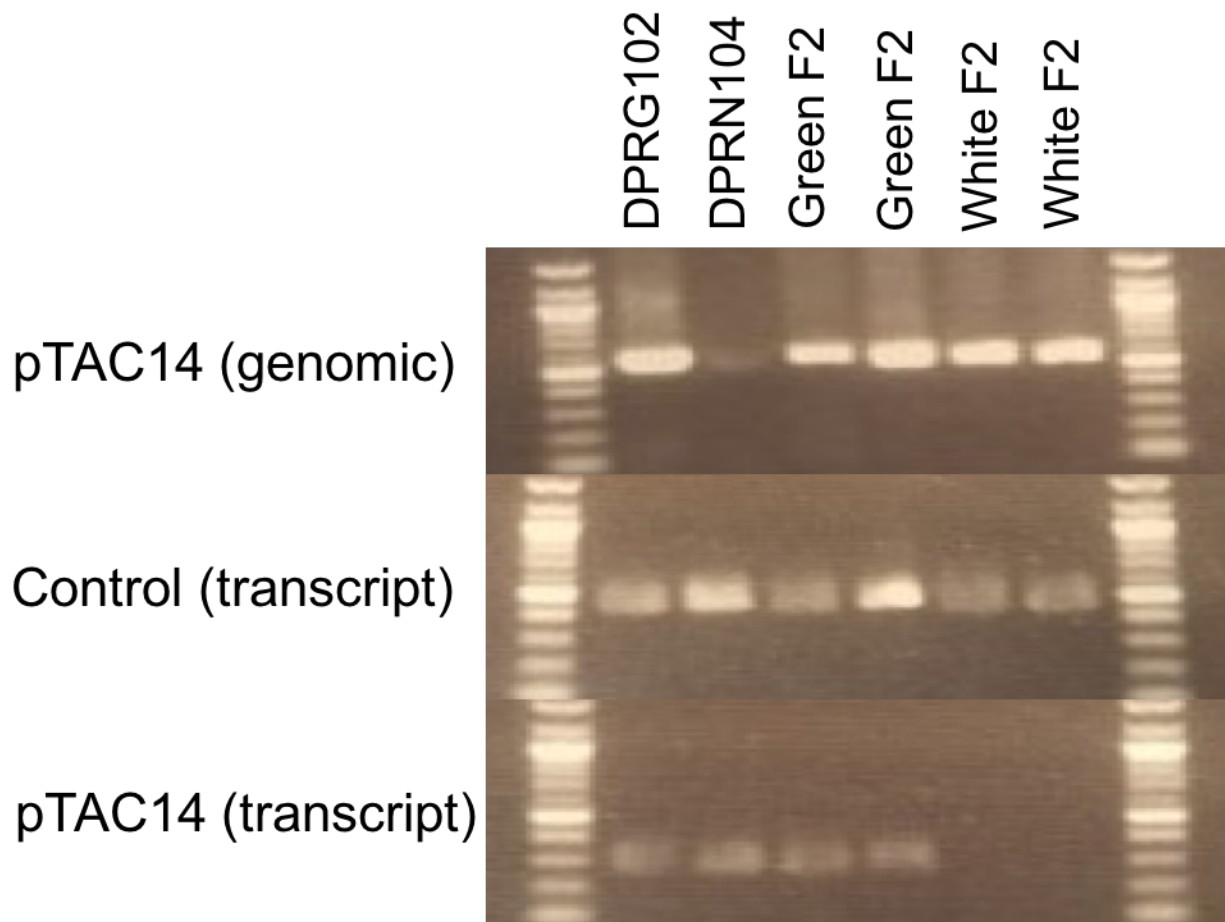


Figure 1.5: pTAC14 is not expressed in white seedlings. Different genotypes listed along top of gel. PCR products for pTAC14 genomic DNA, control transcript expression, and pTAC14 expression shown in three gels. The gel shows that white F2 seedlings do not produce pTAC14 transcripts, despite having a copy present in the genome.

A)	G1:	A	T	A	C	C	T	G	A	A	A
	G2:	G	A	G	T	C	T	G	G	G	G
	N1:	G	A	G	T	T	A	A	A	A	A

B)	<i>M. guttatus</i>				<i>M. nasutus</i>			
	<i>hl13</i>	<i>hl14</i>			<i>hl13</i>	<i>hl14</i>		
	G1	G2			N1	--		

C)	Phenotype	hl13;hl14	Sample Size	Expected	Observed
	GRN	G;G	10	G1;G2	G1;G2
	GRN	G;H	10	G1;G2	G1;G2
	WHT	G;N	16	G1;--	G1;--
	GRN	H;G	10	G1/N1;G2	G1/N1;G2
	GRN	H;H	10	G1/N1;G2	G1/N1;G2
	GRN	H;N	10	G1/N1;--	G1/N1;--
	GRN	N;G	10	N1;G2	N1;G2
	GRN	N;H	10	N1;G2	N1;G2
	GRN	N;N	10	N1;--	N1;--

Figure 1.6: Genetic mapping of pTAC14. A) pTAC14 haplotypes observed in DPRG102 (G1 and G2) and DPRN104 (N1). Sequence after haplotype names corresponds to 10 SNPs that differentiate haplotypes across 760 base pairs of sequence spanning exons 1-4 of pTAC14. B) Predicted position of haplotypes G1 and G2 in *M. guttatus* (DPRG102) and N1 in *M. nasutus* (DPRN104), where “- -” represents a position with no duplicate gene copy present. C) Association mapping of pTAC14. Phenotype is the color of DPRN104xDPRG102 F2 individual, *hl13;hl14* is the genotype of the F2 individual at *hl13* and *hl14* determined with flanking markers(G-homozygous DPRG102, N-homozygous DPRN104, H-heterozygous), sample size is the number of F2s with a particular genotypic combination sampled, expectation is based on predicted positions of pTAC14 haplotypes (B) and observed haplotypes based on sequencing haplotypes. Observed data perfectly corresponds to expectations, indicating that the duplicate copy of pTAC14 in *M. guttatus* (G2) is located at *hl14*.

Chapter 3

DYNAMICS OF HYBRID LETHALITY IN NATURE²

² Zuellig M.P. and A.L. Sweigart. To be submitted to *Evolution*.

Abstract

Theory holds that hybrid incompatibilities will degrade in the presence of interspecific gene flow unless selection within species acts to maintain incompatibility alleles. Therefore, neutrally evolving incompatibilities are expected to degrade and contribute only transiently to reproductive isolation. In this study, we investigate the frequency and distribution of hybrid incompatibility alleles that cause lethality between *Mimulus guttatus* and *M. nasutus*, two closely related species with a history of ongoing gene flow and recurrent hybridization. The hybrid incompatibility alleles underlying lethality were previously characterized genetically and are believed to evolve neutrally in this system. However, we find that incompatibility alleles are widespread in this system and occur at skewed frequencies in different species, indicating that they may be maintained at equilibrium. Furthermore, patterns of haplotype variation associated with incompatibility alleles indicate that selection at incompatibility loci, or closely linked loci, may act to maintain incompatibilities in this system. Our results indicate that hybrid lethality likely contributes to reproductive isolation in *Mimulus* and may constitute a long-term barrier to gene flow between these species.

Introduction

Hybrid incompatibilities (HIs) are believed to play an important role in the origin and maintenance of species by rendering hybrids infertile or inviable, reducing interspecific gene flow (Coyne and Orr 1989, 1997, 2004, Rieseberg et al. 1999, Ramsey et al. 2003, Payseur et al. 2004, Kay 2006, Martin and Willis 2007, Lowry et al.

2008, Teeter et al. 2008, Matsubayashi and Katakura 2009, Dopman et al. 2010, White et al. 2011), but the effect of well-characterized HIs on reproductive isolation has never been demonstrated in natural populations. HIs occur when divergent alleles from each species exhibit negative epistasis when brought together in a hybrid genome, resulting in hybrid dysfunction (Bateson 1909, Dobzhansky 1937, Muller 1942). Numerous studies have demonstrated the prevalence of HIs in plants and animals and targeted genetic studies have revealed the loci, and in some cases, the genes underlying hybrid incompatibilities (Coyne and Orr 2004, Rieseberg and Blackman 2010, Maheshwari and Barbash 2011, Sweigart and Willis 2012, Chen et al. 2016). This work has provided novel insights into the molecular functions of incompatibility alleles and provided some hints about their evolutionary histories. However, we know little about the dynamics of HIs among natural hybridizing species, although in such populations we might directly assess the relevance of HIs in the context of speciation. Do HIs affect rates of gene flow in natural populations? How do HI alleles originate within species? How are they maintained in systems experiencing interspecific gene flow?

In HI systems where interspecific gene flow is ongoing, natural selection is expected to oppose the maintenance of incompatibility alleles whenever defective hybrids are formed. Under this null model, HI alleles that are neutral within species are expected to be purged from the system in favor of compatible alleles, effectively breaking down the barrier to gene flow (Barton and Bengtsson 1986, Gavrillets 1997, Kondrashov 2003, Lemmon and Kirkpatrick 2006, and Bank et al. 2012). The maintenance of incompatibility alleles therefore requires positive selection on

incompatibility alleles *within* species, which counterbalances the negative selection imposed through deleterious hybridizations. This balance of selective forces is modulated by the amount of gene flow in the system and the genetic architecture of the incompatibility, where higher rates of gene flow counter the maintenance of incompatibilities within species (Gavrilets 1997, Bank et al. 2012). Though incompatibility alleles that reside within chromosomal rearrangements (Noor et al. 2001, Navarro and Barton 2003), exist in species with limited dispersal (Langduth et al. 2015, Cushman and Langduth 2016), or exhibit patterns of intra- and inter-genomic conflict (Crespi and Nosil 2013), may experience lower rates of gene flow and therefore may be more difficult to degrade. Thus, understanding the evolutionary forces and genetic context of HI alleles is important for inferring whether HIs play a permanent or transitory role in reproductive isolation.

A number of studies have now shown that HI alleles are often polymorphic within species, where incompatible and compatible alleles exist at a single locus *within* species (Cutter et al. 2012). Polymorphism at HI alleles may reduce the efficacy of a particular HI as a reproductive barrier, by allowing a path of interspecific gene flow at an HI locus and allowing for degradation of incompatible alleles. However, this is contingent on the rate of gene flow in the system and the context of polymorphism in the system (Bank et al. 2012). Polymorphism at a particular locus may stem from processes occurring intraspecifically, where balancing selection (Seidel et al. 2008, Sicard et al. 2015) or drift maintain polymorphism in the system or newly derived selected alleles have not yet fixed (Sweigart and Flagel 2015). Alternatively, polymorphism might result

from interspecific gene flow, where HI alleles are replaced with alleles from heterospecifics. The latter indicates that an HI is actively expressed in natural populations and is either maintained at migration-selection balance or being actively purged from the system. Investigating the source, frequency, and geographic distribution of polymorphic hybrid incompatibility alleles can shed light on the evolutionary forces acting on incompatibility alleles within and between species, their stability in a system, and provide evidence for their relative involvement as reproductive isolating barriers.

Here we take the first step toward understanding the evolutionary genetic basis of an HI causing seedling lethality between two closely related species of yellow monkeyflower, *Mimulus guttatus* and *M. nasutus*. These recently diverged species (200-500kya; Brandvain et al. 2014) co-occur throughout much of their ranges in the western United States, where reproductive isolation between sympatric populations occurs through a number of prezygotic and postzygotic barriers (see Wu et al. 2008 for a review). HIs in the *Mimulus* system exhibit a large amount of polymorphism (Vickery 1978, Case and Willis 2006, Martin and Willis 2010) with polymorphism in some cases being attributed to variation at particular HI alleles (Christie and MacNair 1987, Sweigart et al. 2007). Despite numerous reproductive barriers, sympatric populations of these species can be found throughout their shared range, where interspecific hybridization rates observed in natural populations ranged from ~0-2% (Martin and Willis 2007) and genome-wide estimates of hybridization indicate historical and ongoing gene flow between the two species (Brandvain et al. 2014, Kenney and Sweigart 2016).

Introgression in this system is largely asymmetric, with alleles from *M. nasutus* moving into the *M. guttatus* genome (Sweigart and Willis 2003, Martin and Willis 2007, Brandvain et al. 2014, Kenney and Sweigart 2016). This system of incomplete isolation with recurrent hybridization is ideal for investigating the evolutionary forces that drive the evolution of hybrid incompatibilities within populations and affect their eventual maintenance or degradation.

In a previous study, we genetically mapped two recessive nuclear loci, *h13* and *h14*, that cause hybrid lethality between inbred lines of *M. guttatus* (DPRG102) and *M. nasutus* (DPRN104). Plants that are homozygous for DPRG102 alleles at *h13* and DPRN104 alleles at *h14* always die in the cotyledon stage, due to a lack of chlorophyll production. In this study, we cross the inbred lines used in our previous experiment to wild-collected plants from throughout the *Mimulus* range and screen the F2 progeny for the presence of hybrid lethal seedlings. Furthermore, we genotype green and white F2 seedlings at markers linked to *h13* and *h14* to confirm that hybrid lethality is caused by these two loci and to collect haplotype data at these loci from natural populations.

These data allow us to address the following questions: 1) Are HI alleles fixed within species?, 2) Are HI alleles fixed within populations?, 3) Are HI alleles associated with geography?, 4) Is there evidence that hybrid lethality acts as a reproductive barrier in nature?, 5) Is hybrid lethality a stable or transient reproductive barrier in *Mimulus*?

Methods

Mimulus Lines

To investigate the frequency and distribution of *h13* and *h14* alleles in natural populations of *M. guttatus* and *M. nasutus*, we collected plants of each species from populations throughout their overlapping range in the western United States. Seed stocks used in this experiment were composed of wild-pollinated seeds (N=11, half-sib maternal families) and selfed-seed from wild-collected plants (N=22, full-sib maternal families). Table 2.1 includes information about the source and composition of seed stocks used in our experiment. Additionally, previously generated inbred ‘*tester*’ lines of *M. guttatus* (DPRG102) and *M. nasutus* (DPRN104) that are homozygous for *h13* and *h14* incompatibility alleles, respectively, were used as paternal parents in all of our crosses.

All plants used in this study were planted on moist Fafard-3B potting soil, cold stratified for one week at 2C, allowed to germinate at 27C in the University of Georgia greenhouses, and each seedling was subsequently transplanted to its own 2.5” pot. All plants were grown under 16 hour days and flowered approximately one month after germination.

Crossing Design

We assessed the presence/absence of hybrid lethality alleles in wild-collected plants by assessing the green:white ratio of F2 families in crosses between wild plants and our two *tester* lines, which are homozygous for different hybrid lethality alleles. We

assessed hybrid lethality loci in 1-8 plants per seed stock, where more crosses increase the likelihood of observing multiple hybrid lethality alleles in a single seed stock. We set an upper limit of observing two distinct hybrid lethality alleles per seed stock, despite the possibility of observing more alleles per locus in our half-sib seed stocks. Flowers from wild plants were used emasculated prior to pollen dehiscence and pollinated with either DPRG102 or DPRN104 pollen. A single F1 seed from each cross was subsequently germinated and self-fertilized to generate a F2 family. We tallied the green:white seedling ratio of F2 families, counting a minimum 87 F2 seedlings per cross (FDR <0.01). Crosses that segregate 15:1 green:white F2 seedlings presumably carry *h/13* or *h/14*, depending on whether they were crossed to DPRN104 or DPRG102, respectively.

Genotyping

Previous work in *Mimulus guttatus* suggests that multiple independent incompatibilities that cause seedling lethality segregate in this system (Willis 1992), so the presence of white seedlings in the F2 progeny alone is not sufficient for identifying lines carrying *h/13* and *h/14*. However, since lethality loci are mapped in our system (Chapter 1), we overcame this problem by genotyping green and white seedlings from each F2 population at markers linked to *h/13* and *h/14*. White F2 seedlings caused by the *h/13:h/14* incompatibility always exhibit genotypic distortion at these loci, whereas green seedlings from the same population will segregate near-Mendelian frequencies at linked markers. For instance, white seedlings from a cross between wild-collected “plant 1” and DPRG102 (homozygous for *h/13*), are expected to be homozygous for DPRG102

alleles at *h/13*-linked markers and homozygous for “plant 1” alleles at *h/14*-linked markers, if hybrid lethality is caused by the *h/13:h/14* incompatibility. Whereas green F2s and white seedlings caused by a separate incompatibility will not exhibit this type of distortion (see Figure 2.1 for more details).

In addition to confirming that white seedlings are caused by the *h/13:h/14* incompatibility, genotyping green and white F2 seedlings at markers linked to each locus allowed us to investigate patterns of variation in regions associated with compatible and incompatible alleles at each hybrid incompatibility locus. For each wild individual, we recorded the fragment size of the four markers used to genotype green and white F2 seedlings (two-linked markers per hybrid lethality locus). The marker sizes were then combined, to yield a single two-locus ‘haplotype’ for each individual. For instance, wild-collected ‘plant 1’ that carries an incompatible copy of *h/14* and fragment sizes of 283 (at *h/14*-linked marker M241) and 500 (at *h/14*-linked marker M132), carries the two-locus haplotype ‘283500’.

DNA for this experiment was extracted from green and white seedlings using a standard CTAB-chloroform protocol (Doyle and Doyle 1987) modified for use in 96-well format. Genotyping was performed using four exon-primed size polymorphic markers containing 5’ fluorescent tags (6-FAM or HEX) that are linked to *h/13* and *h/14*. Markers M208 and M236 are linked to *h/13* and are separated by approximately 9 centimorgans, whereas markers M132 and M241 flank *h/14* and are separated by approximately 11 centimorgans. A standard touchdown PCR protocol was used for amplification and genotyping was performed on an ABI3730XL automated DNA sequencer at the Georgia

Genomics Facility. Fragment analysis was performed manually using GENEMARKER software (Hulce et al. 2011).

Results

Frequency and distribution of hybrid lethality alleles in nature

We examined *h13* alleles in 18 populations of *M. guttatus* and 9 populations of *M. nasutus*, sampling a total of 54 and 23 alleles in each species, respectively. *M. guttatus* carries a relatively high frequency of *h13* incompatibility alleles (70% global frequency) and the incompatibility allele was present in at least one individual in all 18 populations studied (Figure 2.2). *M. nasutus* also carries *h13* incompatibility alleles, but they are at much lower frequency (13% global allele frequency) and present in only 3 of the populations sampled (Figure 2.2). Sampling the *h14* incompatibility allele revealed a different pattern, with 11 of 18 *M. guttatus* and 8 of 11 *M. nasutus* populations containing at least one individual with this allele. Global allele frequencies of the *h14* incompatibility allele are 28% and 69% in *M. guttatus* and *M. nasutus*, respectively. Both species exhibit high rates of polymorphism at hybrid lethality loci, but *h13* is generally more abundant in *M. guttatus*, where it was originally identified and *h14* is generally more abundant in *M. nasutus*, where it was originally identified. Both incompatibility alleles are found across a broad geographic range and appear to be polymorphic across the range. The presence of incompatibility alleles at both loci within single populations implies that hybrid lethality is expressed at some rate in nature and may contribute to reproductive isolation. Among *M. guttatus* populations for which both

alleles were surveyed, 61.5% (8/13) carry incompatibility alleles at both loci, including two populations that are sympatric with *M. nasutus*. None of the nine *M. nasutus* populations for which both loci were tested carried incompatibility alleles at both loci.

Genetic variation at hybrid lethality loci

To investigate the source of polymorphism at *hl13* and *hl14* in *Mimulus*, we genotyped green and white F2s produced in our population sampling scheme (above) at two markers linked to each locus. Haplotypes were defined as the combination of allele sizes at the first and second marker, where markers were spaced 9 and 11 centimorgans apart at *hl13* and *hl14*, respectively. This provided us with important data regarding the diversity of haplotypes present in compatible versus incompatible alleles at *hl13* and *hl14* and allowed us to infer whether alleles originated in *M. guttatus* or *M. nasutus* (see Figure 2.1 for additional methodological details).

We observed twelve distinct haplotypes at *hl13* in *M. guttatus* and five in *M. nasutus*. We divided the haplotypes observed in compatible versus incompatible lines of each species (Figure 2.3). We observed nine different haplotypes among the 42 *M. guttatus* individuals carrying incompatible alleles at *hl13*, with the most abundant haplotype found in 38% of individuals. In contrast, *M. guttatus* individuals carrying compatible alleles at *hl13* carry five haplotypes, three that are not observed in incompatible *M. guttatus* individuals. Furthermore, the two most abundant *M. guttatus* compatible haplotypes, 422531 and 495526, represent 71% of total haplotype diversity, whereas these same haplotypes represent only 5% of incompatible *M. guttatus* and

73% of compatible *M. nasutus* haplotypes, respectively. A Fishers exact test was used to demonstrate an overabundance of these two haplotypes, that appear to stem from *M. nasutus*, in compatible relative to incompatible *M. guttatus* individuals ($p < 0.0001$). Four haplotypes were observed among compatible *M. nasutus* individuals and three among *M. nasutus* incompatible individuals. Interestingly, 2/5 haplotypes haplotypes observed in *M. nasutus* incompatible individuals were 495529, a haplotype that is otherwise specific to *M. guttatus* and represents the most common haplotype among compatible *M. guttatus* individuals.

We observed seventeen distinct haplotypes at *hl14* in *M. guttatus* and four in *M. nasutus*. We divided the haplotypes observed in compatible versus incompatible lines of each species (Figure 2.4). We observed 16 different haplotypes among the 31 *M. guttatus* individuals carrying compatible alleles at *hl14*, with the most abundant haplotype found in 29% of individuals. In contrast, *M. guttatus* individuals carrying incompatible alleles at *hl14* were dominated by a single haplotype (85% of individuals), which was also present in every *M. nasutus* individual carrying incompatible *hl14* alleles. Three haplotypes were observed among the *M. nasutus* individuals carrying compatible *hl14* alleles, one shared with compatible *M. guttatus*, one shared with incompatible *M. nasutus*, and another unique to the group. A Fishers exact test was used to demonstrate an overabundance of the '283500' haplotype in incompatible relative compatible *M. guttatus* individuals ($p < 0.0001$).

Discussion

The evolution of reproductive barriers is critical for the evolution and maintenance of species. Numerous studies have demonstrated the efficacy of various reproductive barriers at inhibiting gene flow among sympatric species pairs (Coyne and Orr 2004, Lowry et al. 2008), but the evolutionary mechanisms underlying their evolution and maintenance remains largely unknown. Divergent natural selection seems a likely mechanism for the maintenance of prezygotic isolating barriers, as indicated by numerous studies demonstrating local adaptation and evidence for divergent selection on particular traits and loci (Schluter 2009). However, understanding the evolution of postzygotic barriers is less straightforward, since incompatibility phenotypes have no clear relationship with ecology and the phenotypic effects of incompatibility genes are largely unknown. While a number of HI genes (Ting et al. 1998, Presgraves et al. 2003, Brideau et al. 2006, Oliver et al. 2009, Tang and Presgraves 2009) and loci (Sweigart and Flagel 2015) harbor signatures of positive selection, the form of selection is largely unknown and in many cases seems to be conflictual in nature, rather than ecological (Presgraves 2010, Maheshwari and Barbash 2011). Though, in some cases, selection by local pathogen regimes may be a critical driver of adaptive divergence at incompatibility loci (Bomblies et al. 2007, Bomblies 2010) and balancing selection underlies HI loci in at least two systems (Sicard et al., 2015, Seidel et al. 2008). Additionally, a number of HIs appear to result from purely mutational mechanisms, such as reciprocal gene loss and gene transposition (Masly et al. 2006, Bikard et al. 2006, Mizuta et al. 2010, Yamagata et al. 2010), and are likely subject to evolution by genetic

drift (Werth and Windham 1991, Lynch and Force 2000, Lynch and Conery 2000). Thus, many HI systems likely evolve in a manner akin to mutation-order speciation, rather than through ecological selection. This begs the question of how HIs are maintained upon secondary contact? Theoretical models indicate that gene flow will degrade HIs in the absence of positive divergent selection (Gavrilets 1997, Bank et al. 2012). Though, some have postulated that HI alleles ‘locked’ in chromosomal inversions may be shielded from degradation through drastic reductions in gene flow at HI alleles (Noor et al. 2001, Navarro and Barton 2003). Investigating natural variation at genetically characterized HI loci in hybridizing species allows us to identify the effects of gene flow on the maintenance of HIs in natural populations, providing a powerful framework for testing theoretical predictions.

In a previous study, we mapped two loci that underlie hybrid lethality between two sympatric species of *Mimulus* that exhibit patterns of ongoing gene flow. Our previous study shows that hybrid lethality is caused by duplicate gene loss and no evidence of incompatibility alleles being locked within chromosomal inversions exists. Thus, hybrid lethality in *Mimulus* appears to result from mutational evolution and is likely subject to genetic drift in natural populations. Here we investigate the distribution and frequency of hybrid lethality alleles in the range of overlap between *M. guttatus* and *M. nasutus*, demonstrating that both incompatibility alleles are widespread and polymorphic in both species. We then investigate genetic variation in the regions surrounding each locus, providing strong evidence for ongoing gene flow at HI loci. Furthermore, haplotype structure surrounding introgressed HI alleles indicates that selection against

introgression of heterospecific genetic variation occurs at both loci. This intriguing result indicates that some mechanism, possibly divergent selection at HI alleles or closely linked genes, may allow this incompatibility to remain at equilibrium in migration-selection balance.

Natural Variation at Hybrid Lethality Alleles

Our sampling of HI alleles in natural populations of *M. guttatus* and *M. nasutus* revealed that both species harbor a large amount of variation at hybrid lethality alleles. In *M. guttatus*, we found that *h13* and *h14* incompatibilities are both widespread and polymorphic (Figure 2.2, Table 2.1). Incompatible *h13* and compatible *h14* alleles occur at similar frequencies in *M. guttatus*, 70% and 72%, respectively. This indicates that the most common two-locus genotype in *M. guttatus* is incompatible *h13* with compatible *h14*, which is the genotype of *M. guttatus* sampled in our previous mapping experiment (Chapter 1). Furthermore, these two alleles are found in the majority of populations sampled, whereas compatible *h13* and incompatible *h14* were found in fewer populations. Though, it is probable that increased sampling within populations would identify even more populations carrying compatible *h13* and incompatible *h14* alleles, as these alleles were disproportionately observed in populations with the greatest sampling intensity in our study. Our most densely sampled population, DPR, exhibited allele frequencies similar to the global allele frequencies reported for *M. guttatus*, suggesting that the frequency of HI alleles in individual populations might reflect global allele frequencies.

We also found that *h13* and *h14* incompatibilities are widespread and polymorphic in *M. nasutus*, though the most abundant alleles are opposite of those observed in *M. guttatus* indicating that the most common genotypic orientation is compatible *h13* and incompatible *h14* (Figure 2.2, Table 2.1). Furthermore, these two alleles were found in the majority of populations sampled. Additionally, global allele frequencies are largely skewed in this species, with incompatible *h13* alleles occurring at low (13%) frequency and compatible *h14* alleles occurring at relatively intermediate frequency (31%).

The patterns of HI distribution and frequency observed here exhibit similarities and differences with other HIs that have been studied in *Mimulus*. Christie and Macnair (1987) surveyed allelic variation underlying two independent two-locus HI systems that segregate within *M. guttatus* using genetic complementation tests. Both alleles underlying their Cerig/C10 seedling lethality HI appear to be widespread and segregate at intermediate frequencies within populations. In contrast, they observed strong geographic structuring for the separate C7/U8 lethality system, yet incompatibility alleles still segregated at intermediate frequencies in many populations. Similarly, Sweigart et al. (2007) showed that the *M. guttatus hms1* allele, which interacts with the *M. nasutus hms2* allele to cause hybrid male sterility, is geographically constrained to a single population in Oregon where it segregates at intermediate frequency. Their study also demonstrates that the *hms2* allele is widespread and appears fixed in *M. nasutus*, in striking contrast to the HI alleles observed in *M. nasutus* here. Polymorphism at HI alleles appears to be a common feature in *M. guttatus*, perhaps due to the large amount

of segregating variation in this species, its outcrossing life history, or due to ongoing gene flow with heterospecifics. However, polymorphism at HI alleles in *M. nasutus* is somewhat surprising, as this species is highly inbred, contains relatively low levels of genetic variation, and rarely obtains genetic material through introgression.

Our sampling scheme demonstrates that *M. guttatus* carries incompatible *h13* alleles at relatively high frequencies, whereas *M. nasutus* carries incompatible *h14* alleles at relatively high frequencies. Furthermore, species with different incompatibility alleles are in close proximity to one another, including two sympatric populations where the species live within centimeters of one another. Taken together, these data indicate that hybrid lethality has the potential to contribute to reproductive isolation between these species. The presence of polymorphism at both HI alleles in both species also indicates that hybrid lethality may also act as a barrier to intraspecific gene flow. The action of the same incompatibility within and between species indicates that this incompatibility may be transient in nature, as the negative fitness consequences of conspecific mating in combination with selection imposed against interspecific hybrids should rapidly purge the system of incompatibility alleles. Yet, both alleles are observed across a broad geographic range at relatively high frequencies, implying that they might be maintained at equilibrium.

By sampling genetic variation surrounding HI alleles, we provide evidence that introgression is at least partially responsible for polymorphism at HI alleles. At *h13*, we sampled a total of 13 different haplotypes defined by two markers that are separated by 9 centimorgans. We observe 12 of these haplotypes in *M. guttatus*, however the

distribution of haplotypes associated with compatible versus incompatible *h13* alleles in this species show some striking differences (Figure 2.3). Half of *M. guttatus* individuals with incompatible *h13* alleles carry one of six unique haplotypes, while the majority of remaining individuals carry a single haplotype, 495529, that represents the most abundant haplotype in *M. guttatus*. In contrast, *M. guttatus* with compatible *h13* alleles carry haplotypes that are similar to those observed in compatible *h13* *M. nasutus* individuals, implying that compatible *h13* alleles in *M. guttatus* have introgressed from *M. nasutus*. By assuming that haplotype 422531 originates in *M. nasutus*, we show that a disproportionate percentage of *M. guttatus* individuals with compatible *h13* alleles carry genetic material associated with *M. nasutus* (Fishers exact test, $p < 0.0001$).

At *h14*, we sampled a total of 18 different haplotypes defined by two markers separated by 11 centimorgans. We observed a striking contrast in haplotype structure associated with *M. guttatus* compatible versus incompatible alleles (Figure 2.4, Fishers exact test $p < 0.0001$), where incompatible *h14* haplotypes are dominated by the 283500 haplotype, which is fixed in *M. nasutus* incompatible individuals. In contrast, *M. guttatus* compatible individuals carry an assortment of haplotypes and only one individual carries the 283500 haplotype. This pattern implies introgression of incompatible *h14* alleles from *M. nasutus* into *M. guttatus* without rapid decay of haplotype structure.

The presence of large haplotypes associated with different HI alleles is, in itself, a curious observation. In *M. guttatus* linkage disequilibrium rapidly decays (Brandvain et al. 2014), implying that the large *M. nasutus* haplotypes observed in *M. guttatus* are recent in origin. We suspect that two non-mutually exclusive hypotheses explain this

phenomenon. First, *M. guttatus* individuals harboring large fragments of *M. nasutus* variation may represent early generation hybrids that simply have not had enough time to recombine, allowing for the retention of large haplotype blocks in regions associated with HI alleles. Second, selection against *M. nasutus* alleles in regions associated with HI alleles may purge *M. nasutus* variation from the genome before haplotypes ever have the opportunity to recombine, resulting in the presence of only large haplotypes that haven't yet been selected against. Selection presumably favors introgression of compatible *hl13* alleles into *M. guttatus*, which should result in more time for recombination and the production of recombinant haplotypes, whereas selection against incompatible *hl14* alleles should quickly purge *M. nasutus* variation from the *M. guttatus* genome (Figure 2.5). These expectations are born out in our data, though the high number of unrecombined compatible *M. nasutus hl13* haplotypes in *M. guttatus* is perplexing and may indicate that selection against *M. nasutus* alleles in this region, or on *hl13* itself, contributes to the maintenance of incompatible *hl13* alleles in *M. guttatus*.

Our study on hybrid lethality alleles in natural populations of *Mimulus* sheds light on the dynamics of HIs in nature. Hybrid lethality, like many other HIs (Cutter 2012), is polymorphic within species and this polymorphism is likely explained by interspecific hybridization that allows HI alleles to transcend the species barrier. Despite ongoing gene flow, we observe hybrid lethality alleles in populations throughout the species range and haplotype structure consistent with selection against large amounts of introgression at HI loci. This implicates a role for selection, or some other phenomenon, acting on HI alleles within species that contributes to the maintenance of hybrid lethality

in this system. Future studies investigating the fitness effects of HIs within populations, closely linked genes under positive selection, and genetic mechanisms such as transmission ratio distortion will provide valuable insight into the maintenance of hybrid lethality in this system.

References

- Bank, C., R. Bürger, and J. Hermisson. 2012. The limits to parapatric speciation: Dobzhansky–Muller incompatibilities in a continent–island model. *Genetics* 191:845-863.
- Bateson, W. 1909. Heredity and variation in modern lights. *Darwin and modern science* 85:101.
- Bomblies, K. 2010. Doomed lovers: mechanisms of isolation and incompatibility in plants. *Annual review of plant biology* 61:109-124.
- Bomblies, K., J. Lempe, P. Eppe, N. Warthmann, C. Lanz, J. L. Dangel, and D. Weigel. 2007. Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biol* 5:e236.
- Brandvain, Y., A. M. Kenney, L. Flagel, G. Coop, and A. L. Sweigart. 2014. Speciation and introgression between *Mimulus nasutus* and *Mimulus guttatus*. *PLoS Genet* 10:e1004410.
- Brideau, N. J., H. A. Flores, J. Wang, S. Maheshwari, X. Wang, and D. A. Barbash. 2006. Two Dobzhansky-Muller Genes Interact to Cause Hybrid Lethality in *Drosophila*. *Science* 314:1292-1295.
- Case, A. L. and J. H. Willis. 2008. Hybrid male sterility in *Mimulus* (Phrymaceae) is associated with a geographically restricted mitochondrial rearrangement. *Evolution* 62:1026-1039.
- Chen, C., E. Zhiguo, and H.-X. Lin. 2016. Evolution and Molecular Control of Hybrid Incompatibility in Plants. *Frontiers in plant science* 7.
- Christie, P. and M. R. Macnair. 1987. The distribution of postmating reproductive isolating genes in populations of the yellow monkey flower, *Mimulus guttatus*. *Evolution*:571-578.
- Coyne, J. A. and H. A. Orr. 1989. Patterns of speciation in *Drosophila*. *Evolution*:362-381.
- Coyne, J. A. and H. A. Orr. 1997. " Patterns of speciation in *Drosophila*" revisited. *Evolution* 51:295-303.
- Coyne, J. A. and H. A. Orr. 2004. *Speciation*. Sunderland, MA. Sinauer Associates, Inc.
- Crespi, B. and P. Nosil. 2013. Conflictual speciation: species formation via genomic conflict. *Trends in ecology & evolution* 28:48-57.
- Cushman, S. A. and E. L. Landguth. 2016. Spatially Heterogeneous Environmental Selection Strengthens Evolution of Reproductively Isolated Populations in a Dobzhansky–Muller System of Hybrid Incompatibility. *Frontiers in Genetics* 7.
- Cutter, A. D. 2012. The polymorphic prelude to Bateson–Dobzhansky–Muller incompatibilities. *Trends in ecology & evolution* 27:209-218.
- Dobzhansky, T. and T. G. Dobzhansky. 1937. *Genetics and the Origin of Species*. Columbia University Press.
- Dopman, E. B., P. S. Robbins, and A. Seaman. 2010. Components of reproductive isolation between North American pheromone strains of the European corn borer. *Evolution* 64:881-902.
- Doyle, J. and J. Doyle. 1987. Genomic plant DNA preparation from fresh tissue-CTAB method. *Phytochem Bull* 19:11-15.
- Gavrilets, S. 1997. Hybrid Zones With Dobzhansky-Type Epistatic Selection. *Evolution* 51:1027-1035.

- Hulce, D., X. Li, T. Snyder-Leiby, and C. S. Johathan Liu. 2011. GeneMarker® Genotyping Software: Tools to Increase the Statistical Power of DNA Fragment Analysis. *Journal of Biomolecular Techniques* : JBT 22:S35-S36.
- Kay, K. M. 2006. Reproductive isolation between two closely related hummingbird-pollinated neotropical gingers. *Evolution* 60:538-552.
- Kenney, A. M. and A. L. Sweigart. 2016. Reproductive isolation and introgression between sympatric *Mimulus* species. *Molecular ecology* 25:2499-2517.
- Kondrashov, A. S. and M. Morgan. 2003. Accumulation of Dobzhansky-Muller incompatibilities within a spatially structured population. *Evolution* 57:151-153.
- Landguth, E. L., N. A. Johnson, and S. A. Cushman. 2015. Clusters of incompatible genotypes evolve with limited dispersal. *Frontiers in genetics* 6:151.
- Lemmon, A. R. and M. Kirkpatrick. 2006. Reinforcement and the genetics of hybrid incompatibilities. *Genetics* 173:1145-1155.
- Lowry, D. B., J. L. Modliszewski, K. M. Wright, C. A. Wu, and J. H. Willis. 2008. The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 363:3009-3021.
- Lynch, M. and J. S. Conery. 2000. The Evolutionary Fate and Consequences of Duplicate Genes. *Science* 290:1151-1155.
- Lynch, M. and A. G. Force. 2000. The Origin of Interspecific Genomic Incompatibility via Gene Duplication. *The American Naturalist* 156:590-605.
- Maheshwari, S. and D. A. Barbash. 2011. The Genetics of Hybrid Incompatibilities. *Annual Review of Genetics* 45:331-355.
- Martin, N. H. and J. H. Willis. 2007. Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric *Mimulus* species. *Evolution* 61:68-82.
- Martin, N. H. and J. H. Willis. 2010. Geographical variation in postzygotic isolation and its genetic basis within and between two *Mimulus* species. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365:2469-2478.
- Masly, J. P., C. D. Jones, M. A. Noor, J. Locke, and H. A. Orr. 2006. Gene transposition as a cause of hybrid sterility in *Drosophila*. *Science* 313:1448-1450.
- Matsubayashi, K. W. and H. Katakura. 2009. Contribution of multiple isolating barriers to reproductive isolation between a pair of phytophagous ladybird beetles. *Evolution* 63:2563-2580.
- Mizuta, Y., Y. Harushima, and N. Kurata. 2010. Rice pollen hybrid incompatibility caused by reciprocal gene loss of duplicated genes. *Proceedings of the National Academy of Sciences* 107:20417-20422.
- Muller, H. J. 1942. Isolating mechanisms, evolution and temperature. Pp. 71-125. *Biol. Symp.*
- Navarro, A. and N. H. Barton. 2003. Chromosomal Speciation and Molecular Divergence--Accelerated Evolution in Rearranged Chromosomes. *Science* 300:321-324.
- Noor, M. A., K. L. Grams, L. A. Bertucci, and J. Reiland. 2001. Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences* 98:12084-12088.

- Oliver, P. L., L. Goodstadt, J. J. Bayes, Z. Birtle, K. C. Roach, N. Phadnis, S. A. Beatson, G. Lunter, H. S. Malik, and C. P. Ponting. 2009. Accelerated evolution of the Prdm9 speciation gene across diverse metazoan taxa. *PLoS Genet* 5:e1000753.
- Payseur, B. A., J. G. Krenz, and M. W. Nachman. 2004. Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. *Evolution* 58:2064-2078.
- Presgraves, D. C. 2010. The molecular evolutionary basis of species formation. *Nature Reviews Genetics* 11:175-180.
- Presgraves, D. C., L. Balagopalan, S. M. Abmayr, and H. A. Orr. 2003. Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. *Nature* 423:715-719.
- Ramsey, J., H. Bradshaw Jr, and D. W. Schemske. 2003. Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae). *Evolution* 57:1520-1534.
- Rieseberg, L. H. and B. K. Blackman. 2010. Speciation genes in plants. *Annals of Botany*:mcq126.
- Rieseberg, L. H., J. Whitton, and K. Gardner. 1999. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics* 152:713-727.
- Schluter, D. 2009. Evidence for ecological speciation and its alternative. *Science* 323:737-741.
- Seidel, H. S., M. V. Rockman, and L. Kruglyak. 2008. Widespread genetic incompatibility in *C. elegans* maintained by balancing selection. *science* 319:589-594.
- Sicard, A., C. Kappel, E. B. Josephs, Y. W. Lee, C. Marona, J. R. Stinchcombe, S. I. Wright, and M. Lenhard. 2015. Divergent sorting of a balanced ancestral polymorphism underlies the establishment of gene-flow barriers in *Capsella*. *Nature Communications* 6:7960.
- Sweigart, A. L. and L. E. Flagel. 2015. Evidence of natural selection acting on a polymorphic hybrid incompatibility locus in *Mimulus*. *Genetics* 199:543-554.
- Sweigart, A. L., A. R. Mason, and J. H. Willis. 2007. Natural variation for a hybrid incompatibility between two species of *Mimulus*. *Evolution* 61:141-151.
- Sweigart, A. L. and J. H. Willis. 2012. Molecular evolution and genetics of postzygotic reproductive isolation in plants. *F1000 biology reports* 4:23.
- Tang, S. and D. C. Presgraves. 2009. Evolution of the *Drosophila* nuclear pore complex results in multiple hybrid incompatibilities. *Science* 323:779-782.
- Teeter, K. C., B. A. Payseur, L. W. Harris, M. A. Bakewell, L. M. Thibodeau, J. E. O'Brien, J. G. Krenz, M. A. Sans-Fuentes, M. W. Nachman, and P. K. Tucker. 2008. Genome-wide patterns of gene flow across a house mouse hybrid zone. *Genome research* 18:67-76.
- Ting, C.-T., S.-C. Tsaur, M.-L. Wu, and C.-I. Wu. 1998. A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science* 282:1501-1504.
- Vickery Jr, R. K. 1978. Case studies in the evolution of species complexes in *Mimulus*. Pp. 405-507. *Evolutionary biology*. Springer.
- Werth, C. R. and M. D. Windham. 1991. A Model for Divergent, Allopatric Speciation of Polyploid Pteridophytes Resulting from Silencing of Duplicate-Gene Expression. *The American Naturalist* 137:515-526.

- White, M. A., B. Steffy, T. Wiltshire, and B. A. Payseur. 2011. Genetic dissection of a key reproductive barrier between nascent species of house mice. *Genetics* 189:289-304.
- Willis, J. H. 1992. Genetic analysis of inbreeding depression caused by chlorophyll-deficient lethals in *Mimulus guttatus*. *HEREDITY-LONDON* 69:562-562.
- Wu, C., D. Lowry, A. Cooley, K. Wright, Y. Lee, and J. Willis. 2008. *Mimulus* is an emerging model system for the integration of ecological and genomic studies. *Heredity* 100:220-230.
- Yamagata, Y., E. Yamamoto, K. Aya, K. T. Win, K. Doi, Sobrizal, T. Ito, H. Kanamori, J. Wu, T. Matsumoto, M. Matsuoka, M. Ashikari, and A. Yoshimura. 2010. Mitochondrial gene in the nuclear genome induces reproductive barrier in rice. *Proceedings of the National Academy of Sciences* 107:1494-1499.

Table 2.1 Frequency of *h/13:h/14* incompatibility haplotypes identified in natural populations

Name	Latitude	Longitude	<i>h/13</i> Freq. (# sampled)	<i>h/13</i> haplotypes (# sampled)	<i>h/14</i> Freq. (# sampled)	<i>h/14</i> haplotypes (# sampled)
<i>M. guttatus</i>						
ARB	37.4608	-121.7989	0.5 (1/2)	I_495528 (1) C_495528 (1)	0 (0/1)	C_289464 (1)
BHC	38.1436	-120.7394	1 (1/1)	I_495531 (1)	-	-
BHI	41.2319	-123.9883	-	-	0 (0/1)	C_289464 (1)
BSR	40.7308	-124.3675	-	-	0 (0/1)	C_289464 (1)
CVR	38.3958	-123.2261	1 (1/1)	I_495528 (1)	-	-
DTC	39.1649	-120.1736	1 (1/1)	I_422529 (1)	-	-
EMH	37.4489	-121.7653	0.33 (1/3)	C_422531 (1) C_495526 (2)	0.2 (1/5)	C_287483 (4)
HWF	38.0008	-120.6504	1 (1/1)	I_495528 (1)	0 (0/1)	I_283500 (1)
LCU	47.6844	-123.4428	-	-	1 (1/1)	I_283500 (1)
LSG	43.9298	-122.7836	-	-	1 (1/1)	I_283500 (1)
MAC	38.1456	-121.0520	1 (1/1)	I_495528 (1)	-	-
MCC	37.5852	-120.2702	1 (1/1)	I_495531 (1)	1 (2/2)	I_283482 (1)
MED	37.8260	-120.3329	0.71 (17/24)	I_495528 (7) I_489528 (6) I_495531 (3) I_422528 (1) I_493529 (1)	0.16 (4/25)	I_283500 (4) C_283498 (1) C_289457 (1) C_289465 (5) C_289469 (1) C_289490 (1) C_289503 (1) C_292464 (1) C_292485 (1) C_292490 (1) C_296485 (2) C_296500 (1)
MTL	37.8426	-119.9748	1 (1/1)	C_422526 (1) I_489528 (1)	0 (0/1)	-
NAF	37.9266	-120.7258	0.67 (2/3)	I_426526 (1) I_489528 (1) I_495528 (1)	0 (0/2)	C_286485 (2)
NAN	37.9266	-120.7258	1 (1/1)	-	0.5 (1/2)	I_286486 (1)
NAP	37.9266	-120.7258	1 (1/1)	I_495528 (1)	0.4 (2/5)	I_295469 (1) C_292485 (1)
ONE	45.5793	-122.0736	0.5 (1/2)	I_495528 (1)	0.5 (1/2)	I_283500 (1) C_289406 (1)
RFA	46.8533	-123.4467	-	-	1 (1/1)	I_283500 (1)
RHN	37.8455	-120.4643	1 (1/1)	I_489528 (1)	0 (0/2)	C_291464 (1)
RHR	37.8578	-120.4610	1 (1/1)	I_495528 (1)	0 (0/1)	-

TUO	37.8143	-119.8613	0.33 (2/6)	I_493528 (1) I_493518 (1) C_422531 (1)	0.33 (1/3)	I_283500 (1) C_283500 (1)
UMQ	43.6553	-123.8162	1 (1/1)	I_495528 (1)	0.5 (1/2)	I_283500 (1) C_289465 (1)
USF	37.8222	-120.1018	1 (1/1)	I_495528 (1)	0 (0/1)	C_287496 (1)
<u>M. nasutus</u>						
BRD	37.7167	-119.6489	-	-	0.5 (1/2)	I_283500 (1)
BRG	38.0517	-120.4711	0 (0/5)	C_422531 (4)	0.8 (4/5)	I_283500 (4)
BRN	38.0517	-120.4711	0 (0/1)	-	-	-
MED	37.8260	-120.3329	0 (0/1)	C_422528 (1)	1 (2/2)	I_283500 (1)
MOK	38.3895	-120.5263	0.5 (1/2)	I_495526 (1) C_495526 (1)	0 (0/1)	-
MTB	37.9224	-119.7682	-	-	1 (1/1)	I_283500 (1)
MTC	37.9227	-119.8044	0.25 (1/4)	I_495528 (1) C_422531 (2)	0 (0/3)	C_283469 (2)
NFM	38.4708	-120.3723	1 (1/1)	I_495526 (1)	0 (0/1)	C_283500 (1)
PNM	36.4946	-121.1830	-	-	1 (1/1)	I_283500 (1)
SFM	38.3691	-120.5468	0 (0/1)	-	1 (2/2)	I_283500 (2)
TUO	37.8143	-119.8613	0 (0/2)	-	1 (2/2)	I_283500 (2)
TUP	37.8572	-119.8673	0 (0/6)	C_481531 (2) C_422531 (1)	0.83 (5/6)	I_283500 (5) C_286483 (1)

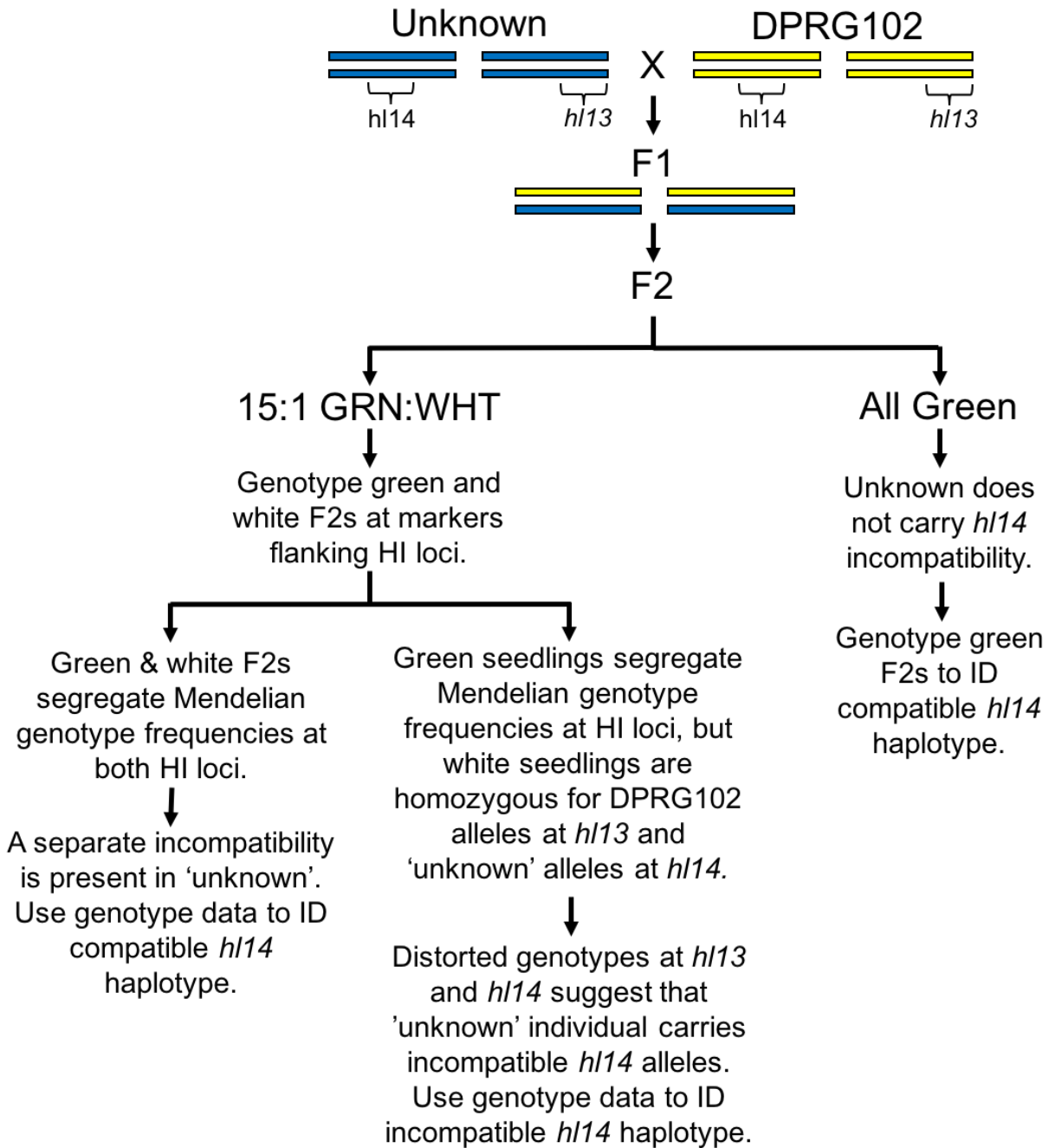


Figure 2.1: Crossing design for characterizing presence/absence of hybrid lethality alleles in unknown individuals. Individuals with unknown genotypes at *hI13* and *hI14* were crossed with 'tester' lines homozygous for these alleles. DPRG102 (*M. guttatus*) is homozygous for incompatible alleles at *hI13* and compatible alleles at *hI14*, whereas

DPRN104 (*M. nasutus*) is homozygous for compatible alleles at *h/13* and incompatible alleles at *h/14*. ‘Tester’ lines always served as paternal parents in our crosses and F1 progeny were self-fertilized to produce a F2 family. A minimum of 87 F2 seedlings, corresponding to a false discovery rate of 1% (probability of not observing white seedlings when they are present), from each cross were counted and the green:white ratio recorded. Eight green F2 seedlings and as many white seedlings as possible were then collected from each F2 family and genotyped at markers flanking *h/13* and *h/14*. White seedlings caused by the *h/13:h/14* incompatibility exhibit highly distorted (i.e., non-Mendelian) genotypes at hybrid lethality loci, whereas green seedlings exhibit Mendelian genotype frequencies. Genotyping F2s at two linked markers also allowed us to identify haplotypes associated with compatible and incompatible hybrid lethality allele in each species. Haplotypes are the unique allele size combinations obtained by genotyping individuals at two markers linked to each hybrid lethality locus. For instance, if an individual carried an incompatible allele at *h/14* and it carried an allele size 283bp at M241 and allele sized 500bp at M132, the individuals haplotype would be ‘*h/14_incompatible_283500*’ (or I_283500, as in table 2.1).

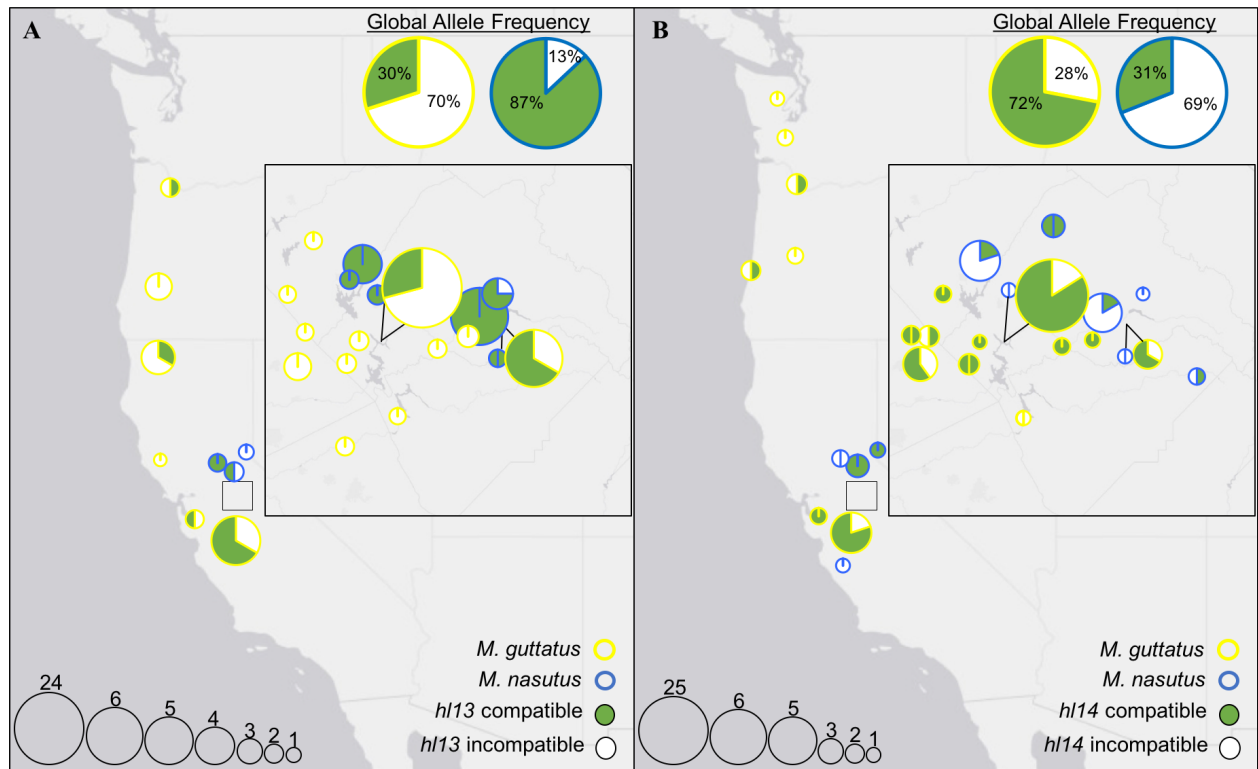


Figure 2.2: Distribution and frequency of *h13* (A) and *h14* (B). Circles represent populations sampled across the western United States. Circle size corresponds to the number of alleles sampled in a given population (bottom left), outer color represent species, inner color corresponds to incompatible or compatible alleles (see key bottom right). Global allele frequencies for *M. guttatus* and *M. nasutus* are shown in upper right corner.

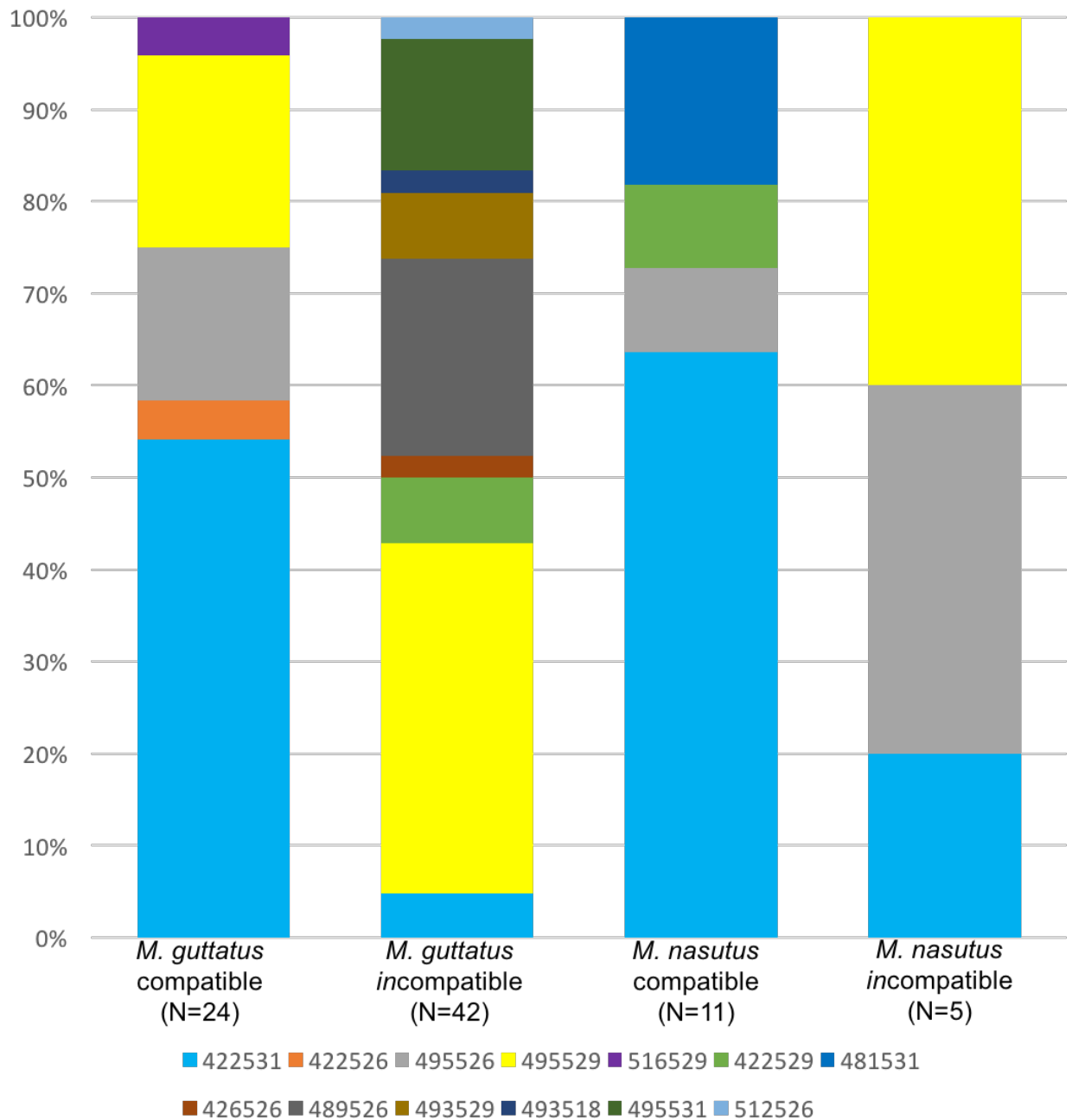


Figure 2.3: Haplotype variation at *h/13*. Bars represent haplotypes associated with compatible and incompatible *h/13* alleles in *M. guttatus* and *M. nasutus*. Bars are stacked to 100% with samples sizes (number of haplotypes sampled) shown below bars. Different colors represent distinct haplotypes generated with markers spaced 9 centimorgans apart and linked to the *h/13* allele

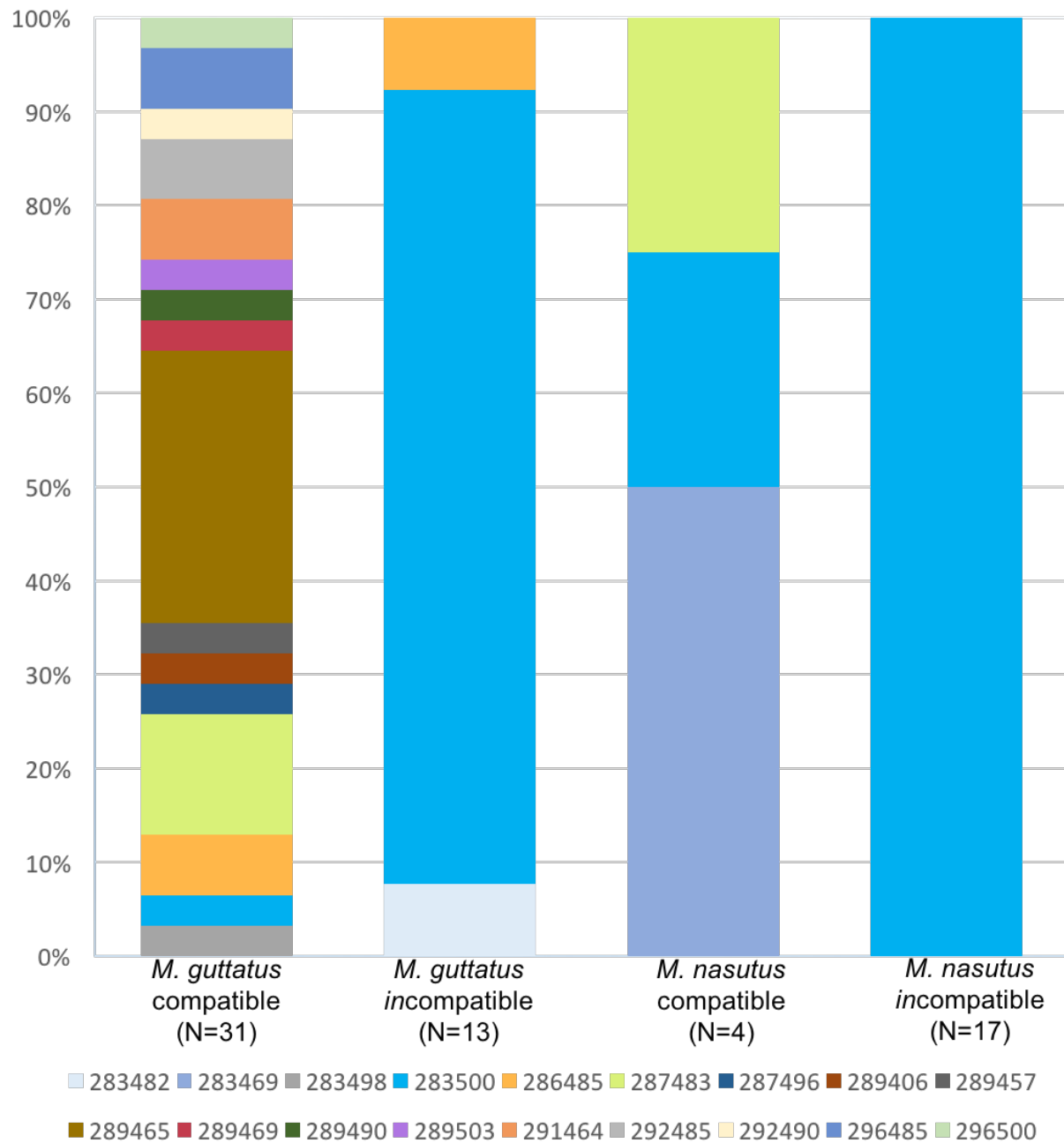


Figure 2.4: Haplotype variation at *h/14*. Bars represent haplotypes associated with compatible and incompatible *h/14* alleles in *M. guttatus* and *M. nasutus*. Bars are stacked to 100% with samples sizes (number of haplotypes sampled) shown below bars. Different colors represent distinct haplotypes generated with markers spaced 11 centimorgans apart and flanking the *h/14* allele.

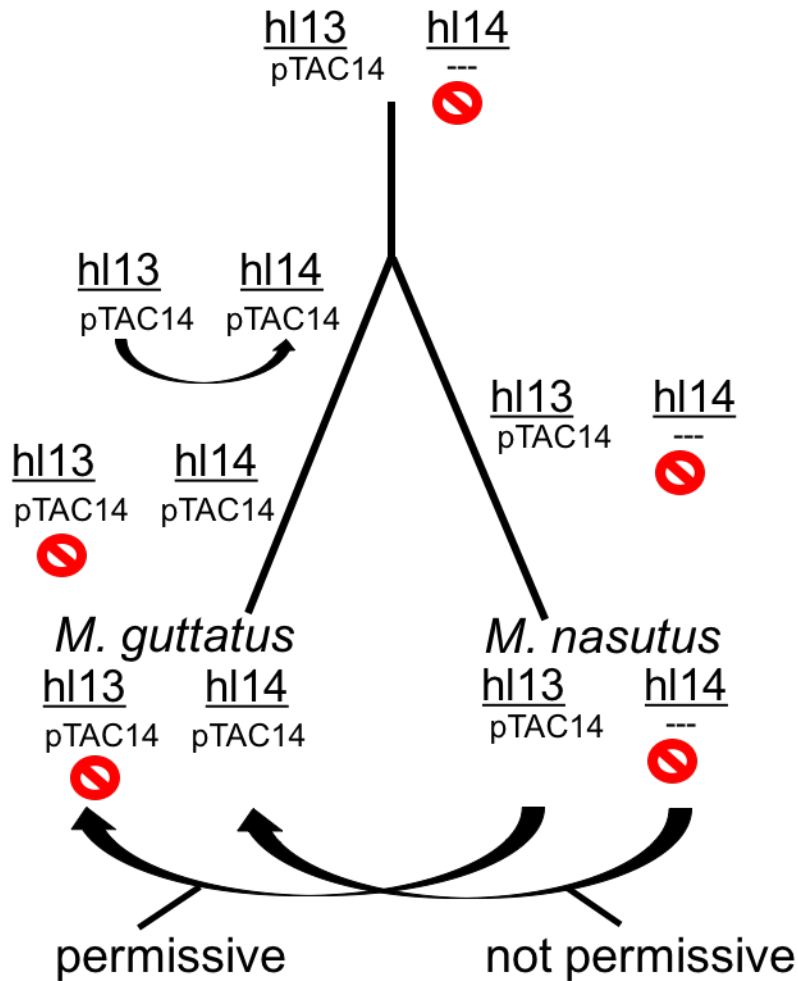


Figure 2.5: Permissiveness of gene flow at hybrid lethality loci. Incompatible alleles are marked with red circle. Gene flow is largely unidirectional from *M. nasutus* into *M. guttatus* in this system. Incompatible *M. nasutus* *hl14* alleles rarely introgress into *M. guttatus* background, because this creates the incompatible genotype. Compatible *M. nasutus* *hl13* alleles create no incompatibility and should therefore easily introgress into *M. guttatus*.

CHAPTER 4

A CRITICAL PHOTOSYNTHESIS GENE UNDERLIES HYBRID LETHALITY IN

***MIMULUS*³**

³Zuellig M.P. and A.L. Sweigart. To be submitted to *Nature*.

Abstract

Hybrid incompatibilities represent a common class of reproductive barriers that limit interspecific gene flow by rendering hybrid progeny unfit. Identification of incompatibility genes can shed light on the evolutionary and genetic mechanisms that underlie their evolution, yielding important insight into the speciation process. Several studies have now identified incompatibility genes, providing important clues about evolutionary forces and molecular mechanisms that cause incompatibilities. However, no incompatibility genes have ever been identified in their most relevant context, between sympatric species that regularly hybridize in nature. In this study, we take a transcriptomics approach to investigate the genetic basis of hybrid lethality between closely related and sympatric species of *Mimulus* that regularly hybridize in nature. We provide strong evidence that duplication of the *Arabidopsis thaliana* homolog of pTAC14 in *M. guttatus*, followed by nonfunctionalization of the ancestral copy, provided the genomic context for hybrid lethality.

Introduction

The production of unfit hybrids is a common mode of reproductive isolation among species (Coyne and Orr 2004) that is largely attributed to the presence of hybrid incompatibilities (HIs) that have evolved between species (Bateson 1909, Dobzhansky 1937, Muller 1942). Genes underlying HIs have now been identified in several systems, including *Drosophila* (Sawamura et al. 1993, Ting et al. 1998, Phadnis and Orr 2009, Tang and Presgraves 2009), *Mus* (Mihola et al. 2009), *Saccharomyces* (Lee et al. 2008,

Chou et al. 2010), *Arabidopsis* (Bikard et al. 2009, Smith et al. 2011), *Caenorhabditis* (Seidel et al. 2008), *Xiphophorus* (Wittbrodt et al. 1989) and *Oryza* (Mizuta et al. 2010, Yamagata et al. 2010). These studies shed light on the evolutionary and genetic mechanisms that give rise to incompatibilities within species and provide important information about their evolutionary histories (Presgraves 2010, Maheshwari and Barbash 2011). However, there remain a glaring lack of studies that investigate the genetic basis of HI between closely related species that experience ongoing gene flow and are therefore more likely to impose reproductive barriers in nature. Furthermore, many HIs are expected to degrade with ongoing gene flow (Gavrilets 1997, Bank et al. 2012), so those that remain intact between hybridizing species may exhibit properties associated with the maintenance of HIs.

Here we investigate the genetic basis of a HI that causes hybrid lethality between *Mimulus guttatus* and *M. nasutus*, two closely related and sympatric species that regularly hybridize in nature (Martin and Willis 2007, Brandvain et al. 2014, Kenney and Sweigart 2016). In a previous study we showed that hybrid lethality is caused by two recessive loci, *hl13* and *hl14*, that prevent chlorophyll production in hybrid plants homozygous for *M. guttatus* alleles at *hl13* and homozygous for *M. nasutus* alleles at *hl14* (Chapter 1). Furthermore, we showed that homologs of the *Arabidopsis thaliana* plastid transcriptionally active chromosome protein, pTAC14, are duplicated in *M. guttatus* and occur at both *hl13* and *hl14* in this species, whereas only a single copy is present in *M. nasutus* at *hl13*. Knockouts of pTAC14 in *A. thaliana* closely resemble hybrid lethality in *Mimulus* and die shortly after germination in the cotyledon stage due

to lack of chlorophyll production (Gao et al. 2011). pTAC14 is a core component of the plastid-encoded prokaryotic-type RNA polymerase (PEP) complex, which is the dominant RNA polymerase involved in transcribing photosynthesis genes in the chloroplast genome (Liere and Borner 2007, Pfannschmidt et al. 2015). Loss of any essential member of the PEP complex results in loss of PEP activity, which results in a characteristic overexpression of housekeeping genes encoded in the chloroplast genome that are transcribed by the nuclear-encoded phage-type RNA polymerase (NEP) and underexpression of genes transcribed by PEP (Pfannschmidt et al. 2015).

In this study, we perform RNAseq on green and white *Mimulus* seedlings and compare patterns of differential expression across various regions of the genome. We first investigate genome-wide patterns of differential expression to characterize the broad effects of hybrid lethality on transcription. We then investigate expression at *h/13* and *h/14* to investigate the hypothesis that homologs of pTAC14 are not expressed in white seedlings. Finally, we investigate patterns of differential expression across the chloroplast genome to determine whether PEP and NEP transcribed genes are misexpressed in a manner consistent with pTAC14 knockouts in *A. thaliana*.

Methods

Plant material and RNA processing

In this experiment, we compared gene expression in white F2 seedlings against seedlings of their parents and green F2 siblings. We generated inbred lines of *M. guttatus* (DPRG102) and *M. nasutus* (DPRN104) through five generations of inbreeding

with single seed descent. These lines were intercrossed to produce reciprocal F1 plants, which were then self-fertilized to produce reciprocal F2 families. We planted parental and F2 seeds on moist Fafard-3B potting soil, cold-stratified seeds at 4C for two weeks, and allowed seedlings to germinate at 23C under 16-hour light cycles in a growth chamber. Seedlings with fully expanded cotyledons began emerging after three days. We collected fully expanded seedlings in pools of ten for each seedling class study (DPRG102, DPRN104, green F2, white F2) directly into 2mL Eppendorf tubes filled with liquid nitrogen. We then extracted RNA from pools of ten seedlings using the Zymo Quick-RNA microprep kit (Zymo Research, Irvine, USA) and estimated RNA concentration using a qubit fluorometer (Life Technologies, Paisley, UK). RNA was subsequently submitted to the Duke University Sequencing Facility, where stranded mRNA libraries were prepared and samples were sequenced across a single lane of Illumina Hiseq 4000 with single-end 50bp reads. Our analysis involved three replicates of DPRG102 and DPRN104 seedlings (i.e., three RNA extractions containing ten seedlings per extraction), five replicates of green F2 seedlings, and six replicates of white F2 seedlings.

Data Processing and Analysis

We performed differential expression analysis using the cufflinks pipeline (Trapnell et al. 2012) with few modifications. We aligned raw reads to the *M. guttatus* IM62 reference genome (www.phytozome.jgi.doe.gov) with Tophat2 (Kim et al. 2013), which resulted in an average of 19 million read alignments per sample. We then

assembled transcriptomes in cufflinks, using the IM62 reference transcriptome to guide assembly. Normalized differential expression analyses were performed using the cuffdiff and cuffnorm programs, where differential gene expression was calculated for all pairwise comparisons between DPRG102, DPRN104, green F2s, and white F2s. We used phytomine (www.phytozome.jgi.doe.gov) to identify *Arabidopsis thaliana* homologs for genes in the *M. guttatus* genome, as top blast hits in the *A. thaliana* genome. We also performed gene ontology enrichment and protein domain enrichment analyses in phytomine using default settings with Bonferroni correction, while general data management, partitioning, and graphing was performed in R (www.R-project.org).

Results

Broad Patterns of Differential Expression: White vs. Green Seedlings

To obtain general information about patterns of differential expression between white seedlings (white F2s) and green seedlings (DPRG102, DPRN104, green F2s), we first isolated genes that were differentially expressed in all white/green comparisons, but were not differentially expressed in green/green comparisons. This analysis identified 1,092 genes that are consistently differentially expressed between white and green seedlings, representing approximately 3.9% of *Mimulus* genes. Of this gene set, 209 genes are underexpressed and 883 are overexpressed in white seedlings (Figure 3.1). Gene ontology analyses show that underexpressed genes are strongly enriched for photosynthetic processes and are localized to the chloroplast, whereas overexpressed

genes exhibit less overall enrichment and do not appear to have general properties that indicate involvement in photosynthesis (Table 3.1).

Differential Expression at Hybrid Lethality Loci

In a previous study, we fine-mapped two loci, *hl13* and *hl14*, that cause hybrid lethality. Here, we investigated patterns of expression in genes within each of these loci to identify potential candidate hybrid lethality genes. Five genes in the *hl13* interval and a single gene at *hl14* are differentially expressed between white seedlings and all green seedlings (Figure 3.2). Two genes, Migut.M02023 and Migut.O00467, are the most underexpressed genes in our comparison of white and green seedlings (see Figure 3.1) and occur at *hl13* and *hl14*, respectively. We previously demonstrated that these genes are duplicated in *M. guttatus*, but exist in single copy at *hl13* in *M. nasutus* (i.e., *M. nasutus* does not carry Migut.O00467). We therefore extracted reads aligning to both genes in DPRG102, DPRN104, and green F2s and reassembled transcripts based on SNP variation within reads in Geneious (Biomatters, www.geneious.com). We found that DPRG102 and DPRN104 contain only a single transcript each, while green F2s contain a combination of these transcripts, indicating that reads in our initial analysis aligned to both duplicates equally well. Furthermore, we found that the DPRG102 transcript corresponds with sequence variation at Migut.O00467, indicating that all transcripts observed in our analysis came from this gene. We provide a corrected analysis of expression at *hl13* and *hl14* (Figure 3.3), which clearly shows that

Migut.M02023 is not expressed in DPRG102 or white seedlings, whereas Migut.O00467 is not expressed in DPRN104 or white seedlings.

Misregulation of chloroplast genome

The genes Migut.M02023 and Migut.O00467 are the only *Mimulus* homologs of the *Arabidopsis thaliana* gene pTAC14, which is a critical PEP-associated protein (PAP) of the plastid-encoded prokaryotic-type RNA polymerase (PEP). Knockouts of pTAC14 (Gao et al. 2011) and other PAPs result in a characteristic chlorotic lethal phenotype in *A. thaliana* (Pfannschmidt et al. 2015), which is very similar to the hybrid lethal phenotype of *Mimulus* hybrids. PEP plays a critical role in transcribing many genes encoded in the chloroplast genome, including genes critical for assembly of chloroplasts, photosystem I, and photosystem II (Liere and Borner 2007). The core PEP RNA polymerase Rpo genes are located in the chloroplast genome and must be transcribed by the nuclear-encoded phage-type RNA polymerase (NEP) to initiate PEP activity in the developing plant (Mullet 1993, Demarsy et al. 2012). PEP and NEP both transcribe genes in the chloroplast genome, where some genes are specifically transcribed by a single polymerase and others are transcribed by both (Demarsy et al. 2006). While the interaction of NEP and PEP is unknown, knockouts of PEP and PAPs exhibit upregulation of NEP-specific chloroplast genes and downregulation of PEP-specific genes (Gao et al. 2011, Schroter et al. 2010).

We hypothesized that if hybrid lethality is due to a lack of pTAC14 transcripts in *Mimulus* white seedlings, we would observe a concomitant increase in regulation of

NEP-specific chloroplast genes and downregulation of PEP-specific chloroplast genes. Across the 54 chloroplast genes assessed here, 70% were differentially regulated in white relative to green seedlings (Figure 3.4), indicating massive misregulation of the chloroplast genome in white seedlings. Furthermore, we assessed differential expression at genes known to be regulated by PEP, NEP, and both polymerases (Figure 3.5). White seedlings exhibit greatly reduced transcription at PEP-regulated genes, increased transcription at NEP-regulated genes, and no pattern at genes regulated by both polymerases.

Discussion

Characterizing the genetic basis of reproductive barriers is a powerful approach to understanding the mechanistic and evolutionary basis of species formation. Hybrid incompatibilities represent a broad class of reproductive barriers that play an important role in speciation by reducing hybrid fitness and reducing gene flow between species. However, little is known about the genes or genetic mechanisms that give rise to HIs. Here, we investigate the genetics of a HI that causes lethality between two closely related species of *Mimulus* and provide compelling evidence that paralogous copies of a critical plastid RNA polymerase accessory protein, pTAC14, are responsible for lethality in this system.

We identified a total of 1,092 genes that were consistently differentially expressed in white seedlings relative to green seedlings. Approximately 81% of these genes were upregulated in white seedlings and enriched for heat shock binding proteins

and glutathione peroxidase activity, which play important roles in cellular homeostasis, repair, and stress response (Bela et al. 2015, Wang et al. 2004), which likely correspond to the stress induced by lack of photosynthetic activity in developing seedlings. The remaining 19% of genes were downregulated and strongly enriched for photosynthetic genes, indicating that a basal regulator of genes involved in photosynthesis or chloroplast development may be disrupted in these plants.

In a previous study, we fine-mapped two loci, *hl13* and *hl14*, that cause hybrid lethality in *Mimulus*. We investigated gene expression within these intervals, finding that white seedlings exhibit differential expression at five genes within *hl13*, but only a single gene within *hl14*. Interestingly, the two most downregulated genes in our analysis of differential expression between white and green seedlings (Figure 3.1) are Migut.M02023 and Migut.O00467, which occur within *hl13* and *hl14*, respectively (Chapter 1). These genes represent paralogous copies of the *Arabidopsis thaliana* gene pTAC14, which is a critical accessory protein of PEP RNA polymerase (Gao et al. 2011). Migut.M02023 and Migut.O00467 are both present in *M. guttatus*, but only a single copy (Migut.M02023) is present in *M. nasutus* (Chapter 1). However, *M. guttatus* (DPRG102) only produces a single transcript that corresponds to Migut.O00467 (Figure 3.4). Furthermore, white seedlings, which are homozygous for *M. guttatus* alleles at *hl13* and homozygous for *M. nasutus* alleles at *hl14*, produce no transcript for either gene. Thus, white seedlings completely lack expression of the *Mimulus* homolog of pTAC14.

Knockouts of pTAC14, PEP RNA polymerase, and other essential PEP RNA polymerase accessory proteins exhibit several characteristic phenotypes, including albino seedlings that die at the cotyledon stage, downregulation of PEP regulated chloroplast genes, and upregulation of NEP regulated chloroplast genes (Gao et al. 2011). While many of the genes encoded in the chloroplast genome can be regulated by both NEP and PEP, the transcriptional activity of these polymerases appears to exhibit a large degree of partitioning, particularly early in development (Schroter et al. 2010). In general, NEP is believed to play an important role in regulating ‘housekeeping’ genes early in seedling development, such as those involved in ATP synthesis, ribosome production, and the core PEP RNA polymerase *rpo* genes (Mullet 1993, Demarsy et al. 2012). PEP appears to take over a primary role in transcribing chloroplast genes, including photosystem I, photosystem II, and cytochrome b6f, as greening and chloroplast development proceeds and continues to maintain a dominant transcriptional role throughout the lifetime of the plant. In general, expression patterns of chloroplast genes in white seedlings are consistent with our expectations for pTAC14 knockouts. Furthermore, we identified a subset of genes that is specifically regulated by PEP, NEP, and both polymerases and compared transcription among green seedlings and white seedlings for these genes. Our data clearly demonstrate that PEP-regulated genes are downregulated, NEP-regulated genes are upregulated, and genes controlled by transcribed both polymerases show no clear expression patterning (Figure 3.5).

Our data strongly indicate that the *Mimulus* homolog of pTAC14 is responsible for hybrid lethality in *Mimulus*. This finding has important implications for the evolution

and maintenance of hybrid incompatibilities. Inactivation of duplicate genes is largely seen as a mutational process, whereby relaxed selection on redundant genes allows for the accumulation of deleterious mutations (Werth and Windham 1993, Lynch and Force 2000). It is difficult to imagine positive selection acting on the incompatible copy of pTAC14 (Migut.M02023) at *hl13* in the *M. guttatus* or the non-existent copy at *hl14* in *M. nasutus*. As such, the evolution of this HI is likely due to largely neutral evolutionary processes. Other studies have observed HIs that have evolved in a similar fashion (Bikard et al 2009, Mizuta et al. 2010, Yamagata et al. 2010), indicating that gene duplication may be a common factor in the evolution of HIs. However, theory holds that neutral HIs should be purged from systems where gene flow is ongoing (Gavrilets 1997, Bank et al. 2012), begging the question of how incompatibilities are maintained in systems where low levels of hybridization are common. Our system is unique in this sense, as other systems where gene duplication underlying HIs has been uncovered experience extremely low levels of interspecific gene flow and may therefore be maintained through allopatry alone. Interestingly, we observe that hybrid lethality alleles are polymorphic, widespread, and possibly maintained in equilibrium in both *M. guttatus* and *M. nasutus* (Chapter 2), begging the question of how a neutrally evolving incompatibility can be maintained within this system? With the genes in hand, this and other questions concerning the evolution and maintenance of HI alleles can be directly tested in a naturally occurring system where HIs may actually contribute to reproductive isolation.

References:

- Bank, C., R. Bürger, and J. Hermisson. 2012. The Limits to Parapatric Speciation: Dobzhansky–Muller Incompatibilities in a Continent–Island Model. *Genetics* 191:845-863.
- Bateson, W. 1909. Heredity and variation in modern lights. *Darwin and modern science* 85:101.
- Bela, K., E. Horváth, Á. Gallé, L. Szabados, I. Tari, and J. Csiszár. 2015. Plant glutathione peroxidases: emerging role of the antioxidant enzymes in plant development and stress responses. *Journal of plant physiology* 176:192-201.
- Bikard, D., D. Patel, C. Le Metté, V. Giorgi, C. Camilleri, M. J. Bennett, and O. Loudet. 2009. Divergent Evolution of Duplicate Genes Leads to Genetic Incompatibilities Within *A. thaliana*. *Science* 323:623-626.
- Brandvain, Y., A. M. Kenney, L. Flagel, G. Coop, and A. L. Sweigart. 2014. Speciation and introgression between *Mimulus nasutus* and *Mimulus guttatus*. *PLoS Genet* 10:e1004410.
- Chou, J.-Y., Y.-S. Hung, K.-H. Lin, H.-Y. Lee, and J.-Y. Leu. 2010. Multiple molecular mechanisms cause reproductive isolation between three yeast species. *PLoS Biol* 8:e1000432.
- Coyne, J. A. and H. A. Orr. 2004. *Speciation*. Sunderland, MA. Sinauer Associates, Inc.
- Demarsy, E., F. Buhr, E. Lambert, and S. Lerbs-Mache. 2012. Characterization of the plastid-specific germination and seedling establishment transcriptional programme. *Journal of experimental botany* 63:925-939.
- Demarsy, E., F. Courtois, J. Azevedo, L. Buhot, and S. Lerbs-Mache. 2006. Building up of the plastid transcriptional machinery during germination and early plant development. *Plant physiology* 142:993-1003.
- Dobzhansky, T. and T. G. Dobzhansky. 1937. *Genetics and the Origin of Species*. Columbia University Press.
- Gao, Z.-P., Q.-B. Yu, T.-T. Zhao, Q. Ma, G.-X. Chen, and Z.-N. Yang. 2011. A functional component of the transcriptionally active chromosome complex, *Arabidopsis* pTAC14, interacts with pTAC12/HEMERA and regulates plastid gene expression. *Plant physiology* 157:1733-1745.
- Gavrilets, S. 1997. Hybrid Zones With Dobzhansky-Type Epistatic Selection. *Evolution* 51:1027-1035.
- Kenney, A. M. and A. L. Sweigart. 2016. Reproductive isolation and introgression between sympatric *Mimulus* species. *Molecular ecology* 25:2499-2517.
- Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, and S. L. Salzberg. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology* 14:R36.
- Lee, H.-Y., J.-Y. Chou, L. Cheong, N.-H. Chang, S.-Y. Yang, and J.-Y. Leu. 2008. Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species. *Cell* 135:1065-1073.
- Liere, K. and T. Börner. 2007. Transcription and transcriptional regulation in plastids. Pp. 121-174. *Cell and molecular biology of plastids*. Springer.

- Lynch, M. and A. G. Force. 2000. The Origin of Interspecific Genomic Incompatibility via Gene Duplication. *The American Naturalist* 156:590-605.
- Maheshwari, S. and D. A. Barbash. 2011. The Genetics of Hybrid Incompatibilities. *Annual Review of Genetics* 45:331-355.
- Martin, N. H. and J. H. Willis. 2007. Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric *Mimulus* species. *Evolution* 61:68-82.
- Mihola, O., Z. Trachtulec, C. Vlcek, J. C. Schimenti, and J. Forejt. 2009. A mouse speciation gene encodes a meiotic histone H3 methyltransferase. *Science* 323:373-375.
- Mizuta, Y., Y. Harushima, and N. Kurata. 2010. Rice pollen hybrid incompatibility caused by reciprocal gene loss of duplicated genes. *Proceedings of the National Academy of Sciences* 107:20417-20422.
- Muller, H. J. 1942. Isolating mechanisms, evolution and temperature. Pp. 71-125. *Biol. Symp.*
- Mullet, J. E. 1993. Dynamic regulation of chloroplast transcription. *Plant Physiology* 103:309.
- Pfannschmidt, T., R. Blanvillain, L. Merendino, F. Courtois, F. Chevalier, M. Liebers, B. Grüber, E. Hommel, and S. Lerbs-Mache. 2015. Plastid RNA polymerases: orchestration of enzymes with different evolutionary origins controls chloroplast biogenesis during the plant life cycle. *Journal of experimental botany* 66:6957-6973.
- Phadnis, N. and H. A. Orr. 2009. A single gene causes both male sterility and segregation distortion in *Drosophila* hybrids. *science* 323:376-379.
- Presgraves, D. C. 2010. The molecular evolutionary basis of species formation. *Nature Reviews Genetics* 11:175-180.
- Sawamura, K., M. Yamamoto, and T. K. Watanabe. 1993. Hybrid lethal systems in the *Drosophila melanogaster* species complex. II. The Zygotic hybrid rescue (Zhr) gene of *D. melanogaster*. *Genetics* 133:307-313.
- Sawamura, K. and M.-T. Yamamoto. 1997. Characterization of a reproductive isolation gene, zygotic hybrid rescue, of *Drosophila melanogaster* by using minichromosomes. *Heredity* 79.
- Schröter, Y., S. Steiner, K. Matthäi, and T. Pfannschmidt. 2010. Analysis of oligomeric protein complexes in the chloroplast sub-proteome of nucleic acid-binding proteins from mustard reveals potential redox regulators of plastid gene expression. *Proteomics* 10:2191-2204.
- Seidel, H. S., M. V. Rockman, and L. Kruglyak. 2008. Widespread genetic incompatibility in *C. elegans* maintained by balancing selection. *science* 319:589-594.
- Smith, L. M., K. Bomblies, and D. Weigel. 2011. Complex evolutionary events at a tandem cluster of *Arabidopsis thaliana* genes resulting in a single-locus genetic incompatibility. *PLoS Genet* 7:e1002164.
- Tang, S. and D. C. Presgraves. 2009. Evolution of the *Drosophila* nuclear pore complex results in multiple hybrid incompatibilities. *Science* 323:779-782.
- Ting, C.-T., S.-C. Tsaur, M.-L. Wu, and C.-I. Wu. 1998. A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science* 282:1501-1504.
- Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim, D. R. Kelley, H. Pimentel, S. L. Salzberg, J. L. Rinn, and L. Pachter. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protocols* 7:562-578.

- Wang, W., B. Vinocur, O. Shoseyov, and A. Altman. 2004. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends in plant science* 9:244-252.
- Werth, C. R. and M. D. Windham. 1991. A Model for Divergent, Allopatric Speciation of Polyploid Pteridophytes Resulting from Silencing of Duplicate-Gene Expression. *The American Naturalist* 137:515-526.
- Wittbrodt, J., D. Adam, B. Malitschek, W. Mäueler, F. Raulf, A. Telling, S. M. Robertson, and M. Scharl. 1989. Novel putative receptor tyrosine kinase encoded by the melanoma-inducing Tu locus in *Xiphophorus*. *Nature* 341:415-421.
- Yamagata, Y., E. Yamamoto, K. Aya, K. T. Win, K. Doi, Sobrizal, T. Ito, H. Kanamori, J. Wu, T. Matsumoto, M. Matsuoka, M. Ashikari, and A. Yoshimura. 2010. Mitochondrial gene in the nuclear genome induces reproductive barrier in rice. *Proceedings of the National Academy of Sciences* 107:1494-1499.

Table 3.1: Gene ontology enrichment of differentially expressed genes.

Ontology	GO Term, Name	p-value
<u>Underexpressed</u>		
Biological Process:	GO:0019684, photosynthesis light reaction	4.9e-11
	GO:0006091, generation of precursor metabolites and energy	1.8e-10
	GO:0015979, photosynthesis	3.5e-10
	GO:0009765, photosynthesis light harvesting	6.2e-6
	GO:0009072, aromatic amino acid family metabolic process	9.1e-5
	GO:0009767, photosynthetic electron transport chain	3.0e-4
	GO:0009073, aromatic amino acid family biosynthetic process	4.0e-4
	GO:0022900, electron transport chain	0.0010
	GO:0008652, cellular amino acid biosynthetic process	0.0090
	GO:0009579, thylakoid	0.0019
Cellular Compartment:	GO:0044436, thylakoid part	0.0019
	GO:0034357, photosynthetic membrane	0.0097
Molecular Function:	GO:0045156, transferring electrons within the cyclic electron transport pathway of photosynthesis activity	0.0139
	GO:0046982, protein heterodimerization activity	0.0348
	GO:0016765, transferase activity	0.0432
<u>Overexpressed</u>		
Cellular Compartment:	GO:0005622, intracellular	0.0099
Molecular Function:	GO:0031072, heat shock protein binding	0.0370
	GO:0004602, glutathione peroxidase activity	0.0413

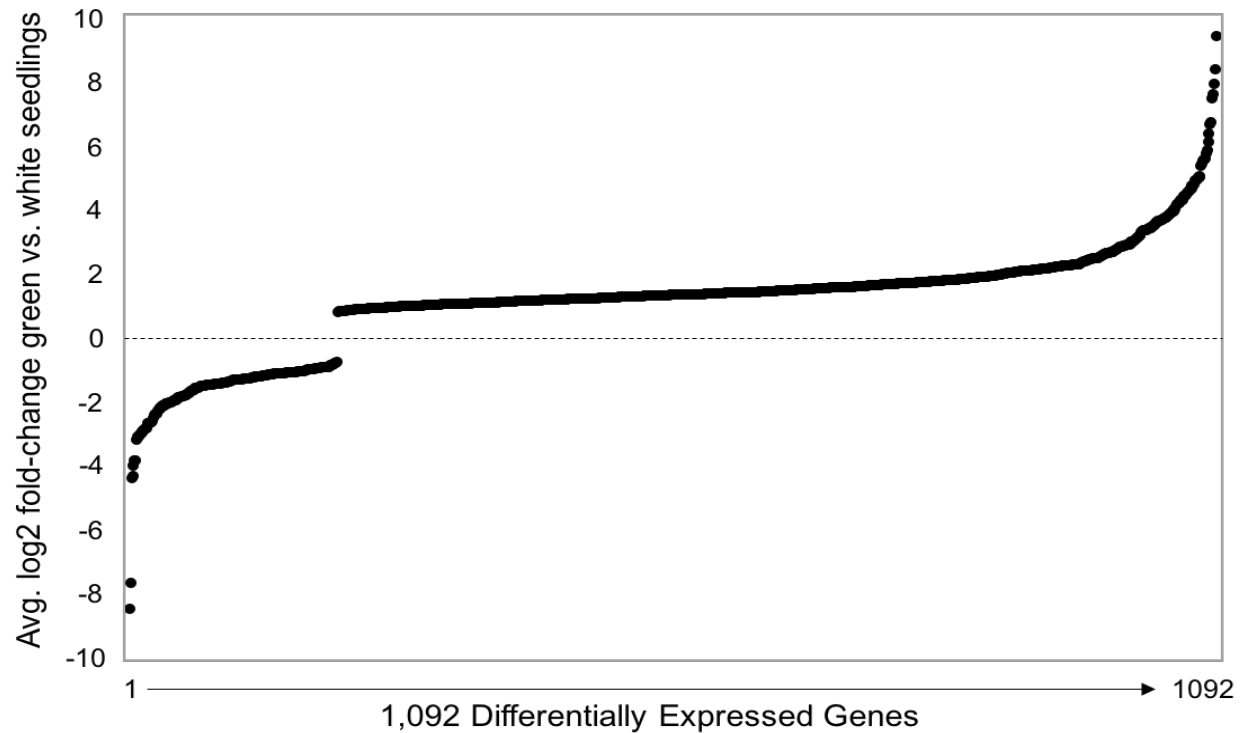


Figure 3.1: Average log2 fold-change among 1,092 genes that are differentially expressed in all pairwise comparisons of white (white F2) and green (DPRG102, DPRN104, green F2) seedlings. Genes above and below line are over- and underexpressed in white seedlings, respectively.

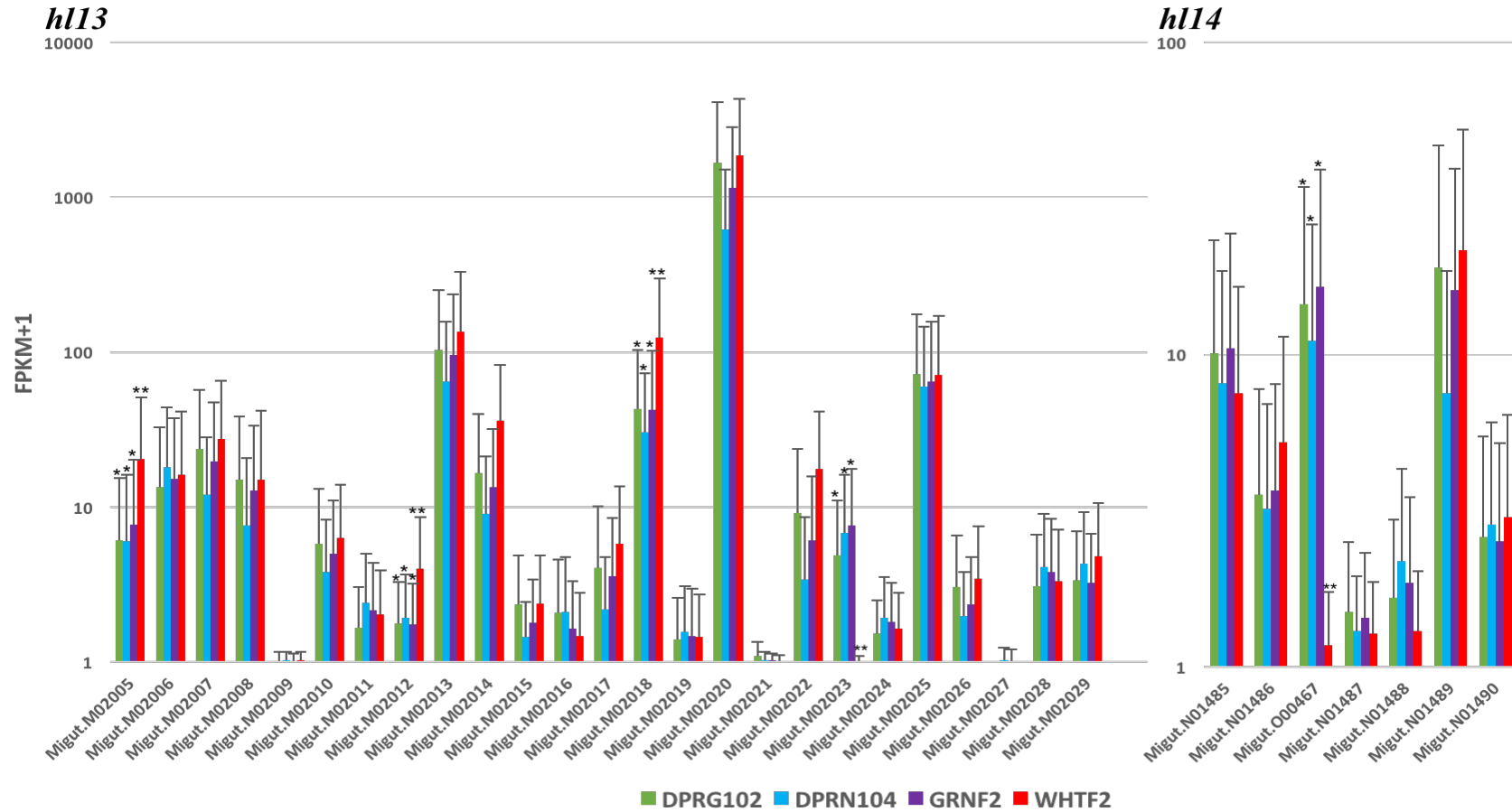


Figure 3.2: Expression of genes within *h13* and *h14* hybrid lethality loci, uncorrected. Genes differentially expressed between white F2s (WHTF2) and all green seedlings are indicated by asterisks. Migut.M02023 and Migut.O00467 located at *h13* and *h14*, respectively, are homologs of *A. thaliana* pTAC14.

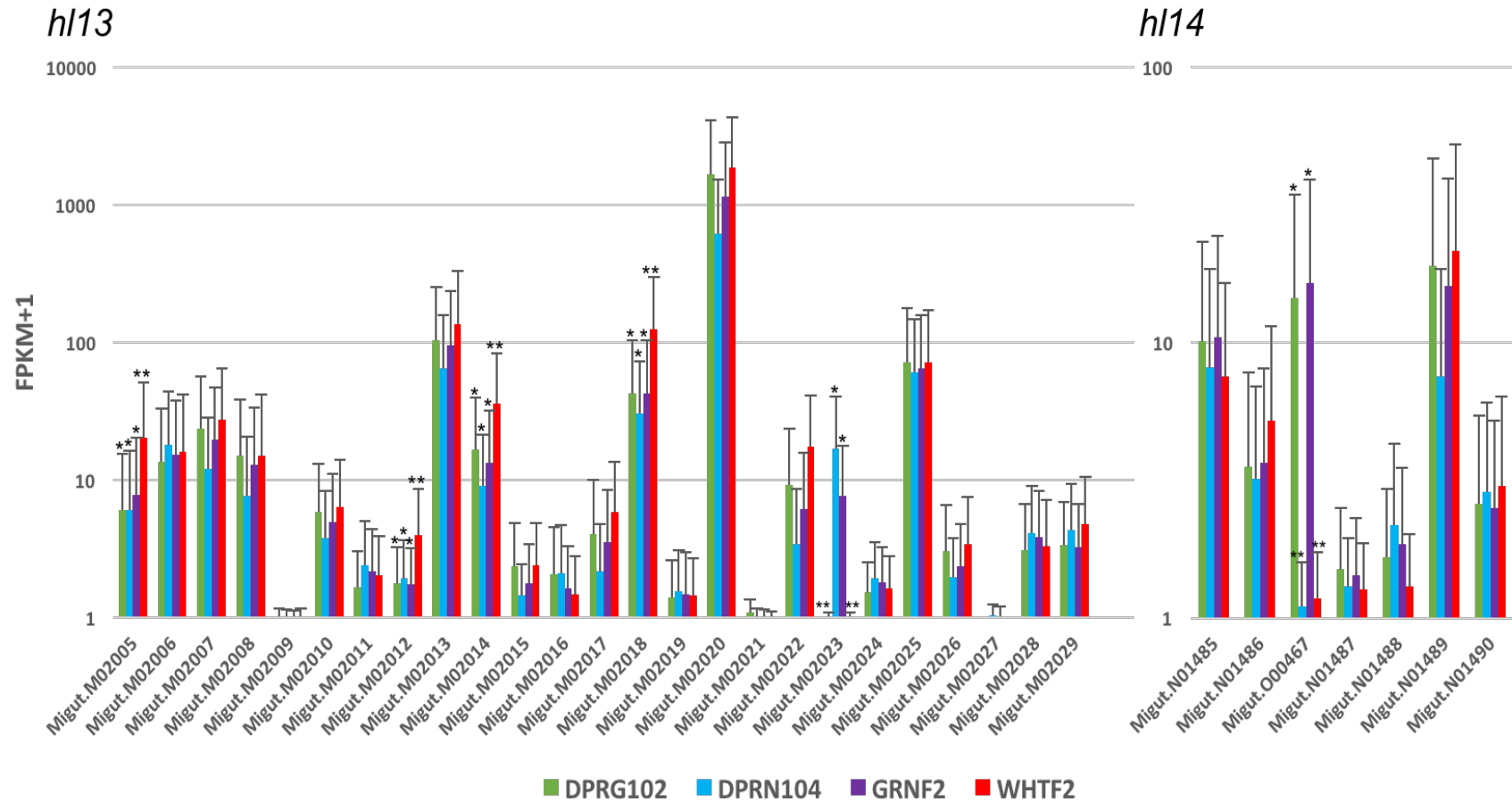


Figure 3.3: Expression of genes within *h13* and *h14* hybrid lethality loci, corrected. Genes differentially expressed between white F2s (WHTF2) and all green seedlings are indicated by asterisks. Migut.M02023 and Migut.O00467 located at *h13* and *h14*, respectively, are homologs of *A. thaliana* pTAC14. Here, transcripts produced from DPRG102 and DPRN104 are aligned to the correct pTAC14 paralog annotated in the *M. guttatus* reference genome.

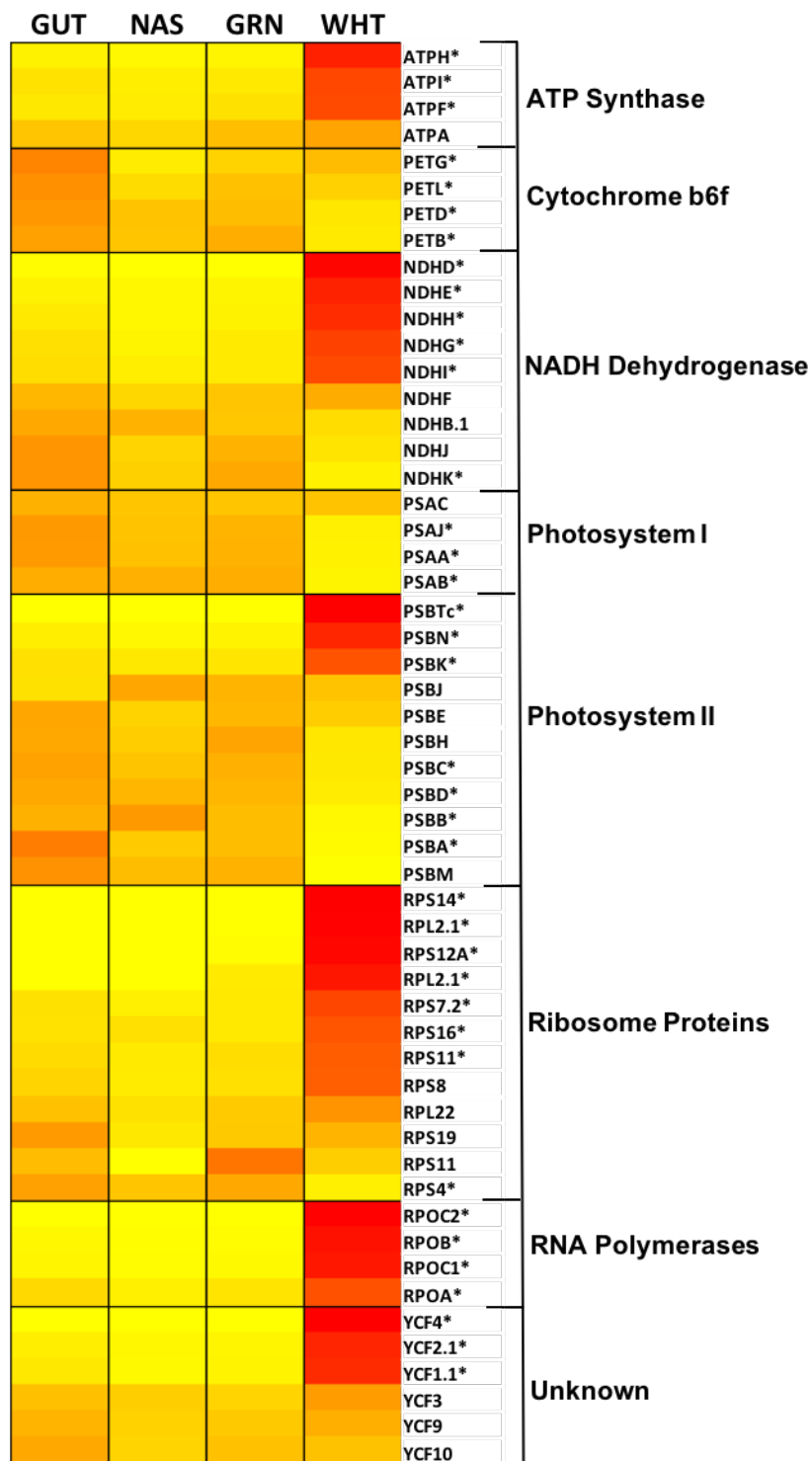


Figure 3.4: Differential expression of chloroplast encoded genes. GUT (DPRG102), NAS (DPRN104), GRN (green F2), WHT (white F2).

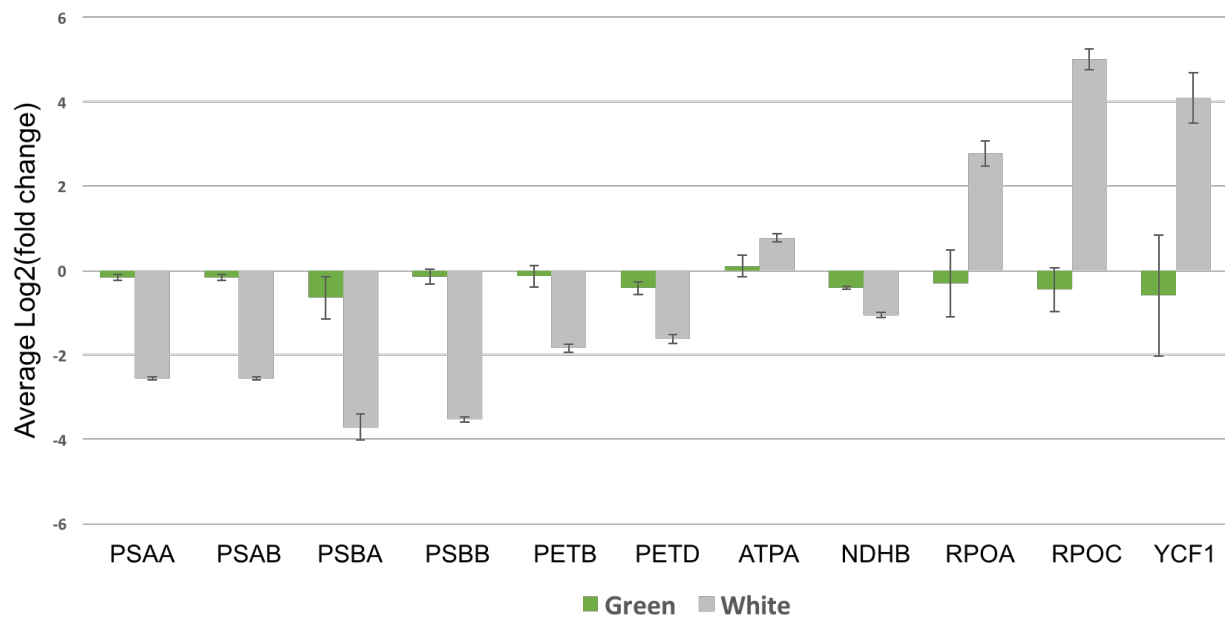


Figure 3.5: Differential expression among chloroplast genes with known PEP and NEP transcriptional activity. Green: average fold change in pairwise comparisons among green seedlings. White: average fold change in pairwise comparisons between white and all green seedlings. PEP transcribed genes: PSAA, PSAB, PSBA, PSBB, PETB, and PETD. NEP and PEP transcribed genes: ATPA and NDHB. NEP transcribed genes: RPOA, RPOC, and YCF1.

CHAPTER 5

CONCLUSION

In this dissertation, I have provided a thorough investigation of a hybrid incompatibility that causes lethality between closely related species of *Mimulus* that naturally hybridize in nature. This work represents one of the first investigations of a hybrid incompatibility that occurs between naturally hybridizing species and is novel in that I provide a detailed explanation of the genes underlying lethality and shed light on the dynamics of hybrid incompatibility in a natural population.

In chapters two and four, I utilized a series of genetic mapping and transcriptomic approaches to identify the genetic basis of hybrid lethality. I demonstrate that hybrid lethality is caused by two unlinked recessive nuclear loci (*hl13* and *hl14*), which both harbor copies of the critical photosynthetic gene pTAC14. pTAC14 is not expressed in hybrid lethal plants and the transcriptome of these plants resembles those of pTAC14 knockouts in *Arabidopsis thaliana*. Furthermore, functional copies of pTAC14 occur at different genomic locations in *M. guttatus* (*hl14*) and *M. nasutus* (*hl13*). In the F2 generation, 1/16th of the hybrid progeny are homozygous for *M. guttatus* alleles at *hl13* and *M. nasutus* alleles at *hl14* and therefore do not harbor a functional copy of this critical gene, causing premature lethality. This work provides important information

about the genes that underlie hybrid incompatibilities and highlights the importance of gene duplications in the evolution of reproductive barriers.

In chapter three, I took a combined crossing and genotyping approach to investigate the distribution, frequency, and dynamics of hybrid lethality (*h13* and *h14*) alleles in natural populations of *Mimulus*. I show that *h13* and *h14* are widespread and polymorphic within *M. guttatus* and *M. nasutus*, implying that incompatibility alleles are not being rapidly purged from this system and may actually be maintained at equilibrium. Furthermore, I show that introgression of incompatible *h14* alleles from *M. nasutus* into *M. guttatus* appears to be restricted, which is likely the result of selection against these alleles when hybrid lethals are formed. Interestingly, I also provide evidence that introgression of compatible *h13* alleles from *M. nasutus* into *M. guttatus* is restricted, which implies that incompatible *h13* alleles may be favored in *M. guttatus*. These observations imply that hybrid lethality may be maintained at equilibrium within *Mimulus* and likely contributes to reproductive isolation within this system.