INVESTIGATING PRE-MATING AND POST-MATING REPRODUCTIVE ISOLATION IN *DROSOPHILA*

Charles J. J. Miller

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Curriculum of Genetics and Molecular Biology in the School of Medicine.

Chapel Hill
2016

Approved by:

Daniel Matute

Jeff Sekelsky

Corbin Jones

Bob Goldstein

Chris Martin

ABSTRACT

Charles J. J. Miller: Investigating pre-mating and post-mating reproductive isolation in *Drosophila*
(Under the direction of Daniel R. Matute)

Reproductively isolating barriers which inhibit gene flow between species can be broadly classified into pre-mating and post-mating barriers. Pre-mating barriers evolve rapidly and are thought to be important for the initiation of speciation, while post-mating barriers evolve more slowly and are thought to be important for the maintenance of separate species after speciation has occurred. Here I present two studies, each examining one type of reproductive barrier in *Drosophila* species. Chapter 2 examines hybrids of *Drosophila melanogaster* with two other species, *Drosophila simulans* (5 million years diverged) and *Drosophila santomea* (15 million years diverged), and examines the effects of temperature in post-mating reproductive isolation in these crosses. Post-mating isolation in hybrids with *D. melanogaster* is temperature sensitive in crosses with *D. simulans*, but not *D. santomea*. These data suggest that divergence time and extrinsic factors both play a significant role in post-mating isolation. Chapter 3 examines the cosmopolitan/Zimbabwe mate choice split in *D. melanogaster*, in which Zimbabwe females only choose to mate with Zimbabwe males. The trait maps strongly to the *rim* locus on chromosome 3R, and Zimbabwe and cosmopolitan lines differ by two SNPs in *rim*. Precise gene replacement with CRISPR/Cas9 reveals that the Zimbabwe allele of *rim* induces a strong Zimbabwe male preference in cosmopolitan flies, suggesting that strong pre-mating isolation can occur as a result of one or two SNP changes.

This work is dedicated to my wife, Cortney Winkle-Miller, and to my parents and my sister. I do not believe a more supportive group of human beings exists on this Earth.

ACKNOWLEDGEMENTS

The completion of this work has been a significant process for me, and not always an easy one. I chose to switch labs (and projects, and fields) in my fourth year of graduate school; while challenging, this has ultimately proved to be one of the best decisions I made while at UNC. I feel blessed to have been able to work with Daniel and my colleagues in the Matute lab, who welcomed me with open arms and helped me to thrive.

To my committee, thank you for your support and guidance. Dr. Corbin Jones, thank you for answering my endless questions about fly genetics and genomics. Dr. Chris Martin, your career guidance has been invaluable. Dr. Bob Goldstein, your insightful questions forced me to never perform at less than my best. Dr. Jeff Sekelsky, your feedback was always profoundly helpful, and you are a better Director of Graduate Studies than any program could ever hope for.

Daniel, you have been a truly extraordinary mentor. You took a chance on a senior graduate student switching thesis labs, who had never worked in evolutionary biology before, and you have taught me so much the last two years. Words cannot express how fortunate I feel to have been given the chance to work with and learn from you. Thank you, thank you.

Aaron, Dave, and Antonio. You guys are great coworkers, equally capable of joking around and shooting the breeze and having excellent conversations about science and research. I'm going to miss working with you.

Lastly, but most importantly, I simply could not have done this without my wife, Cortney. A better partner could not be had. I could not have done this without you by my side. I love you.

PREFACE

Chapter 2 is a manuscript accepted in the journal Genes|Genomes|Genetics. I collected the data, performed the data analysis, and drafted the figures with input from Dr. Daniel Matute. Additional data from Matute et al. 2010, published in Science, was used. Chapter 3 is a manuscript currently in preparation. I performed the experiments and drafted the figures with input from Dr. Daniel Matute. Dr. David Turissini and Dr. Aaron Comeault provided assistance with data analysis and mating experiments, as well as sequence analysis for construct design for the CRISPR/Cas9 modifications performed. Jamie Roebuck performed the microinjections on *Drosophila* embryos to generate the transformant flies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

Bal	Balancer chromosome
BSC	Biological Species Concept
CHC	Cuticular Hydrocarbons
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
D.	*Drosophila*
df	Deficiency chromosome
DMI	Dobzhansky-Muller Incompatibility
F1	1st generation offspring
GSC	Genic Species Concept
HI	Hybrid Incompatibility
kbp	Thousand Base Pairs
M	Cosmopolitan *Drosophila melanogaster*
mel	*Drosophila melanogaster*
san	*Drosophila santomea*
sim	*Drosophila simulans*
Z	Zimbabwe *Drosophila melanogaster*

CHAPTER ONE : INTRODUCTION

What Are Species?

Perhaps the longest running argument in the field of biological science is what, exactly, are species. One of the first formal systems of describing species, Linnaean taxonomy, grouped organisms according to like-characteristics and set forth a naming and classification structure to organize the catalogue of known organisms (Vences et al., 2013). While only the most broad categories of Linnaean taxonomy are still in use today, it has had a profound effect on the way we conceptualize and classify species.

Initially, species were simply classified according to taxonomic characteristics: likeness of appearance, similarity in biological structure, etc. However, this mode of classification eventually fell out of popularity due to its fallibility (one striking example being the red panda *Ailurus fulgens*, initially classified as a bear, and then a type of raccoon, and now considered to be a member of the mustelid superfamily) (Flynn et al., 2000). *Speciation* by Coyne and Orr lists no fewer than 9 potential definitions of species (Coyne and Orr, 2004). The most widely used is the Biological Species Concept, which states that two organisms are considered to be separate species if they cannot produce fertile offspring. This mechanism of classification has had useful impact, allowing us to correctly classify wolves and dogs as being subspecies of the species *Canis lupus* rather than entirely separate groups, but it has also faltered in some respects. For example, some hybrid offspring of lions and tigers ("ligers" or "tigons," depending on the gender of each parent) have been found to be

fertile (Guggisberg, 1975), but lions and tigers are too obviously different to be classified as one species.

Recently, the most widely used method for classifying and grouping species has involved the use of genetic and molecular data to group and classify species based on sequence similarity, frequently referred to as the Genic Species Concept. The GSC uses the wealth of DNA sequencing information currently available to group organisms by similarity in nucleotide sequence and molecular markers as well as by morphological and reproductive traits. More accurate than relying simply on physical characteristics or mating outcome, this has allowed us to expand our conceptualization of species. The biggest weakness of the Genic Species Concept is in identifying nascent species or incipient speciation, in which genetic divergence is not so high as to warrant classification as separate species but divergence is clearly occurring. One such example is in the Zimbabwe and cosmopolitan races of *Drosophila melanogaster* (David and Capy, 1988), which will be discussed in depth later. Thus, a synthesis of the BSC and GSC remains our best method for identifying and classifying species. Modern taxonomies and phylogenies frequently employ a combination of both methods in order to achieve the most accurate classification of species possible (Baker and Bradley, 2006).

Speciation

The process through which one species divides into two (or more) is termed speciation. Put simplistically, speciation is caused by the cessation of gene flow (exchange of genetic material) between two populations of a species. Without gene flow, newly arisen alleles or changes in allele frequency in one population are not reflected in the second. Over time, the accumulation of these differences leads to complete divergence of the species. There are four main modes of speciation, differentiated based on the level of contact and

potential gene flow between the two populations, which will be discussed in detail below: allopatric, peripatric, parapatric, and sympatric (Fig 1.1).

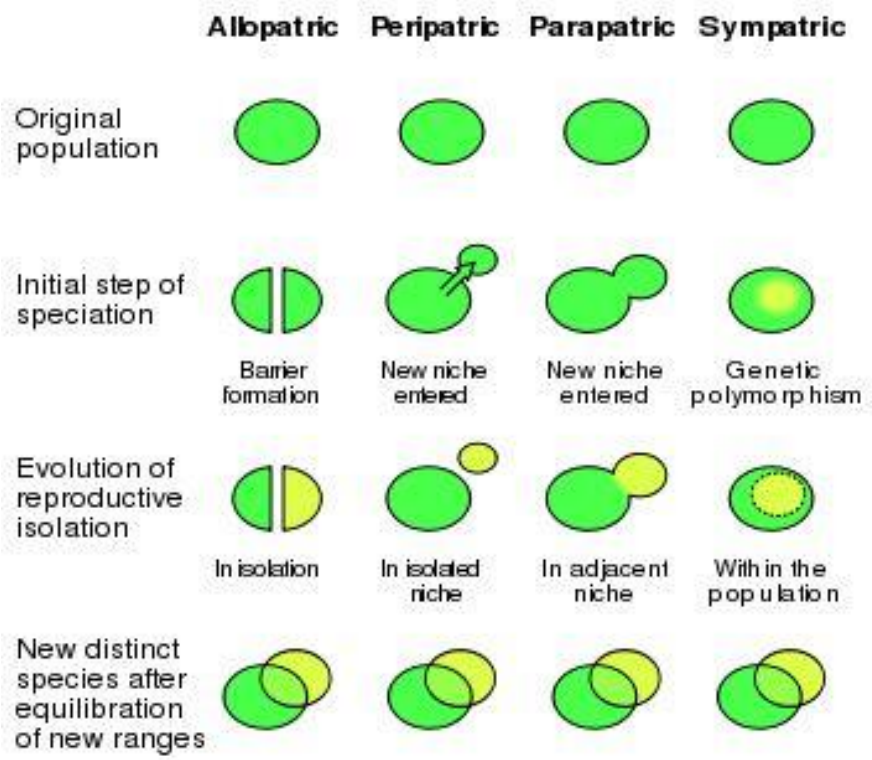

Figure 1.1 Modes of speciation

An illustration of the four modes of speciation: allopatric, peripatric, parapatric, and sympatric. Each involves a different level of geographic isolation between the two populations of the ancestral species, with allopatric being the most complete isolation and sympatric being no isolation at all. (Krempels, 2006)

Allopatric speciation is the "classical" mode of speciation (Butlin et al., 2008), and the easiest to conceptualize. In an allopatric speciation model, two populations of a species are separated by a geographic barrier that entirely precludes gene flow (a common teaching example being a population which finds itself stranded on an island, or a volcanic eruption creating novel physical barriers between formerly connected regions). With gene flow between the populations a physical impossibility, genetic drift and specialization into novel niches inexorably result in the differentiation of the two groups into separate species. This is

likely the main mode of speciation (Butlin et al., 2008, Coyne and Orr, 2004), as other modes require additional barriers to gene flow beyond the physical. Peripatric speciation is a subset of allopatric speciation, and occurs when one population of a species colonizes a new niche and becomes geographically isolated as a result (Provine, 2004). The principle difference between allopatric and peripatric speciation is that in peripatric speciation one of the populations is much smaller, which may have eventual consequences for that population (such as founder or bottleneck effects) (Coyne and Orr, 2004).

Parapatric speciation is a mode of speciation in which two populations are mostly separated but may overlap in range in a narrow contact zone between the two ranges. This unequal distribution may arise through novel geographic barriers which hinder but do not entirely preclude travel, unequal dispersal into a newly colonized area resulting in unusual ranges, or divergence in behavior (Butlin et al., 2008, Coyne and Orr, 2004). The small size of the contact zone results in unequal and/or non-random gene flow, which may not overcome divergence in the populations through genetic drift, niche specialization, or behavioral divergence leading to fixation of novel alleles. Eventually, the reduction in gene flow may lead to the differentiation of the two populations into two species. Notably, if the species maintain their initial ranges after speciation, they may hybridize in the contact zone and maintain limited gene flow through introgression (if F1 hybrids are able to produce progeny with the parental species). This may have eventual fitness consequences for both populations through adaptive introgression (Llopart et al., 2014) or reinforcement (Matute, 2010a).

Sympatric speciation is the most contentious of the evolutionary modes, with some groups suggesting it simply does not occur (Fitzpatrick et al., 2008) while others posit it may

be quite common (Johannesson, 2010). In sympatric speciation, two populations diverge into separate species while overlapping in geographic range and having no a priori barriers to gene flow. This model of speciation is thought to be quite uncommon, because it requires the rapid evolution of an extremely strong barrier to gene flow that is not physical (a behavior barrier is the most likely avenue through which this could occur) (Coyne and Orr, 2004). In addition to likely being quite rare, it is difficult to prove that two species have diverged in sympatry: the only proven instance is that of two species of *Arecaceae* palm tree on an oceanic island (Savolainen et al., 2006). One potential mechanism might be the colonization of a novel niche by one population in an organism with mating behaviors tightly tied to foraging behaviors (e.g. *Drosophila*), which may result in cessation of mating as the two populations can longer locate the other as potential mates. Alternatively, the evolution of strong pre-mating barriers might result in the cessation of gene flow as the two populations stop mating and diverge through genetic drift (Wu et al., 1995, Hollocher et al. 1997).

It is important to consider the mode through which two species diverged, as it can have substantial impact on those species and the interactions between species post-divergence. As noted before, species diverging in peripatry may suffer severe founder effects as a result of reduced population size, which may have severely negative consequences for that species (Coyne and Orr, 2004, Provine, 2004). More broadly, whether species diverge in allopatry or sympatry has significant effects on later interactions between those species. Two species which diverged in allopatry are much less likely to have evolved strong pre-mating isolation, as this was never a requisite for cessation of gene flow. Conversely, two species which have diverged in sympatry are much more likely to have evolved strong pre-mating isolation, as this is the only way gene flow could have been reduced enough to allow

speciation to occur (Figure 1.2) (Coyne and Orr, 1989, Coyne and Orr, 1997, Coyne and Orr, 2004). The strength of pre-mating isolation between two species can determine how readily or not those species hybridize (Coyne and Orr, 2004). When studying mating isolation between two species, it is important to consider whether they diverged in allopatry or in sympatry, as well as which state they exist in now. Species which diverged in allopatry are more likely to have weak pre-mating isolation and strong post-mating isolation, for example. This distinction can inform both the types of studies which are appropriate as well as how the data from these studies should be analyzed.

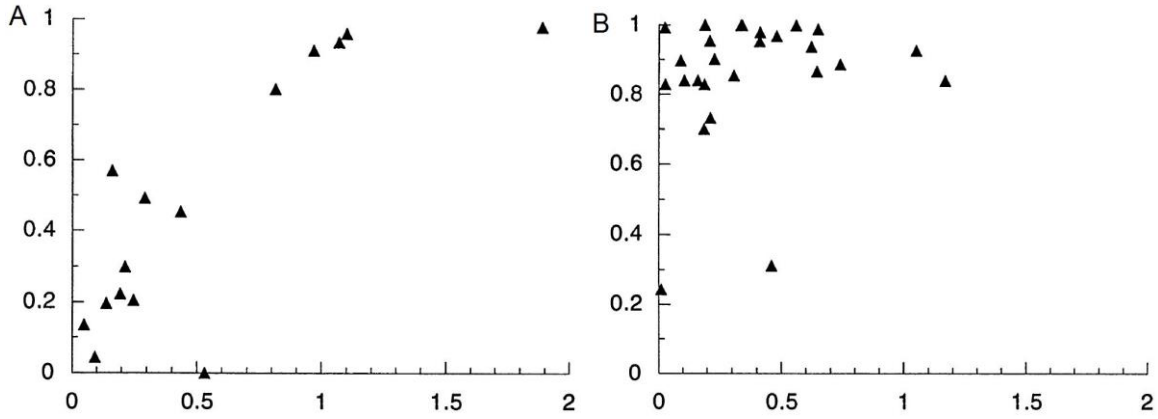

Figure 1.2 Pre-mating isolation is stronglyaffected by proximity

Species existing in allopatry to each other (A) have substantially lower strength of pre-mating isolation than species which exist in sympatry (B). Y-axis: relative strength of pre-mating reproductive isolation, 0 being none and 1 being total isolation. X-axis: Nei's genetic distance between the two species. Figure adapted from Coyne and Orr, 1997.

Reproductive Isolation

Prevalence of mating between populations and the outcomes of mating are necessary concepts in studying and understanding speciation and the divisions between species. The umbrella term reproductive isolation describes three main questions: whether two species will mate, whether they will successfully bear progeny, and what fitness consequences the progeny have. These three types of barrier are termed pre-mating, post-mating pre-zygotic, and post-mating post-zygotic. (Orr, 1995, Coyne and Orr, 2004).

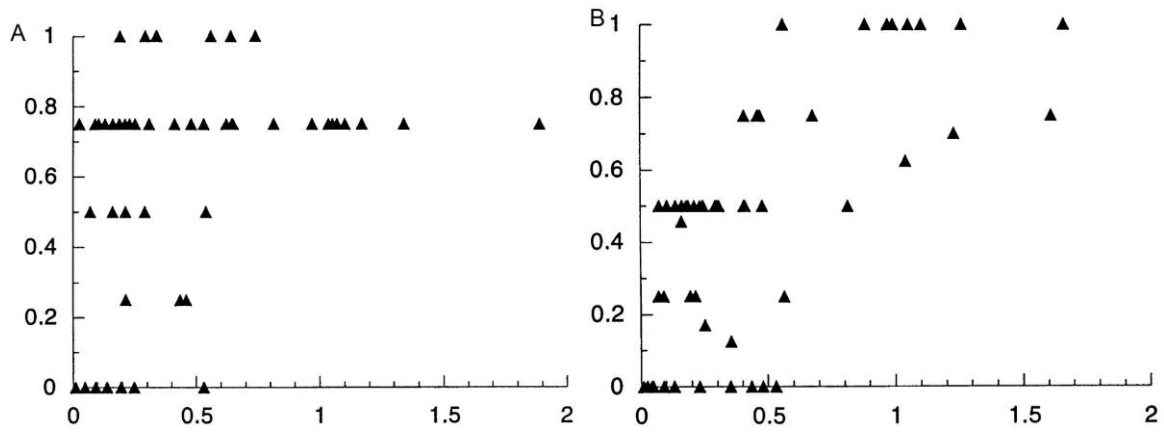

Figure 1.3 Pre-Mating isolation evolves more rapidly than post-mating isolation
Strength of pre-mating isolation (A) between species reaches high levels more quickly than strength of post-mating isolation (B) between species. Y-axis: relative strength of reproductive isolation, 0 being none and 1 being total isolation. X-axis: Nei's genetic distance between the two species. Figure adapted from Coyne and Orr, 1997.

Pre-Mating Isolation

Pre-mating barriers describe the set of barriers which cause two species not to mate. These barriers can involve simple differences such as physical incompatibilities between the species, or complex behavioral differences (Arthur and Dyer, 2015). Pre-mating barriers are crucial to our study of speciation, as they are thought to evolve more rapidly than either type of post-mating barrier (Figure 1.3) (Coyne and Orr, 1989, Coyne and Orr, 1997). Because of

this, it is likely that pre-mating barriers drive speciation. Pre-mating barriers are also weaker than post-mating barriers, however. For this reason, they are thought to be crucial for the initiation of speciation but less important for its maintenance (Rosenblum et al., 2012, Cenzer et al., 2015).

Pre-mating barriers evolve more strongly in sympatry than in allopatry (Coyne and Orr, 1989, Coyne and Orr, 1997), which has two main functional consequences. First, in any case except total allopatry, pre-mating barriers will necessarily be drivers of speciation, due to their rapid and strong evolution. Secondly, evolution of pre-mating barriers following speciation can have fitness consequences for species. Comeault et al. 2016 shows that evolution of pre-mating barriers between populations of separate species existing in sympatry can have fitness consequences for those species. In this case, *D. yakuba* males from areas of sympatry with *D. santomea* sire fewer progeny than allopatric males when mated to *D. yakuba* females, suggesting that the evolution of pre-mating barriers with *D. santomea* has negatively affected the fitness of sympatric *D. yakuba* (Comeault et al., 2016).

Post-Mating Pre-Zygotic Isolation

Post-mating barriers are slower to evolve than pre-mating barriers (Figure 1.3) (Coyne and Orr, 1989, Coyne and Orr, 1997), but are important to prevent fusion of nascent species after divergence has occurred (Rosenblum et al., 2012). They can be sorted into two main categories: pre-zygotic and post-zygotic (Coyne and Orr, 2004). Pre-zygotic barriers occur when species mate successfully, but are unable to successfully form a zygote. These barriers typically involve incompatibilities between the male and female reproductive tracts (Ahmed-Braimah, 2016, Coyne and Orr, 2004). Frequently, sperm fail to fertilize the egg (an important distinction from successful fertilization followed by zygotic or embryonic

lethality). Post-mating pre-zygotic barriers seem to evolve between the onset of pre-mating barriers and the evolution of post-mating post-zygotic barriers (Alipaz et al., 2001).

Post-Mating Post-Zygotic Isolation

Post-zygotic barriers occur when a zygote forms but there is a fitness consequence for the hybrid offspring. This frequently involves reductions in hybrid fertility or total lethality of the hybrid organism (Maheshwari and Barbash, 2011, Coyne and Orr, 1997). It is not uncommon for one gender of offspring to be fertile while the other is sterile, or one gender to be viable but sterile while the other is inviable. Almost always, the heterogametic sex is the more severely affected in these cases, a pattern termed Haldane's Rule (Turelli and Orr, 1995, Delph, 2016). These post-zygotic barriers are importance for maintenance of species, as they both hinder gene flow between the species and provide strong negative fitness consequences for hybrids (Rosenblum et al., 2012). Post-zygotic isolation frequently results from the accumulation of alleles in each species which become deleterious when interacting in hybrids. These novel interactions are termed Dobzhansky-Muller Incompatibilities (Dobzhansky et al., 1942).

Dobzhansky-Muller Incompatibilities occur when loci which have diverged in the parental species have novel interactions with deleterious consequences in hybrids. If the ancestral state of a pair of loci is A-B, and they diverge to a-B and A-b in the two parental species, the hybrid genotype a-b will have negative consequences due to novel interactions between alleles "a" and "b" (Fig 1.4). DMIs are thought to underlie most deleterious hybrid phenotypes (Sawamura, 2016), and have been extensively studied, though rarely conclusively proven. One example is the *hybrid male rescue* gene found in *Drosophila*, which has been shown to have lethal interactions in hybrids between *Drosophila melanogaster* and related species (Barbash et al., 2000, Cooper and Phadnis, 2016). Much

work on hybridization has been performed in *Drosophila* species, due to their ease of use as a model system, ability to hybridize even with distantly related species (e.g. *Drosophila melanogaster* is able to hybridize with *Drosophila santomea*, from which it diverged 15 million years ago), and the enormous array of genetic tools available for use in *Drosophila* species (Matute et al., 2009, Turissini et al., 2015).

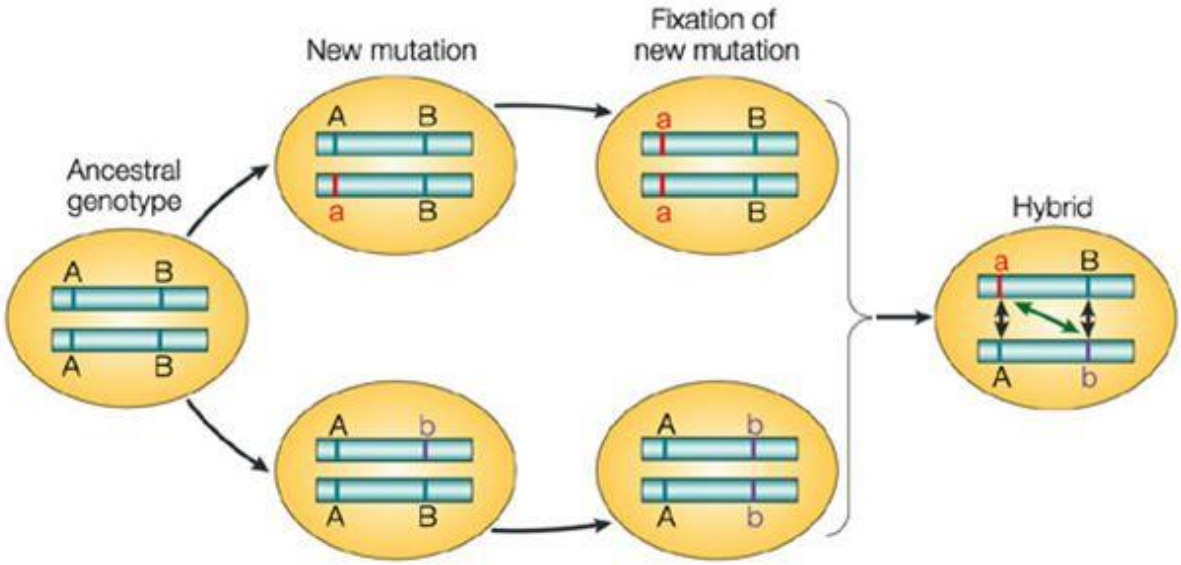

Figure 1.4 Schematic of the evolution of Dobzhansky-Muller Incompatibilities

Dobzhansky-Muller Incompatibilities arise when two populations separate and begin to diverge through genetic drift. Over time, the accumulation of mutations results in allelic combinations which have negative fitness consequences in hybrid offspring due to novel interactions between these alleles. (Johnson, 2008)

Hybridization

Hybrids occur when two separate species mate and produce viable offspring; as a result, their genetics are often quite unusual and can result in unexpected effects such as hybrid gigantism (Reisinger, 1929). Typically, however, hybrids display phenotypes intermediate to the two parental species. If the two parental species are sufficiently specialized, this can have deleterious consequences for the hybrid (Coyne and Orr, 2004). It

is theorized that these intermediate phenotypes may also allow hybrids to colonize novel ecological niches and undergo parapatric or peripatric speciation (Coyne and Orr, 2004, Schumer et al., 2014). Hybrid zones have been found in nature in surprising frequency (Taylor et al., 2015), and the study and observation of hybrid zones have been very helpful for our understanding of speciation and evolution (Coyne and Orr, 2004).

Many species that do not hybridize in nature can be made to hybridize in the lab, and hybrids have been extensively researched (Mallet, 2007). Studying hybrids is useful not only for understanding the genetics of the two parental species, but also advances our understanding of the genetic interactions that maintain barriers between species and reduce fitness of hybrid offspring, impeding gene flow (Orr et al., 2007). Generally, hybrids are overall less fit than their parental species (if hybrids were more fit, we would expect that the division between the two species would slowly disappear as offspring of the hybrids eventually out-compete the pure species) (Inoue and Watanabe, 1979). Fitness consequences in hybrids frequently involve one sex being sterile, generally the heterogametic sex (Turelli and Orr, 1995, Delph, 2016) (alternatively, one sex is sterile, and the other dies before reaching adulthood).

The *Drosophila melanogaster* Species Subgroup

The *melanogaster* species subgroup comprises a total of 9 species (Fig 1.5) with divergence times as high as 15 million years (*D. melanogaster* and *D. santomea*) and as low as 1 million years (*D. santomea* and *D. yakuba*) (Coyne et al., 2004, Turissini et al., 2015). Species display highly varied levels of ecological specialization. *D. melanogaster* and *D. simulans*, for example, are highly cosmopolitan species able to breed and flourish in a wide variety of environments and temperatures, and can subsist on many substrates (Austin and Moehring, 2013). *D. santomea*, found only on the islands of Sao Tome and Principe, is a

temperature specialist endemic to the cloud forest highlands on the islands where it lives (Turissini et al., 2015). *D. sechellia* is perhaps the most specialized, subsisting solely on the toxic fruit *Morinda citrifolia*. (Yassin et al., 2016). *Drosophila* species have been instrumental to our understanding of reproductive isolation, including determining how quickly DMIs evolve between diverged species (Matute et al., 2010).

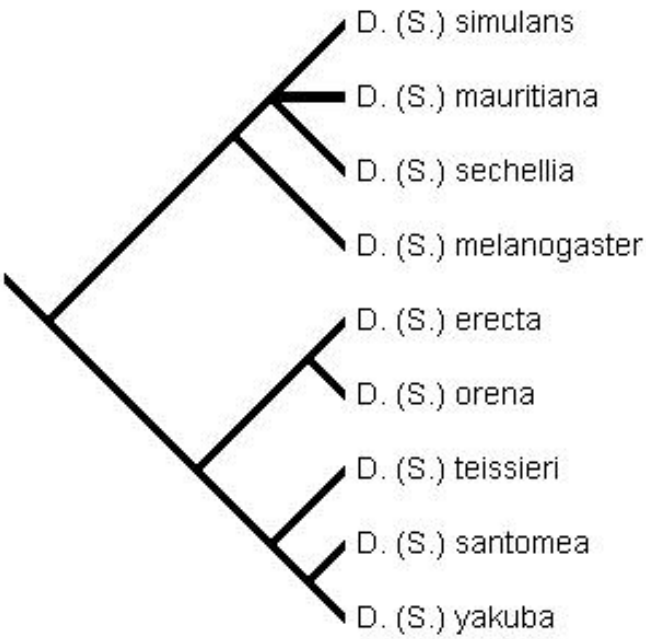

Figure 1.5 The *Drosophila melanogaster* species subgroup

The *Drosophila melanogaster* species subgroup comprises a total of 9 species with divergence times varying between 1 to 15 million years, and generalist as well as specialist species. It is arguably the most studied group of species in the world. (van der Linde, 2006)

Studying Reproductive Isolation in the *melanogaster* Species Subgroup

Drosophila have been an extraordinarily useful research tool in understanding both the genetic basis of hybrid sterility/inviability and the genetics of reproductive isolation. *D. melanogaster* and its sister species have been especially helpful in this regard, particularly hybrids between *D. melanogaster* and *D. simulans* (Stanley et al., 1980). Researchers have mapped the genetic bases of male hybrid lethality to an interaction between three major

genes: *Hmr*, *Lhr*, and *gfzf* (Barbash et al., 2000, Cooper and Phadnis, 2016). This example confirms the Dobzhansky-Muller model of hybrid incompatibilities, and has been enormously informative for our understanding of epistatic interactions between gene products and the role they play in reproductive isolation. Additionally, two alleles affecting hybrid females in this cross have also been mapped: *Nup96* (Presgraves, 2003) and *Nup160* (Tang and Presgraves, 2009).

Hybrids between *D. santomea*, *D. yakuba*, and *D. teissieri* have also been highly informative. *D. teissieri* and *D. yakuba* diverged roughly 2.4 million years ago; *D. santomea* and *D. yakuba* diverged 1 million years ago (Turissini et al., 2015). Hybrids between any pair of these species display identical phenotypes: females are fertile, while males are viable yet sterile, an unexpected pattern, as the increased divergence time between *D. teissieri* and its sister species would normally result in a more severe phenotype for hybrid offspring. When the fertile F1 females are crossed to pure species male of either parental species, a small proportion of males become fertile (increasing in size with each successive backcross). Some work mapping the basis of male hybrid infertility in this clade has been performed, and found that the trait maps strongly to the X chromosome (Moehring et al., 2006). This remains a promising avenue of research to identify loci involved in DMIs.

In addition to their used in mapping the basis of post-zygotic isolation, the *melanogaster* group has also been useful for mapping pre-mating isolation. For example, altering the profile of cuticular hydrocarbons produced by *D. simulans* and *D. sechellia* has been found to affect mate choice in males and females. The Zimbabwe/cosmopolitan mate choice system of strong pre-mating isolation within *Drosophila melanogaster* has also been extensively studied (Wu et al., 1995, Hollocher et al., 1997, Greenberg et al., 2003).

In this thesis I answer two important questions in the study of reproductive isolation. One, is post-zygotic isolation affected by environmental factors, and if so how strongly? Two, what is the genetic basis of pre-mating isolation in the Zimbabwe and cosmopolitan races of *Drosophila melanogaster*? My results answer these two longstanding questions and enhance our understanding of the mechanisms and effects of reproductive isolation, both between and within species.

REFERENCES

1. Ahmed-Braimah, Y. H. (2016). Multiple genes cause postmating prezygotic reproductive isolation in the Drosophila virilis group. *Genes Genomes Genetics*, *X*, 1–17. http://doi.org/10.1534/g3.116.033340

2. Alipaz, J. A., Wu, C., & Karr, T. L. (2001). Gametic incompatibilities between races of Drosophila melanogaster. *Proceedings of the Royal Society of London*, *268*(July 2000), 789–795. http://doi.org/10.1098/rspb.2000.1420

3. Arthur, N.J., & Dyer, K. A. (2015). Asymmetrical sexual isolation but no postmating isolation between the closely related species Drosophila suboccidentalis and Drosophila occidentalis. *BMC Evolutionary Biology*, *15*(March 2015), 1-9. http://doi.org/10.1186/s12862-015-0328-y

4. Austin, C. J., & Moehring, A. J. (2013). Optimal temperature range of a plastic species, Drosophila simulans. *Journal of Animal Ecology*, *82*, 663–672. http://doi.org/10.1111/1365-2656.12041

5. Baker, R. J., & Bradley, R. D. (2006). Speciation in mammals and the genetic species concept. *Journal of Mammalogy*, *87*, 643–662.

6. Barbash, D. A., Roote, J., & Ashburner, M. (2000). The Drosophila melanogaster hybrid male rescue gene causes inviability in male and female species hybrids. *Genetics*, *154*, 1747–1771.

7. Butlin, R. K., Galindo, J., & Grahame, J. W. (2008). Sympatric, parapatric or allopatric: the most important way to classify speciation? *Philosophical Transactions of the Royal Society B*, *363*(=), 2997–3007. http://doi.org/10.1098/rstb.2008.0076

8. Cenzer, M. L. (2016). Adaptation to an invasive host is driving the loss of a native ecotype. *Evolution*, (70), 2296–2307. http://doi.org/10.1111/evo.13023.This

9. Comeault, A. C., Venkat, A., & Matute, D. R. (2016). Correlated evolution of male and female reproductive traits drive a cascading effect of reinforcement in Drosophila yakuba. *Proceedings of the Royal Society B*, *283*

10. Cooper, J. C., & Phadnis, N. (2016). A genomic approach to identify hybrid incompatibility genes. *Fly*, *10*, 142–148. http://doi.org/10.1080/19336934.2016.1193657

11. Coyne, J. A., & Orr, H. A. (1989). Patterns of Speciation in Drosophila. *Evolution*, *43*(2), 362–381.

12. Coyne J. A., O. H. A. (1997). "Patterns of Speciation in Drosophila" Revisited. *Evolution*, *51*(1), 295–303.

13. Coyne, J. A., & Orr, H. A. (2004). *Speciation* (First Edit). Sinauer Associates.

14. David, J. R., & Capy, P. (1988). Genetic variation of Drosophila melanogaster natural populations. *Trends in Genetics*, *4*(4), 106–111.

15. Delph, L. F., & Jeffery, P. (2016). Haldane's Rule: genetic bases and their support. *Journal of Heredity*, 383–391. http://doi.org/10.1093/jhered/esw026

16. Dobzhansky, T., Holz, A. M., & Spassky, B. (1942). Genetics of natural populations. VIII. Concealed variability in the second and the fourth chromosomes of Drosophila pseudoobscura and its bearing on the problem of heterosis. *Genetics*, *27*, 463–490.

17. Fitzpatrick, B. M., Fordyce, J. A., & Gavrilets, S. (2008). What, if anything, is sympatric speciation? *Journal of Evolutionary Biology*, *21*, 1452–1459. http://doi.org/10.1111/j.1420-9101.2008.01611.x

18. Flynn, J. J., Nedbal, M. A., Dragoo, J. W., & Honeycutt, R. L. (2000). Whence the red panda? *Molecular Phylogenetics and Evolution*, *17*, 190–199. http://doi.org/10.1006/mpev.2000.0819

19. Greenberg, A. J., Moran, J. R., Coyne, J. A., & Wu, C.-I. (2003). Ecological adaptation during incipient speciation revealed by precise gene replacement. *Science*, *111*, 1754–1758.

20. Guggisberg, C. A. W. (1975). *Wild Cats of the World* (First Edit). Taplinger Pub Co.

21. Hollocher, H., Ting, C.-T., Pollack, F., & Wu, C.-I. (1997). Incipient speciation by sexual isolation in Drosophila melanogaster: variation in mating preference and correlation between sexes. *Evolution*, *51*(4), 1175–1181.

22. Inoue, Y., & Watanabe, T. K. (1979). Inversion polymorphisms in Japanese natural populations of Drosophila melanogaster. *Japanese Journal of Genetics*, *54*, 69–82.

23. Johannesson, K. (2010). Are we analyzing speciation without prejudice? *Annals of the New York Academy of Sciences*, *1206*, 143–149. http://doi.org/10.1111/j.1749-6632.2010.05701.x

24. Johnson, N. A. (2008). Hybrid incompatibility and speciation. *Nature Education*, *1*, 20.

25. Krempels, D. (2006). Spatial aspects of speciation. Retrieved October 20, 2016, from https://en.wikipedia.org/wiki/File:Speciation_modes.svg

26. Llopart, A., Herrig, D. E., Brud, E., & Lein, Z. S. (2014). Sequential adaptive introgression of the mitochondrial genome in Drosophila yakuba and Drosophila santomea. *Molecular Ecology*, *23*, 1124–1136. http://doi.org/10.1111/mec.12678

27. Maheshwari, S., & Barbash, D. A. (2011). The genetics of hybrid incompatibilities. *Annual Review of Genetics*, *45*, 331–355. http://doi.org/10.1146/annurev-genet-110410-132514

28. Mallet, J. (2007). Hybrid speciation. *Nature*, *446*, 279–283. http://doi.org/10.1038/nature05706

29. Matute, D. R., Butler, I. A., & Coyne, J. A. (2009). Little effect of the tan locus on pigmentation in female hybrids between Drosophila santomea and D. melanogaster. *Cell*, *139*, 1180–1188. http://doi.org/10.1016/j.cell.2009.10.033

30. Matute, D. R., Butler, I. A., Turissini, D. A., & Coyne, J. A. (2010). A test of the snowball theory for the rate of evolution of hybrid incompatibilities. *Science*, *329*, 1518–1522.

31. Matute, D. R. (2010a). Reinforcement of gametic isolation in Drosophila. *PLoS Biology*, *8*, 1–11. http://doi.org/10.1371/journal.pbio.1000341

32. Matute, D. R. (2010b). Reinforcement can overcome gene flow during speciation in Drosophila. *Current Biology*, *20*, 2229–2233. http://doi.org/10.1016/j.cub.2010.11.036

33. Moehring, A. J., Llopart, A., Elwyn, S., Coyne, J. A., & Mackay, T. F. C. (2006). The genetic basis of postzygotic reproductive isolation between Drosophila santomea and D yakuba due to hybrid male sterility. *Genetics*, *233*, 225–233. http://doi.org/10.1534/genetics.105.052985

34. Orr, H. A., Masly, J. P., & Phadnis, N. (2007). Speciation in Drosophila: from phenotypes to molecules. *Journal of Heredity*, *98*, 103–110. http://doi.org/10.1093/jhered/esl060

35. Orr, H. A. (1995). The Population Genetics of Speciation: the Evolution of Hybrid Incompatibilities. *Genetics*, *139*, 1805–1813.

36. Orr, H. A., Madden, L. D., Coyne, J. A., Goodwin, R., & Hawley, R. S. (1997). The Developmental Genetics of Hybrid Inviability: A Mitotic Defect in Drosophila Hybrids. *Genetics*, *145*, 1031–1040.

37. Presgraves, D. C. (2003). A fine-scale genetic analysis of hybrid incompatibilities in Drosophila. *Genetics*, *163*, 955–972.

38. Provine, W. B. (2004). Anecdotal, historical, and critical commentaries on genetics. *Genetics*, *167*, 1041–1046.

39. Reisinger, L. (1929). Bastardierung. *Zoologischer Anzeiger*, *81*, 254–260.

40. Rosenblum, E. B., Sarver, B. A. J., Brown, J. W., Des Roches, S., Hardwick, K. M., Hether, T. D., … Harmon, L. J. (2012). Goldilocks meets Santa Rosalia: an ephemeral speciation model explains patterns of diversification across time scales. *Evolutionary Biology*, *39*, 255–261. http://doi.org/10.1007/s11692-012-9171-x

41. Savolainen, V., Anstett, M., Lexer, C., Hutton, I., Clarkson, J. J., Norup, M. V, … Baker, W. J. (2006). Sympatric speciation in palms on an oceanic island. *Nature*, *441*, 9–12. http://doi.org/10.1038/nature04566

42. Sawamura, K. (2016). Genome-wide analyses of hybrid incompatibility in Drosophila. *Advanced Techniques in Biology and Medicine*, *4*, 10–12. http://doi.org/10.4172/2379-1764.1000159

43. Schumer, M., Rosenthal, G. G., & Andolfatto, P. (2014). How common is homoploid hybrid speciation? *Evolution, 68*, 1553-60. doi: 10.1111/evo.12399.

44. Stanley, M., Parsons, P. A., Spence, G. E., & Weber, L. (1980). Resistance of species of the Drosophila melanogaster subgroup to environmental extremes. *Australian Journal of Zoology*, *28*, 413–421.

45. Tang, S., & Presgraves, D. C. (2009). Evolution of the Drosophila nuclear pore complex results in multiple hybrid incompatibilities. *Science*, *323*, 779–782.

46. Turelli, M., & Orr, H. A. (1995). The dominance theory of Haldane's Rule. *Genetics*, *140*, 389–402.

47. Turissini, D. A., Liu, G., David, J. R., & Matute, D. R. (2015). The evolution of reproductive isolation in the Drosophila yakuba complex of species. *Journal of Evolutionary Biology*, *28*, 557–575. http://doi.org/10.1111/jeb.12588

48. van der Linde, K. (2006). Drosophila melanogaster subgroup phylogeny. Retrieved October 20, 2016, from https://en.wikipedia.org/wiki/File:Drosophila_melanogaster_subgroup_phylogeny.jpg

49. Vences, M., Guayasamin, J. M., Miralles, A., & Riva, I. (2013). To name or not to name: criteria to promote economy of change in Linnaean classification schemes. *Zootaxa*, *3636*, 201–244.

50. Wu, C., Hollocher, H., Begun, D. J., Aquadro, C. F., Xu, Y., & Wu, M.-L. (1995). Sexual isolation in Drosophila melanogaster : A possible case of incipient speciation. *Proceedings of the National Academy of Sciences*, *92*(March), 2519–2523.

51. Yassin, A., Debat, V., Bastide, H., Gidaszewski, N., David, J. R., & Pool, J. E. (2016). Recurrent specialization on a toxic fruit in an island Drosophila population. *PNAS*, *113*, 4771–4776. http://doi.org/10.1073/pnas.1522559113.

CHAPTER TWO : THE EFFECTS OF TEMPERATURE ON *DROSOPHILA* HYBRID FITNESS[1]

Introduction

Reproductive barriers hamper gene flow between species (Coyne and Orr 2004). Depending on when in the reproductive cycle barriers occur, they can be classified as prezygotic or postzygotic. Phenotypes that prevent the successful formation of a zygote, such as certain behavioral or gametic incompatibilities, can lead to prezygotic isolation (reviewed in Coyne and Orr 2004). Conversely, postzygotic isolation manifests as defects in hybrids and includes a range of phenotypic defects such as developmental breakdown and behavioral abnormalities (reviewed in Maheshwari and Barbash 2011). In its most extreme form, postzygotic isolation results in hybrid inviability.

The evolution of postzygotic isolation is crucial to speciation for at least three reasons. First, even though comparative studies have strongly suggested that prezygotic isolation tends to evolve faster than postzygotic isolation (Coyne and Orr 1989, Orr et al. 1997, Mendelson 2003, Rabosky and Matute 2013), they are often not strong enough to prevent the fusion of nascent species (Rosenblum et al. 2012, Comeault et al. 2015, Cenzer 2016). Postzygotic barriers are more robust and are often crucial to maintaining separation of species. Second, hybrid defects can also influence the evolution of other barriers to gene flow (reviewed in Servedio and Noor 2003, Hopkins 2013). For example, in the process of reinforcement, prezygotic isolation becomes stronger in areas of sympatry due to indirect

[1] A manuscript accepted for publication in Genes|Genomes|Genetics

selection on hybrids with deleterious phenotypes (Noor and Servedio 2003, Hudson and Price 2014). Finally, studying postzygotic isolation, and other traits that reduce fitness in hybrids, can reveal how much functional divergence has occurred between the genomes of the parent species, furthering our understanding of the processes that initiate and maintain separation of species (Coyne and Orr 2004, Orr et al. 2007, Rosenblum et al. 2012).

Postzygotic isolation frequently results from Dobzhansky-Muller incompatibilities (reviewed in Coyne and Orr 2004, and Nosil and Schluter 2011) According to the Dobzhansky-Muller model, deleterious epistatic interactions between alleles from different species reduce fitness in hybrids (Muller et al. 1937, Dobzhansky et al. 1942, Orr 1995, Coyne and Orr 2004). The model requires at least two interacting loci that evolve separately in allopatric populations. Postzygotic isolation arises as a collateral effect when the species come into secondary contact and hybridize. For example, the ancestral alleles at a pair of loci are "a-b," and two allopatric populations subsequently diverge into "a-B" and "A-b." The hybrid genotype "A-B" has deleterious consequences due to the interaction between the "A" and "B" alleles, which were only present together in the hybrid. Dobzhansky-Muller incompatibilities (DMIs) are frequently implicated in the defects observed in many interspecies hybrids, and thus are crucial to understanding how species form and persist over time.

Several mapping efforts have succeeded in characterizing the developmental defects underlying hybrid inviability as well as the causal alleles in some cases (reviewed in Nosil and Schluter 2011, Maheshwari and Barbash 2012, Sawamura 2016). These studies have revealed two general trends regarding the evolution of postzygotic isolation. First, sex chromosomes frequently harbor alleles that lead to sterility and inviability in hybrids (Masly

and Presgraves 2007, Carrington et al. 2011), which may explain a pattern known as 'Haldane's rule': when hybrids have a defect, the heterogametic sex is typically more severely affected (Orr et al. 1997, Delph and Demuth 2016). Second, hybrid incompatibilities accumulate at an exponential rate through a process known as the snowball theory, a key prediction of the Dobzhansky-Muller model (Orr 1995, Orr and Turelli 2001, Moyle and Nakazato 2010, Matute et al. 2010, Matute and Gavin-Smyth 2014, Wang et al. 2015).

Drosophila hybrids have been crucial for our understanding of the genetic basis of hybrid inviability (Orr et al. 2007, Aruna et al. 2009). In particular, studying crosses between *D. melanogaster* females and *D. simulans* males has been one of the most informative for investigating the genetic basis of postzygotic isolation. *D. simulans* is thought to have originated in Southeast Africa and is widespread around the globe, and has a similar thermal tolerance and niche preference to *D. melanogaster* (Stanley et al. 1980, Austin and Moehring 2013). Interspecific crosses between *D. melanogaster* females and *D. simulans* males produce only sterile hybrid females; male offspring die as larvae (Sturtevant 1920, Inoue and Watanabe 1979). The genetic basis of hybrid male lethality has been finely mapped and at least three loci, one on each major chromosome, have been found to be involved in the epistatic interaction responsible for male hybrid inviability. Different alleles are fixed in the gene triad *Hmr/Lhr/gfzf* between *D. melanogaster* and *D. simulans*, and their interaction in hybrid offspring is deleterious (Barbash et al. 2000; Phadnis et al. 2015; Cooper and Phadnis 2016). Additionally, two alleles influencing the viability of hybrid females have also been mapped: *Nup96* (Presgraves 2003), and *Nup160* (Tang and Presgraves 2009).

Drosophila melanogaster can also hybridize with species to which it is even more distantly related than *D. simulans* (Matute et al. 2009a, Turissini et al. 2016). The cross

between *D. melanogaster* and *D. santomea* also produces only hybrid females (Matute et al. 2009a); males fail to develop the distal half of the abdomen and die as embryos (Matute and Gavin-Smyth 2014). This cross is the most divergent known to produce hybrid progeny in *Drosophila* (Matute et al. 2010). *Drosophila santomea* is endemic to the highlands of São Tomé, a volcanic island off the coast of Cameroon (Lachaise et al. 2000). On the extinct volcano of Pico de São Tomé, *D. santomea* occupies the mist forests of the island at high elevations where it is thought to breed on figs of the endemic subspecies *Ficus chlamydocarpa fernandesiana* (Lachaise et al. 2000, Llopart et al. 2005a, Llopart et al. 2005b). Within the *melanogaster* species subgroup, *D. santomea* and *D. simulans* have very different life history traits, whereas *D. simulans* and *D. melanogaster* are more similar (Capy and Gibert 2004). For example, *D. melanogaster* and *D. simulans* are both globally distributed (Capy and Gibert 2004), but *D. santomea* is restricted to the high altitudes of São Tomé. Similarly, *D. melanogaster* and *D. simulans* are temperature generalists, while *D. santomea* is a temperature specialist.

In previous studies of hybrids between *D. melanogaster* and *D. simulans*, the penetrance of a few hybrid inviability alleles has been found to be largely independent of environmental factors (Barbash et al. 2000, Presgraves et al. 2003, Tang and Presgraves 2009). Nonetheless, other hybrid inviability loci might be affected by extrinsic factors (Coyne et al. 1998, Presgraves et al. 2003). For example, temperature has been shown to affect the magnitude of hybrid inviability in several clades (*Tribolium* beetles: Wade et al. 1999, Dowling et al. 2007; *Nasonia* wasps: Bordenstein et al. 2001, Koevoets et al. 2012; *Nicotiana*: Yamada et al. 2000, Muralidharan et al. 2014). Crosses between *D. melanogaster* and *D. simulans* have been used to identify genomic regions in *D. simulans* associated with

hybrid inviability at two different temperatures (Coyne et al. 1998). Similarly, hybrids between *D. melanogaster* and *D. mauritiana* have revealed that alleles from *D. melanogaster* may also have different effects at different temperatures (Cattani and Presgraves 2012). Finally, temperature dependent rescue of male inviability by mutant *Hmr* has been shown in hybrids of *D. melanogaster* with both *D. simulans* and *D. mauritiana* (Hutter and Ashburner 1987). However, we know little regarding whether the same type of variance in penetrance occurs in other interspecific hybrids.

We tested whether environmentally dependent inviability can be observed in two *Drosophila* interspecific hybrids: *D. melanogaster/D. santomea* F1 females (*mel/san*) and *D. melanogaster/D. simulans* F1 females (*mel/sim*). Given that *D. santomea* is a temperature specialist (Matute et al. 2009a) and *D. simulans* is a generalist (Capy and Gibert 2004), we explored whether the penetrance of recessive inviability alleles in hybrids with *D. melanogaster* was affected by temperature. Our expectation was that hybrid inviability should be strongly affected by both the identity of the species involved in the interspecific crosses and the temperature at which hybrids developed. We hypothesized that *mel/san* hybrids would be much more strongly affected by temperature than *mel/sim* hybrids. Our results indicate that even though the penetrance of particular loci is affected by temperature in both the *mel/san* and *mel/sim* crosses, hybrid viability is more affected by temperature in *mel/sim* hybrids than in *mel/san* hybrids.

Results and Discussion

We identified the genomic regions containing recessive hybrid incompatibilities in the genomes of hybrids between *D. melanogaster* and either *D. santomea* or *D. simulans* when those hybrids were reared at 18ºC. Recessive hybrid incompatibility alleles at 24ºC were previously mapped in both hybrids (*mel/san* and *mel/sim;* Matute et al. 2010). We first

report the results for each species independently and then compare the results between species and temperatures.

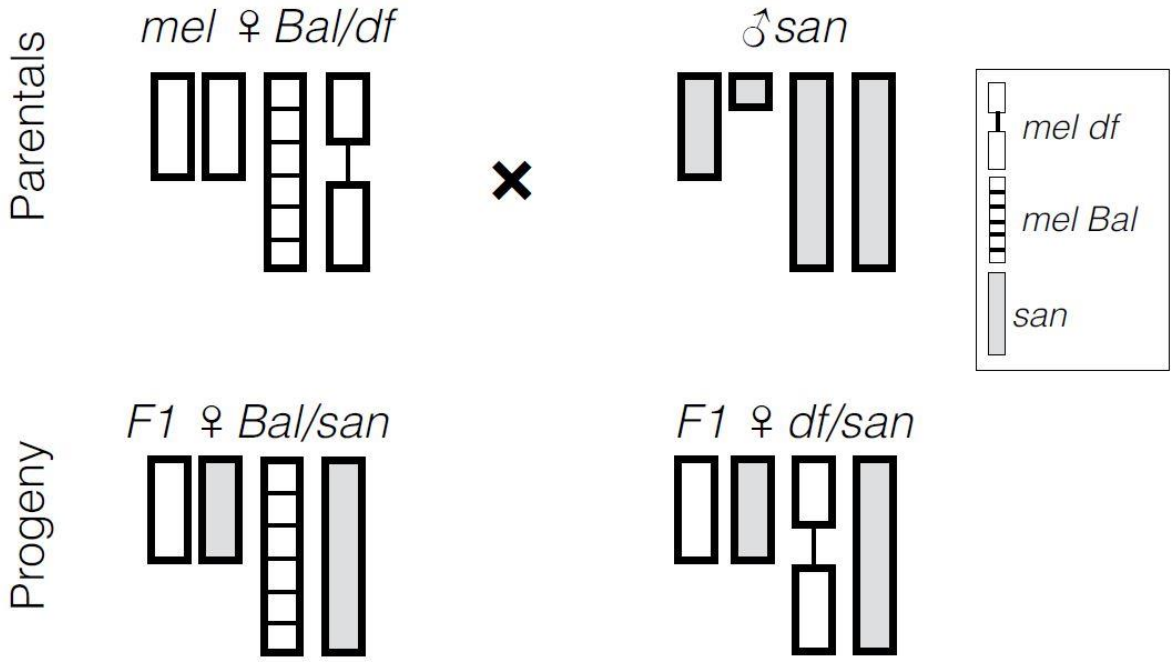

Figure 2.1 Deficiency mapping approach to detect alleles involved in hybrid inviability
A significant dearth of *df/san* individuals compared to their *Bal/san* sisters indicates that the deficiency uncovers a lethal, or semilethal allele involved in hybrid inviability. *D. melanogaster* balancer chromosomes are shown as striped bars; deficiency chromosomes are shown as a line connecting two bars. *D. santomea* chromosomes are shown in light grey. Sex chromosomes are shown as shorter bars than autosomes, and *Y* is shown as shorter than the *X*.

D. santomea

We used 223 *D. melanogaster* deficiency stocks (spanning 78.22% of euchromatic regions) and found 91 that caused partial or complete hybrid incompatibility when crossed to *D. santomea* at 18ºC (Figure 2). We compared our results with the map of inviability alleles at 24ºC, where 90 deficiencies caused hybrid inviability. We found that the slight plurality of deficiencies (56 deficiencies) caused hybrid inviability at both 24ºC and 18ºC. Thirty-five regions cause inviability only at 18ºC and 34 regions cause inviability only at 24ºC. The

overlap of incompatibilities between temperatures was significant (randomization tests: $P < 1 \times 10^{-4}$; Table S2). The same result is found if we assess the effect of temperature for the minimum number of hybrid incompatibilities (correcting for overlapping deficiencies which may share a common deficiency rather than represent several unique deficiencies). We found that the slight plurality of regions (43 regions), caused hybrid inviability at both 24º C and 18ºC. Twenty-nine regions cause inviability only at 18ºC and 31 regions cause inviability only at 24ºC. The overlap of incompatibilities between temperatures was also significant (randomization tests: $P < 1 \times 10^{-4}$). This is particularly interesting because we find the opposite pattern in the *mel/sim* cross (see below).

Figure 2.2 Deficiency mapping of hybrid incompatibilities in the *D. santomea* genome at two different temperatures

Light blue: hybrid inviability only at 18ºC. Green: hybrid inviability only at 24ºC. Dark blue: hybrid inviability at both temperatures. Deficiencies not causing hybrid inviability are not shown.

We found no difference in the relative density of incompatibilities across chromosomes. This was true for loci that cause hybrid inviability at only 18ºC (χ^2 = 2.6154, p-value = 0.913), at only 24ºC (χ^2 = 3.587, P = 0.609), and at both temperatures (χ^2 = 0.789, P = 0.977).

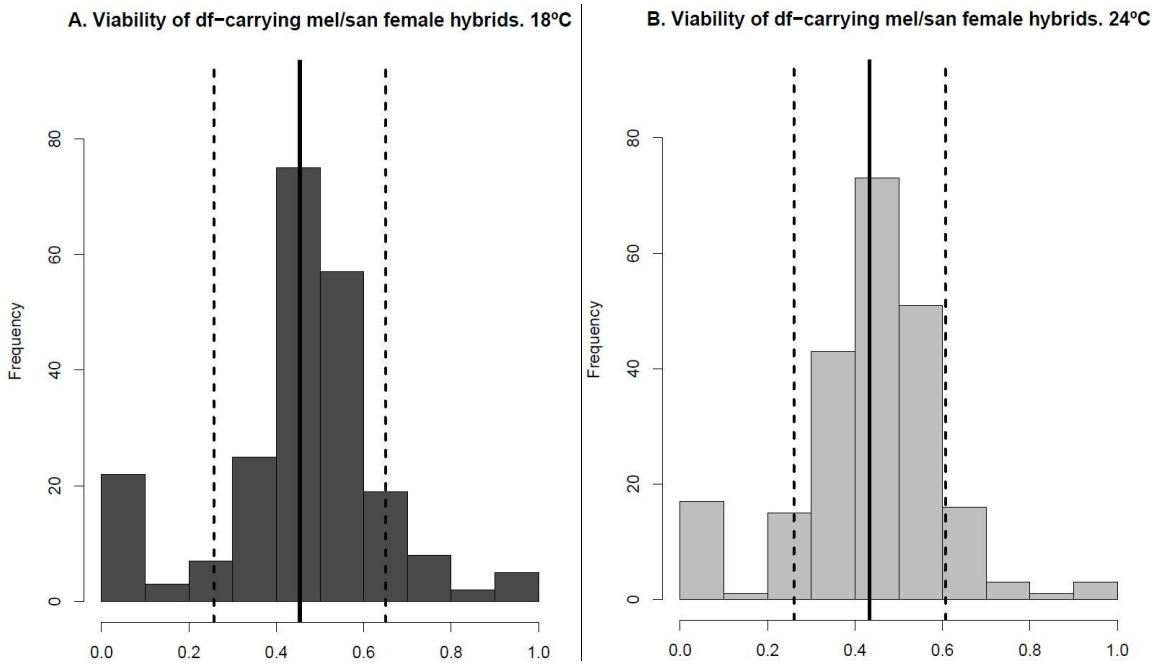

Figure 2.3 Relative fitness frequencies of the df-carrying hybrids in *mel/san* hybrids at two different temperatures

A. *mel/san* 18ºC. B. *mel/san* 24ºC. Black solid lines in each panel show the mean fitness of the *df*-carrying hybrids. Black dashed lines show the mean ± standard deviation of the mean. The X axis shows relative viability of deficiency-carrying progeny (observed *df*-carrying progeny/observed *Bal*-carrying progeny + observed *df*-carrying progeny) while the Y axis shows the number of stocks having a given level of viability of deficiency carrying offspring.

D. simulans

We used the same panel of 223 *D. melanogaster* deficiencies to detect hybrid incompatibilities in the *D. simulans* genome. At 18ºC we found 7 deficiency stocks that caused partial or complete hybrid incompatibility when crossed to *D. simulans* (Figure 2). We compared these results with the map of inviability alleles at 24ºC, where 17 deficiencies lead to HI. Of the previously reported deficiencies that uncovered hybrid incompatibilities,

we found that 16 of these 17 regions caused hybrid inviability at only 24ºC and not at 18ºC. There was no significant overlap of incompatibilities between temperatures (randomization tests: $P < 0.4312$; Table S3). The same result is found when we assess the effect of temperature on the minimal number of hybrid incompatibility regions (correcting for overlapping regions of deficiencies which may all uncover the same recessive lethal allele): only one region causes hybrid inviability at both temperatures; fourteen hybrid incompatibility alleles cause hybrid inviability at only 24ºC. Notably, we found a group of six deficiencies that only cause hybrid inviability at 18ºC, which corresponds to at least 5 hybrid incompatibility regions.

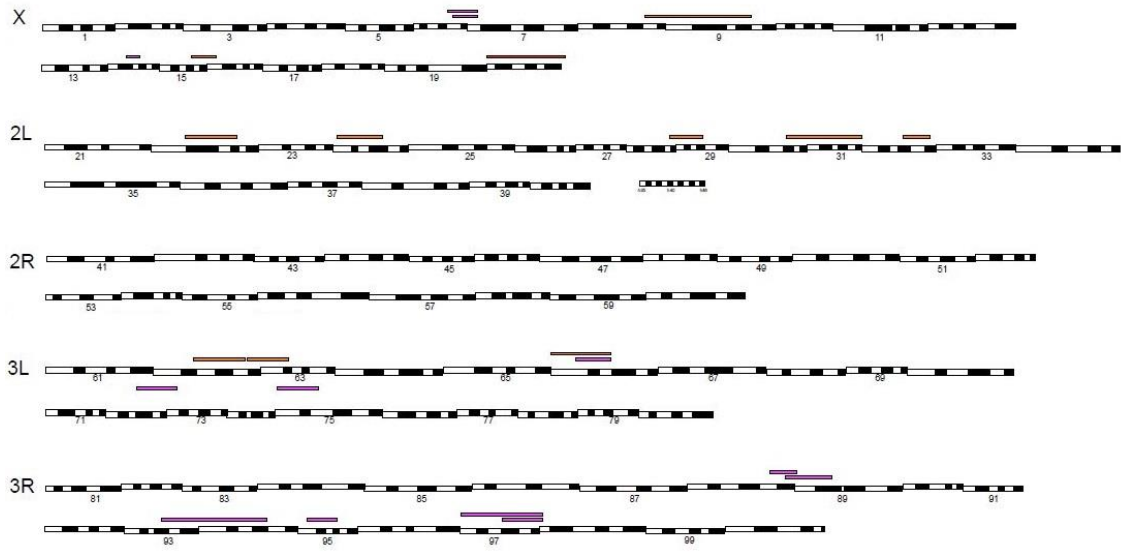

Figure 2.4 Deficiency mapping of hybrid incompatibilities in the *D. simulans* genome at two different temperatures

Orange: hybrid inviability only at 18ºC. Pink: hybrid inviability only at 24ºC. Red: hybrid inviability at both temperatures. Deficiencies not causing hybrid inviability are not shown.

Effect of temperature in HI alleles

We next compared the average effect size of exposing recessive alleles in the *D. santomea* genome in *mel/san* hybrid females at the two temperatures. This constitutes a test for the effect of temperature on the penetrance of alleles involved in hybrid inviability. First, we looked at the fitness distributions at the two temperatures for both hybrids: *mel/san* (Figure 3) and *mel/sim* (Figure 5). Even though 69 of 125 individual loci cause hybrid inviability at only one temperature in *mel/san* hybrids (Table S2, see above), the genome-wide effect of temperature on the fitness of *df*-carrying hybrids is modest and non-significant (mean difference between fitness at 18ºC and 24ºC= -0.0120; 95% CI: [-0.0488, 0.0089]; paired t-test; t = -1.3659, df = 222, P = 0.1734). In *mel/sim* hybrids, we found that temperature has a strong effect on the fitness of *df*-carrying hybrids, and that these hybrids are more viable at 24ºC (mean difference between fitness at 18ºC and 24ºC= -0.0256; 95% CI: [-0.0426, -0.0086]; paired t-test; t = -2.968, df = 222, P = 3.326×10^{-4}).

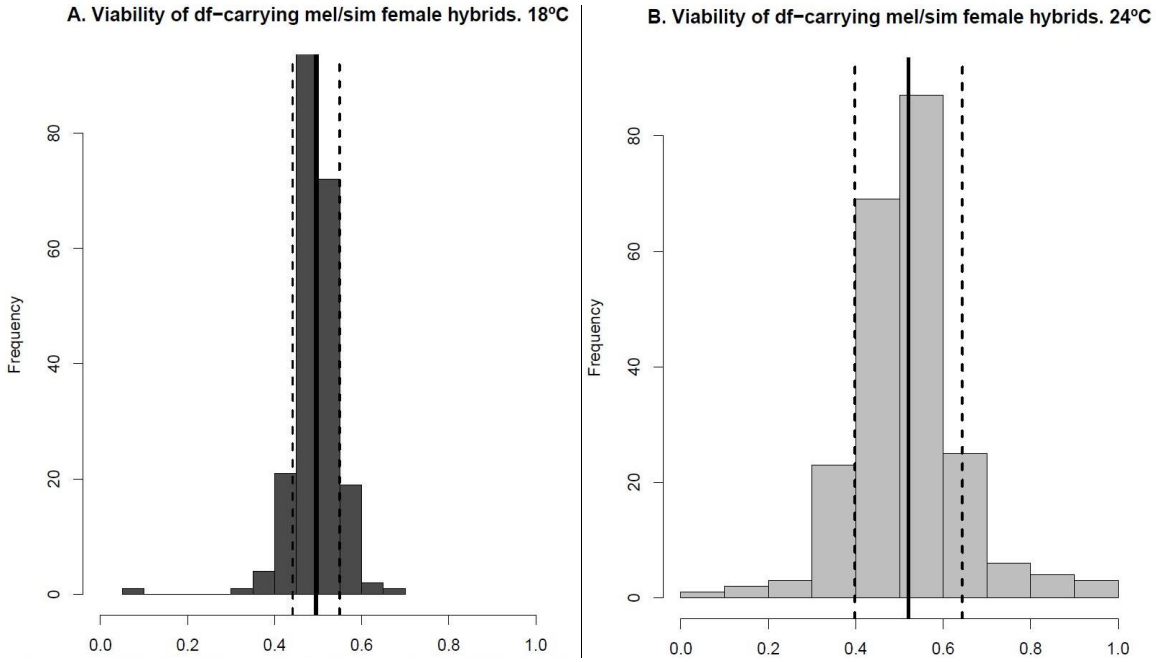

Figure 2.5 Fitness frequencies of the df-carrying hybrids in *mel/sim* hybrids at two different temperatures

A. *mel/sim* 18ºC. B.*mel/sim* 24ºC. Black solid lines in each panel show the mean fitness of the *df*-carrying hybrids. Black dashed lines show the average ± standard deviation from the mean. The X axis shows relative viability of deficiency-carrying progeny (observed *df*-carrying progeny/observed *Bal*-carrying progeny + observed *df*-carrying progeny) while the Y axis shows the number of stocks having a given level of viability of deficiency carrying offspring.

We also fit a linear model to jointly assess the relative contributions of paternal species and temperature. We found that the fitness distribution of *df*-carrying hybrids differs significantly between *mel/san* and *mel/sim* hybrids (Table 2.1, species effect). We also found that the magnitude of hybrid inviability is not affected by the rearing temperature *per se* (Table 2.1, temperature effect), but it is affected by the interaction between the parental species and rearing temperature (Table 2.1, temp × species interaction). To quantify the importance of the species, we fit a linear model dependent only on the temperature × species interaction. We found that temperature affects inviability differently between species ($F_{3,888}$ = 16.018 P = 3.785×10^{-10}). *df*-carrying *mel/sim* hybrids are more fit on average than *df*-

carrying *mel/san* at 18ºC (Linear contrasts with multiple comparison corrections: viability of *df/sim* - viability of *df/san* at 18ºC: Estimate = 0.0415; t-value= 2.980; P = 0.0157). Similarly, *df*-carrying *mel/sim* hybrids are also more fit than *df*-carrying *mel/san* hybrids at 24ºC (viability of *df/sim* - viability of *df/san* at 24ºC: Estimate = 0.08711; t-value= 6.252, P < 0.001).

Table 2.1 Levels of heterogeneity at relative viability of df(i)/(j) hybrids, where (i) represents a deficiency stock and (j) represents either of the two parental species

Linear model: HI ~ temp + species + (temp × species)					
	Degrees of freedom	Sum of squares	Mean square error	F value	Pr(>F)
Species	1	0.9226	0.92258	42.6207	$< 1.116 \times 10^{-10}$
Temperature	1	0.0018	0.00177	0.0818	0.77495
Temperature x Species	1	0.1159	0.11586	5.3523	0.02092
Residuals	888	19.2219	0.02165		

A linear model shows that the two fixed effects (temperature, species of the father), and the interaction between these two effects determines the viability of hybrids.

Effect of chromosomal location

We next explored whether temperature caused differences in the magnitude of hybrid inviability between *X*-linked and autosomal regions in the two interspecific hybrids. We found that temperature-dependent viability is not contingent on chromosomal location and that temperature has similar effects on *X*-linked and autosomal alleles in both hybrids (chromosome × species × temperature interaction: $F_{2,880} = 2.3132$, P = 0.0995). Given the large number of possible pairwise comparisons (66 comparisons, Table S4), we restricted our analyses to six comparisons, all within species and only accounting for the interaction term.

Pairwise comparisons within species confirmed that the effect of temperature on the penetrance of hybrid inviability alleles is minimal in both types of hybrids (*mel/sim* and *mel/san*) hybrids and that none of the three chromosomes is more prone to show differential hybrid inviability when raised at different temperatures (Table 2, rows 4-6).

Hybrid inviability is one of the most extreme phenotypes of reproductive isolation and constitutes both an important barrier to gene flow and an important mechanism for completing speciation (Coyne and Orr, 2004, Noor and Feder, 2006, Edmands, 2007). Although it has generally been considered to be more environmentally independent than prezygotic isolation (Coyne and Orr, 2004, Sobel and Randle, 2009), the penetrance of hybrid inviability is affected by extrinsic factors such as temperature (Wade and Johnson, 1994; Wade et al., 1999). In this report, we measured the penetrance of hybrid inviability alleles in two interspecific *Drosophila* crosses at two different temperatures. While temperature has a stronger effect on the penetrance of hybrid incompatibility loci in *mel/sim* hybrids than in *mel/san* hybrids, the overall results from both crosses suggest that temperature plays an important role in hybrid inviability. We found that *mel/san* hybrids have many more incompatibilities than *mel/sim* hybrids, as expected based on their longer divergence time. We found that most hybrid inviability alleles in *mel/san* hybrids are deleterious at both temperatures. Our results strongly indicate that the penetrance of these incompatibilities is independent from temperature (at least at the two assessed temperatures). Yet, there are alleles that cause inviability only at 18ºC or only at 24ºC, indicating that postzygotic isolation in this cross can still be affected by extrinsic factors. In *mel/sim* hybrids we found the opposite pattern; the magnitude of hybrid inviability is strongly dependent on the temperature at which the hybrids are raised.

Table 2.2 Pairwise comparisons (Tukey HSD test) from a linear model show that in *mel/san* hybrids only chromosome two is marginally affected by temperature, and the effect size is modest

Linear hypothesis	Mean 1	Mean 2	Estimate	Standard error	t-value	Pr(>|t|)
X.*sim*.24ºC - *X*.sim.18ºC == 0	0.4975	0.4465	0.0402	0.0286	1.407	0.9612
2.*sim*.24ºC - 2.sim.18ºC == 0	0.5278	0.4964	0.0314	0.0206	1.524	0.9323
3.*sim*.24ºC - 3.*sim*.18ºC == 0	0.4986	0.4917	0.0068	0.0244	0.279	1.0000
X.*san*.24ºC - *X*.*san*.18ºC == 0	0.4014	0.5377	0.0451	0.0286	1.579	0.9144
2.*san*.24ºC - 2.*san*.18ºC == 0	0.4430	0.5027	-0.0598	0.0206	-2.901	0.1399
3.*san*.24ºC - 3.*san*.18ºC == 0	0.4113	0.4229	-0.0116	0.0244	-0.475	1.0000

The first column shows the pairwise comparisons (Chromosome.Species.Temperature). Mean 1 refers to the mean of the first category listed in the comparison; Mean 2 refers to the mean of the second category. The effect of temperature was not significant in either *mel/sim* or *mel/san* hybrids.

Only one of the identified loci causes hybrid inviability at both 18ºC and 24ºC, indicating that different sets of loci affect HI at different temperatures. Given these data, our initial hypothesis that *D. santomea*'s temperature specialization would cause *mel/san* hybrids to be more affected by temperature than *mel/sim* hybrids is unlikely to be correct. If temperature had a strong effect on hybrid incompatibilities in *mel/san* hybrids, we would expect to see far more temperature-dependent hybrid incompatibilities than temperature-independent hybrid incompatibilities, a pattern we do not observe.

When we evaluated the mean effect size of *D. santomea* recessive hybrid inviability alleles in *mel/san* hybrids, we found that the mean viability of *df*-carrying hybrids is similar

at 24ºC and at 18ºC, a somewhat surprising result. We expected that *mel/san* hybrids would be more temperature sensitive due to the narrow temperature range inhabited by *D. santomea* (Matute et al. 2009a, Turissini et al. 2016). The mean magnitude of hybrid inviability in *mel/sim* hybrids, and unlike the pattern observed in *mel/san* hybrids, is contingent on temperature, and *df*-carrying hybrids do better at 24ºC than at 18ºC. This result is surprising because, unlike *D. santomea, D. simulans* is a widely cosmopolitan species that is able to breed at a range of temperatures (Austin and Moehring 2013), and we expected that *D. simulans* hybrids would be less affected by temperature.

A possible explanation for this pattern is that hybrids between highly divergent species (*mel/san*) are less likely to be affected by temperature because their genomes contain a larger number of loci with potentially deleterious interactions (Orr 1995, Matute et al. 2010, Moyle and Nakazato 2010, Wang et al. 2015). Alternatively, increased divergence time between species likely leads to an increase in the number of loci involved in HI, which might be expected to lead to reduced temperature sensitivity. In such cases, the penetrance of hybrid incompatibilities might be less likely to be affected by environmental factors due to the very large number of deleterious interactions. Even with a moderate reduction in the number of interactions at a lower temperature, many other deleterious interactions will remain and cause hybrid inviability. Conversely, more recently diverged species will have fewer deleterious interactions, and so may be more strongly affected by temperature as each single interaction plays a larger role in hybrid inviability. It is also possible that *D. santomea*'s temperature specialization has resulted in lower variability among alleles involved in thermal preference/thermal tolerance. This may result in lower variability of outcomes between temperatures because each allele has similar fitness at each temperature.

Temperature dependent HI alleles in *mel/sim* hybrids could hypothetically serve as an intermediate state for gene flow between populations, allowing successful production of progeny under only certain conditions. This is an unlikely explanation, however, as hybrids between *D. melanogaster* and *D. simulans* or *D. santomea* are inviable, or sterile, and have never been observed in nature.

Our results have one caveat. We cannot address whether the penetrance of alleles involved in hybrid incompatibility is more, or less pronounced in interspecific hybrids from parents with a restricted thermal niche than in interspecific hybrids with a wide thermal niche. Our experiment does not allow us to disentangle the effects of genetic distance between hybrids and the identity of the examined species. An ideal test would involve comparing the penetrance of recessive hybrid inviability alleles between pairs of hybrids whose parents have roughly equivalent genetic distances. To study highly divergent hybrids, one could study hybrids between *D. melanogaster* with *D. santomea* and between *D. melanogaster* with *D. yakuba*. Since *D. santomea* and *D. yakuba* are sister species, their levels of divergence from *D. melanogaster* are roughly equivalent (Turissini et al. 2016). However, multiple attempts to hybridize *D. melanogaster* and *D. yakuba* have failed and when hybridization has succeeded, the protocol is onerous and unlikely to be applicable to a genome-wide mapping approach (Sanchez and Santamaria 1997). Another possibility is to compare the viability of *mel/san* hybrids with hybrids between *D. melanogaster* and *D. teissieri*, which is related to *D. santomea* and *D. yakuba*. A second set of potentially informative crosses would be *D. melanogaster* with *D. simulans* and *D. melanogaster* with *D. sechellia* (or *D. mauritiana*). The *simulans/sechellia/mauritiana* triad might also be useful

to assess whether there are interactions between the mitochondrial and endosymbiont genomes, the nuclear genome, and the temperature at which the hybrids are raised.

Several studies have found that hybrid defects are more common and more severe at high temperatures. The genetic underpinnings of such interactions remain unknown, although potential explanations have included differences in molecular kinetics and a high correlation of thermal tolerance alleles with DMIs. Our finding that *df*-carrying *mel/sim* hybrids have higher overall viability at 24ºC than 18ºC is surprising. Previous work examining the interaction of temperature and hybrid viability has found that hybrid viability decreases at higher temperatures (Koevoets et al. 2012), disagreeing with our finding. In the case of *df*-carrying *mel/san* hybrids, the influence of temperature on viability was negligible. This suggests that at the very least, the interaction of temperature and HI alleles is complex and likely varies depending on the species pair.

One possible explanation for this unexpected result is temperature dependent haploinsufficiency. *df*-carrying hybrids have only a single copy of each gene located within the particular deficiency they carry. Though these regions are known not to cause haploinsufficiency when hemizygous in the parental species, it is unknown if these regions will be haploinsufficient in *mel/sim* hybrids but not in *mel/san* hybrids. The single copy of the gene product at these loci may be sufficient in the hybrids when reared at 24ºC but suffer too great a loss of function and become insufficient due to the reduced kinetics at 18ºC. These loci may become haploinsufficient in the hybrids at the lower temperature due to reduced function of the gene product at 18ºC. In this case, haploinsufficiency would be contingent on genetic background (i.e., the identity of the hybrid), suggesting species-specific epistatic interactions and not generalized haploinsufficiency. It is also possible that

these regions harbor temperature dependent recessive lethal alleles segregating naturally in *D. simulans*, but not *D. santomea*, although this is an unlikely explanation for the observed pattern, as both species are capable of breeding at 18ºC (Matute et al. 2009b, Austin and Moehring 2013). If 18ºC-dependent recessive lethal variants segregated naturally in these species at frequencies high enough to be detected by our mapping, we would expect to see substantial reductions in fitness in these species when reared at 18ºC.

Our results show that temperature can play a significant role in the penetrance of hybrid inviability, and that the effect of temperature varies depending on the species pair. Overall, our results and those from similar reports suggest that we should not think of hybrid inviability solely as the product of genetic interactions in the hybrid offspring, but rather must consider the phenomenon of hybrid inviability within the broader environmental and organismal context in which it is observed.

Materials and Methods

We crossed *D. melanogaster* females carrying a chromosomal deficiency with either *D. simulans* or *D. santomea* males in order to map recessive hybrid incompatibility alleles. Larvae were reared at 18ºC. We compared the results of our mapping with a previous study that identified hybrid inviability loci at 24ºC for these two species pairs. We describe each step as follows.

Species and stocks

We used one outbred stock for *D. santomea* and one for *D. simulans*. These stocks were generated by combining males and females from multiple isofemale lines. The *D. santomea* stock SYN2005 was generated by mixing 6 isofemale lines collected in the highlands of São Tomé. *Drosophila simulans* FC was created by J. Coyne and has been previously reported (Coyne et al. 1998, Matute and Gavin-Smyth 2014). All lines were

reared on standard cornmeal/Karo/agar medium at 24ºC under a 12 h light/dark cycle in 100mL bottles. Adults were allowed to oviposit for one week, after which time the bottles were cleared. We added 1mL of propionic acid (0.5% V/V) solution to the vials and provided a pupation substrate (Kimberly Clark, Kimwipes Delicate Task; Irving, TX). At least 10 bottles of each species were kept in parallel to guarantee the collection of large numbers of virgins.

Drosophila melanogaster deficiency stocks were purchased from Bloomington stock center in five batches, one for each chromosomal arm. Once quarantined, stocks were expanded in 200mL plastic bottles containing cornmeal food. We let females oviposit; when larvae were observed in the bottles, they were monitored daily for black pupae. All flies were kept at 24ºC under a 12 hour light/dark cycle. Table S1 lists all the stocks used in this report.

Virgin collection

To cross *D. melanogaster* deficiency stocks to male *D. santomea* or *D. simulans*, we needed virgin females from each *D. melanogaster* mutant stock. We kept *D. melanogaster* deficiency stocks in 300mL plastic bottles with cornmeal fly food. Once dark pupae were observed, bottles were cleared every 12 hours. Females from these mutant stocks were collected as virgins within 8 hours of eclosion under CO_2 anesthesia and kept for three days in single-sex groups of 20 flies in 30mL, corn meal food-containing vials. Males were also collected daily from kimwiped bottles but were not necessarily virgins. They were kept in all-male vials (20 individuals per vial). On day four, we assessed whether there were larvae in the media in both the female and male vials. If the inspection revealed any progeny, the vial was discarded. If the vials had no larvae, the virgin individuals were used for crosses.

Deficiency mapping

We used deficiency mapping to detect recessive alleles from the *D. santomea* genome involved in hybrid inviability (HI) (Coyne et al. 1998, Presgraves 2003). Our crossing design detects recessive partners of a DMI in a species crossable with *D. melanogaster* by uncovering recessive deleterious alleles with null alleles of a genomic region from *D. melanogaster*. The approach involves crossing females from *Drosophila melanogaster* (*mel*) stocks containing known genomic deletions, or "deficiencies" (*df*, Bloomington Drosophila Fly Stock Center), maintained as heterozygotes against a balancer (*Bal*) chromosome carrying a dominant homozygous lethal mutation, to *D. santomea* (*san*) males (Figure 1). On day 4 after virgin collection, males and females were mixed in a 30mL plastic vial with cornmeal fly food. The ratio of females to males was always 1:2 and at least 10 females were used per cross. To maximize the lifespan of flies, we maintained all crosses with the vial laying on its side for the duration of the assay. Vials were inspected every five days to check for progeny. We transferred the parents to a new vial when we observed either larvae or dead embryos. The old vial was tended by dampening the media with propionic acid and adding tissue paper (Kimwipes, Kimtech Science) for the larvae to pupate upon. We performed at least 20 replicates per cross and on average 10 of them produced progeny. Crosses were kept until no more progeny were produced from each vial.

Assessment of hybrid inviability

We measured the effect of each hemizygous region (those expressing *san or sim* recessive alleles) on the viability of hybrid female offspring (Figure 2). If a *D. melanogaster* deficiency uncovered a completely lethal recessive region of the *D. santome*a genome (one which caused lethality in F_1 hybrids), this cross would produce *Bal/san* but not *df/san* hybrid females (Figure 1; Coyne et al. 1998, Matute et al. 2010). If the *D. melanogaster* deficiency

uncovered a recessive region of the *D. santomea* genome that compromised hybrid fitness but did not cause complete lethality, then, this cross would produce an excess of *Bal/san* compared to *df/san* hybrid females (as assessed by a χ^2 test, 1 degree of freedom). Cases in which *Bal/san* hybrids are significantly more common than *df/san* hybrids indicate epistatic interactions between a recessive *san* allele (exposed when hemizygous) and a dominant factor in the *mel* genome (Coyne et al. 1998). This allowed us to measure hybrid inviability quantitatively instead of as a binary trait. All crosses were kept at 18ºC once started.

Counting hybrid inviability alleles

The minimal number of HI alleles was determined by counting the number of overlapping deficiencies associated with HI. If two deficiencies overlap and both cause HI, it can be assumed that they share a locus involved in HI. To assess whether the density of hybrid incompatibility alleles was uniform across chromosomal arms, we compared the observed number of hybrid incompatibility alleles with the expectations from a uniform distribution (i.e., same number of hybrid incompatibilities alleles in the five chromosomal arms) using Pearson's χ^2 test with simulated p-values (based on 2,000 replicates, library 'stats'; R 2014).

Effect of temperature

To assess whether temperature affected the viability of different hybrid genotypes, we measured HI at 18ºC and compared it with the magnitude of HI at 24ºC (data for HI at 24ºC were previously published in Matute et al. 2010). In order to minimize the effect of different genetic backgrounds, we only compared hybrid inviability between stocks that had been evaluated at both temperatures and in both hybrid crosses. First, we compared the mean viability of *df*-carrying hybrid individuals of each genotype (i.e., deficiency) at the two

temperatures using paired t-tests. We did two tests, one for each interspecific cross (R, library 'stats'; R Core Team 2014).

Next, we fit two linear models to the data in order to analyze the interaction between hybrid genotype, deficiency, and temperature. First, we fit a linear model in which the viability of the *df*-carrying genotype was the response; the temperature (18ºC and 24ºC) and the hybrid genotype (*mel/sim* and *mel/san*) were fixed effects. We also included the interaction between temperature and hybrid genotype. The linear model followed the form:

$$\boldsymbol{viab(df)_i \sim temp_i + genotype_j + temp \times genotype_{ij} + Error_{ij}}$$

The linear model was fit with the function 'lm' (R, library 'stats'; R Core Team 2014). Pairwise posthoc comparisons were done with a Tukey Honest Significant Difference (HSD) test using the function 'glht' (R, library 'multcomp'; Hothorn et al. 2008). We also fit a linear model that included only the interaction between temperature and genotype:

$$\boldsymbol{viab(df)_i \sim temp \times genotype_{ij} + Error_{ij}}$$

Effect of chromosome

We fit a linear model to test whether the sex chromosomes and autosomes had different effects on HI. Since we were only interested in assessing whether temperature affected the fitness of *df*-carrying hybrids differently in the two hybrids, the linear model had three fixed effects: location (chromosome), temperature (18ºC, and 24ºC), and genotype (*mel/san* and *mel/sim*). The model also included all possible interactions between the effects:

$$\boldsymbol{viab(df)_i \sim chromosome_i + temperature_j + genotype_k} \\ \boldsymbol{+ chromosome \times temperature_{ij} + chromosome \times genotype_{ik}} \\ \boldsymbol{+ temperature \times genotype_{jk}} \\ \boldsymbol{+ chromosome \times temperature \times genotype_{ijk} + Error_{ijk}}$$

The model was fit with the function 'lm' (R, library 'stats'; R Core Team 2014). Pairwise comparisons were done with a Tukey Honest Significant Difference test using the function 'glht' (R, library multcomp; Hothorn et al. 2008) in a linear model that only included the interaction effect between deficiency location, temperature, and genotype:

$$\boldsymbol{viab(df)_i \sim chromosome \times temperature \times genotype_{ijk} + Error_{ij}k}$$

Supplemental Data

Supplemental Table 2.1 Mutant strains used in this report

The 'Stock' column refers to the number of the mutant in Flybase. The table also includes all the raw data (counts) for the two types of interspecific crosses.

Stock	cytology	*Bal/san* progeny at 18ºC	*df/san* progeny at 18ºC	Proportion *df/san* 18ºC	*Bal/sim* progeny at 18ºC	*df/sim* progeny at 18ºC	Proportion *df/sim* 18ºC
1329	1A1--2A	76	31	0.289719626	98	104	0.514851485
25058	1A5--1B12	37	0	0	78	104	0.571428571
25062	1D1--2A3	206	115	0.358255452	134	156	0.537931034
9054	2E1--3A2	107	153	0.588461538	204	231	0.531034483
26569	2F2--3A4	181	325	0.64229249	55	47	0.460784314
935	2F6--3C5	266	270	0.504672897	29	21	0.42
8031	3A3--3A8	101	160	0.61302682	174	199	0.533512064
9348	3A8--3B1	85	65	0.433333333	205	234	0.533029613
8948	3B1--3C5	145	0	0	99	111	0.528571429
939	3C11--3E4	250	180	0.418604651	104	131	0.557446809
729	3C1--3D6	240	105	0.304347826	209	190	0.476190476
944	4C15--5A2	65	55	0.458333333	202	188	0.482051282
945	5A8--5C6	180	170	0.485714286	196	171	0.465940054
26506	5B6--5D2	104	48	0.315789474	174	156	0.472727273
946	5C2--5D6	168	151	0.473354232	97	78	0.445714286
8947	5C7--5F3	108	63	0.368421053	154	178	0.536144578
7713	5F2--6B2	347	331	0.48820059	185	199	0.518229167
7714	6B2--6C4	70	71	0.503546099	167	200	0.544959128
23670	6C11--6D3	56	41	0.422680412	172	142	0.452229299
9625	6C12--6D8	155	185	0.544117647	132	111	0.456790123

25063	6C2--6C8	80	50	0.384615385	145	167	0.53525641
3196	6E2--7A6	330	325	0.496183206	65	82	0.557823129
947	6E4--7A6	135	145	0.517857143	198	206	0.50990099
8955	7A3--7B2	220	221	0.501133787	200	225	0.529411765
3221	7B2-7C4	55	50	0.476190476	75	91	0.548192771
949	7D1--7D6	56	0	0	205	183	0.471649485
950	7D10--8A5	159	115	0.419708029	165	195	0.541666667
951	7F1--8C6	44	0	0	145	152	0.511784512
3651	8B5--8D9	121	96	0.442396313	97	109	0.529126214
952	8E--9D	543	314	0.366394399	142	104	0.422764228
954	9B1-2;10A1-2	81	95	0.539772727	185	205	0.525641026
7339	9D5-9E8	492	369	0.428571429	164	199	0.548209366
5707	9E3--10A8	467	414	0.469920545	79	96	0.548571429
26556	9E4--9F12	61	77	0.557971014	561	501	0.471751412
25068	9E8--10A3	274	179	0.395143488	311	267	0.461937716
962	10F7--11D1	96	49	0.337931034	253	271	0.517175573
967	11D--12A2	272	448	0.622222222	274	290	0.514184397
727	12A3--12E9	34	102	0.75	185	177	0.488950276
998	12D2--13A5	12	14	0.538461538	184	184	0.5
3347	13F1--14B1	69	105	0.603448276	105	98	0.482758621
125	14B8--14C1	145	133	0.478417266	57	78	0.577777778
991	14F6--15A6	172	254	0.596244131	289	251	0.464814815
25416	15A1--15E2	193	133	0.40797546	161	145	0.473856209
4741	15D3--16A6	331	288	0.465266559	400	310	0.436619718
4953	16A2--16C10	306	271	0.469670711	203	172	0.458666667
970	17A1--18A2	77	61	0.442028986	48	32	0.4

7754	18A2--18A3	64	107	0.625730994	262	251	0.489278752
971	18A5--18D	70	0	0	145	131	0.474637681
7721	18D13--18F2	117	46	0.282208589	172	145	0.457413249
972	18E1--20F	88	1	0.011235955	423	425	0.501179245
977	19F1--20F	156	1	0.006369427	461	515	0.527663934
3714	20A--20F	76	2	0.025641026	304	165	0.351812367
3638	21A1--21B8	247	241	0.493852459	298	256	0.462093863
8672	21B7--21C2	244	247	0.50305499	318	341	0.517450683
6283	21B7--21C3	143	133	0.481884058	68	79	0.537414966
6608	21C3--21C8	36	23	0.389830508	451	402	0.471277843
3084	21D1--22B3	161	106	0.397003745	382	311	0.448773449
24120	22B2--22D4	227	166	0.422391858	288	144	0.333333333
7144	22D2--22F2	139	222	0.614958449	167	155	0.48136646
90	22F4--23C4	111	108	0.493150685	211	208	0.496420048
1567	23C1--23E2	71	128	0.64321608	49	61	0.554545455
6875	23C5--23E2	661	721	0.52170767	178	156	0.467065868
6965	23E5--23F5	151	176	0.5382263	291	283	0.493031359
6507	23F3--24A2	94	105	0.527638191	105	100	0.487804878
5330	24A2--24D4	104	93	0.472081218	146	101	0.408906883
693	24C2--25C9	191	313	0.621031746	56	45	0.445544554
9270	24F4--25A7	271	316	0.538330494	94	111	0.541463415
8835	25C1--25C4	456	311	0.40547588	299	282	0.485370052
8674	25C4--25C8	329	287	0.465909091	415	471	0.531602709
7497	25C8--25D5	163	168	0.50755287	261	243	0.482142857
781	25D2--26B5	251	438	0.635703919	149	167	0.528481013
490	25F3--26D11	116	111	0.488986784	45	54	0.545454545

6299	26B1--26D2	487	496	0.504577823	167	201	0.546195652
6374	26D10--27C1	179	197	0.52393617	188	165	0.467422096
2414	27C1--28A	131	398	0.752362949	199	183	0.479057592
5420	27C2--27C5	101	113	0.528037383	74	99	0.572254335
4956	27E2--28D1	145	356	0.710578842	153	203	0.570224719
7147	28A4--28D9	146	165	0.530546624	205	204	0.498777506
140	28D2--28E5	301	657	0.685803758	106	121	0.533039648
179	28E4--29C1	62	88	0.586666667	413	231	0.358695652
8836	28F5--29B1	141	213	0.601694915	197	209	0.514778325
9298	29B4--29C3	171	146	0.460567823	164	151	0.479365079
2892	29C1--30C9	202	241	0.544018059	205	231	0.529816514
6478	30C3--30F1	184	187	0.504043127	95	76	0.444444444
1045	30D--31F	138	153	0.525773196	341	267	0.439144737
8469	30F5--31B1	701	318	0.312070658	410	541	0.568874869
3366	31B1--32A2	166	111	0.400722022	51	66	0.564102564
9503	31B1--31D9	201	231	0.534722222	67	81	0.547297297
7142	32A1--32D1	114	92	0.446601942	188	154	0.450292398
9505	32C1--32C1	245	413	0.627659574	212	200	0.485436893
7143	32D1--32E1	331	528	0.614668219	431	398	0.480096502
5869	32D1--32F3	186	306	0.62195122	291	187	0.391213389
3079	32F1--33F2	72	51	0.414634146	174	200	0.534759358
6999	34A3--34B9	539	207	0.277479893	560	501	0.472196041
9594	34B4--34C4	166	118	0.415492958	18	31	0.632653061
9506	34C1--34C6	171	153	0.472222222	24	24	0.5
3588	35B4--35F7	46	0	0	67	66	0.496240602
1491	35D1--36A7	192	221	0.535108959	204	201	0.496296296

420	36C2--37C1	34	27	0.442622951	56	52	0.481481481
567	37B2--38D5	183	123	0.401960784	56	67	0.544715447
167	38A6--40B1	96	48	0.333333333	186	191	0.5066313
7531	40A5--40D3	321	451	0.584196891	203	201	0.497524752
9510	40A5--40E5	172	11	0.06010929	205	191	0.482323232
4959	h35--40A1	124	7	0.053435115	56	67	0.544715447
749	h44--42A2	103	58	0.360248447	41	51	0.554347826
1888	42B3--43E18	81	58	0.417266187	34	23	0.403508772
3368	42E--44C	1	21	0.954545455	178	209	0.54005168
198	43F--44D8	141	116	0.451361868	94	100	0.515463918
201	44D1--44F12	99	87	0.467741935	204	183	0.472868217
3591	44F10--45E1	203	198	0.493765586	202	204	0.502463054
4966	45A6--45E3	56	65	0.537190083	147	165	0.528846154
6917	45D3--45F6	190	211	0.526184539	512	493	0.490547264
9410	45F6--46B4	177	71	0.286290323	204	216	0.514285714
1743	46A--46C	129	62	0.32460733	22	19	0.463414634
1702	46C--47A1	71	0	0	104	86	0.452631579
190	47D3--48B2	266	123	0.316195373	205	174	0.459102902
1145	48A3--48C8	56	116	0.674418605	319	301	0.485483871
7145	48C5--48E1	145	121	0.454887218	238	198	0.45412844
7146	48E1--48E10	154	478	0.756329114	58	56	0.49122807
5879	48E12--49B6	48	36	0.428571429	194	199	0.506361323
754	49B2--49E2	71	48	0.403361345	204	203	0.498771499
442	49C1--50D1	116	101	0.465437788	198	402	0.67
6516	50D1--50D7	92	116	0.557692308	311	341	0.523006135
7875	50D4--50E4	24	32	0.571428571	287	302	0.512733447

9496	50E1--50E6	73	58	0.442748092	195	211	0.519704433
7876	50E4--50F6	136	118	0.464566929	174	209	0.545691906
6455	50E6--51E4	156	1	0.006369427	174	151	0.464615385
3518	51D3-52F9	7	176	0.961748634	188	154	0.450292398
3520	52F5--53A1	44	45	0.505617978	173	142	0.450793651
25078	53C1--53C6	248	421	0.629297459	47	54	0.534653465
7445	53D9--54B10	61	41	0.401960784	69	67	0.492647059
7414	54B1--54B10	59	201	0.773076923	419	411	0.495180723
5574	54B16--54B16	245	321	0.567137809	242	230	0.487288136
5680	54B17--C4	211	217	0.507009346	39	42	0.518518519
9596	54B2--54B17	82	78	0.4875	211	246	0.538293217
6780	54E5--55B7	236	463	0.662374821	252	211	0.455723542
1547	55A--55F	41	36	0.467532468	241	267	0.525590551
757	55E2--56C11	27	256	0.90459364	71	78	0.523489933
6866	56C4--56D10	133	643	0.828608247	221	211	0.488425926
6647	56D7--56F12	67	58	0.464	189	195	0.5078125
7896	56F11--56F16	176	223	0.558897243	167	145	0.46474359
3467	56F9--57D12	71	1	0.013888889	154	141	0.477966102
5246	57D2--58D1	166	34	0.17	154	180	0.538922156
282	58D1--59A	29	262	0.900343643	41	32	0.438356164
3909	59A1--59D4	258	273	0.514124294	56	78	0.582089552
7273	59B--59E1	121	398	0.766859345	78	75	0.490196078
1682	59D5--60B8	11	91	0.892156863	104	111	0.51627907
9691	60B8--60C4	122	461	0.790737564	57	45	0.441176471
2604	60C5--60D10	78	91	0.538461538	82	73	0.470967742
9069	60C8--60E8	123	31	0.201298701	41	30	0.422535211

2471	60E2--60E12	76	58	0.432835821	194	160	0.451977401
4961	60F1--60F5	19	76	0.8	120	111	0.480519481
2577	61A--61D3	162	161	0.498452012	314	301	0.489430894
439	61C5--62A8	129	102	0.441558442	151	132	0.466431095
600	62A10--62D5	108	103	0.488151659	242	231	0.488372093
9693	62A11--62B7	104	38	0.267605634	101	89	0.468421053
2400	62B4--62E5	111	86	0.436548223	103	74	0.418079096
6755	62E8--63B6	47	1	0.020833333	204	151	0.425352113
3650	62F--63B10	179	162	0.475073314	138	135	0.494505495
3649	63C2--63F7	67	89	0.570512821	99	138	0.582278481
463	63E6--64A10	75	108	0.590163934	56	71	0.559055118
3686	63F6--64C15	92	92	0.5	104	92	0.469387755
3096	64C--65C	40	52	0.565217391	156	204	0.566666667
4393	65A2;65E1	133	90	0.403587444	204	200	0.495049505
6867	65D4--65E6	100	92	0.479166667	89	142	0.614718615
6964	65E10--65F6	236	192	0.448598131	203	190	0.48346056
1420	65F3;66B10	160	91	0.362549801	138	165	0.544554455
5877	66A17-20;66C1-5	55	80	0.592592593	104	11	0.095652174
6460	66B12-C1;66D2-4	49	55	0.528846154	202	198	0.495
1541	66B8-9;66C9-10	194	170	0.467032967	243	202	0.453932584
3024	66D10--66E2	51	0	0	105	111	0.513888889
4500	66E1-6;66F1-6	55	45	0.45	67	87	0.564935065
7079	66F1-2;67B2-3	92	102	0.525773196	88	96	0.52173913
997	67A2;67D7-13 or 67A5;67D9-13	41	52	0.559139785	92	79	0.461988304
23668	67C7;67D5	21	33	0.611111111	96	90	0.483870968
2612	68C8--69B5	115	40	0.258064516	92	103	0.528205128

5492	69A4--69D6	120	99	0.452054795	104	88	0.458333333
6456	69D4-5;69F5-7	52	42	0.446808511	205	194	0.486215539
3124	70C1-2;70D4-5	320	288	0.473684211	99	78	0.440677966
3126	70D2-3;71E4-5	216	311	0.590132827	104	100	0.490196078
6551	71C2-3;72B1-C1	57	82	0.589928058	203	222	0.522352941
2993	72C1-D1;73A3-4	154	98	0.388888889	311	278	0.471986418
6411	74D3-75A1;75B2-5	136	112	0.451612903	75	100	0.571428571
2608	75A6-7;75C1-2	36	21	0.368421053	84	92	0.522727273
6754	75F10--76A5	89	181	0.67037037	99	99	0.5
8082	75F2;76A1	124	133	0.517509728	75	77	0.506578947
3617	76B1-2;76D5	84	72	0.461538462	300	277	0.480069324
5126	76B4;77B	36	0	0	218	211	0.491841492
2052	77A1;77D1	33	0	0	123	141	0.534090909
3127	77B-C;77F-78A	97	101	0.51010101	35	51	0.593023256
4429	77F3;78C8-9	78	89	0.532934132	78	98	0.556818182
4430	78C5--79A1	211	180	0.460358056	402	378	0.484615385
1990	83C1--84B2	136	79	0.36744186	250	222	0.470338983
1962	85A2;85C1-2	64	51	0.443478261	39	54	0.580645161
1931	85D8-12;85E7-F1	4	0	0	85	77	0.475308642
7080	85F1-2;86C7-8	476	64	0.118518519	84	98	0.538461538
3128	86C1;87B1-5	264	172	0.394495413	101	89	0.468421053
3003	86E2-4;87C6-7	124	12	0.088235294	203	199	0.495024876
383	88E7-13;89A1	128	116	0.475409836	241	200	0.453514739
756	88F9-89A1;89B9-10	63	34	0.350515464	142	111	0.438735178
1920	89B5;89C	171	150	0.46728972	203	189	0.482142857
1467	89B7--89E7	29	41	0.585714286	222	259	0.538461538

4431	89E--91B2	85	0	0	89	99	0.526595745
3011	90F1-F4;91F5	49	64	0.566371681	54	42	0.4375
3012	91F1-2;92D3-6	206	171	0.453580902	99	111	0.528571429
4962	92B3;92F13	271	160	0.371229698	104	100	0.490196078
7413	92F7-93A1;93B3-6	248	151	0.378446115	204	191	0.483544304
26529	93D1;93F14	0	2	1	220	199	0.474940334
2586	94A3-4;94D1-4	28	20	0.416666667	231	241	0.51059322
7674	95A4;95B1	199	107	0.349673203	142	161	0.531353135
2585	95A5-7;95D6-11	331	201	0.377819549	200	188	0.484536082
4432	95D7-D11;95F15	75	107	0.587912088	312	299	0.489361702
2363	95F7--96A18	270	255	0.485714286	86	100	0.537634409
3468	96A2-7;96D2-4	127	121	0.487903226	93	87	0.483333333
5601	96F1;97B1	292	311	0.515754561	111	89	0.445
1910	97A--98A2	196	227	0.536643026	142	160	0.529801325
823	97E3--98A5	41	0	0	203	187	0.479487179
9529	97F1-2;98A	48	43	0.472527473	163	142	0.46557377
7412	98B1-2;98B3-5	181	223	0.551980198	153	138	0.474226804
430	98E3;99A6-8	133	151	0.531690141	142	111	0.438735178
669	99A1-2;99B6-11	55	79	0.589552239	89	74	0.45398773
3547	99B5-6;99E4-F1, 98F;100F	56	48	0.461538462	94	83	0.468926554
3546	99C8;100F5	96	16	0.142857143	102	81	0.442622951

Supplemental Table 2.2 Genome distribution of recessive *Drosophila santomea* hybrid incompatibilities in *mel/san* hybrids

We only scored the effect of deficiencies in five major Muller elements and excluded the dot-chromosome.

Chromosome arm	Overall	Total 24ºC	Unique 24ºC	Total 18ºC	Unique 18ºC	Total both	Unique both
X	28	2	2	10	8	16	11
2L	32	9	9	11	8	12	8
2R	27	10	8	6	5	11	10
3L	19	6	5	6	6	7	6
3R	19	7	7	2	2	10	8
Total	125	34	31	35	29	56	43

Supplemental Table 2.3 Distribution of recessive *Drosophila simulans* hybrid incompatibilities in *mel/sim* hybrids

We only scored the effect of deficiencies in five major Muller elements and excluded the dot-chromosome.elements and excluded the dot-chromosome.

Chromosome arm	Overall	Total 24ºC	Unique 24ºC	Total 18ºC	Unique 18ºC	Total both	Unique both
X	6	2	2	3	2	1	1
2L	5	5	5	0	0	0	0
2R	0	0	0	0	0	0	0
3L	6	3	3	3	3	0	0
3R	6	6	4	0	0	0	0
Total	23	16	14	6	5	1	1

Supplemental Table 2.4 Pairwise comparisons of relative viability of deficiency stocks at different temperatures in *mel/sim* and *mel/san* hybrid F1 females

Bolded cells show pairwise comparisons also shown in Table 2.2.

Linear hypotheses	Estimate	Std. Error	t value	Pr(>|t|)
sim.18.2 - san.18.2 == 0	-0.006276	0.020595	-0.305	1
san.24.2 - san.18.2 == 0	**-0.059755**	**0.020595**	**-2.901**	**0.1399**
sim.24.2 - san.18.2 == 0	0.025103	0.020595	1.219	0.987
san.18.3 - san.18.2 == 0	-0.079837	0.0226	-3.533	0.0217
sim.18.3 - san.18.2 == 0	-0.010969	0.0226	-0.485	1
san.24.3 - san.18.2 == 0	-0.091446	0.0226	-4.046	<0.01
sim.24.3 - san.18.2 == 0	-0.004148	0.0226	-0.184	1
san.18.*X* - san.18.2 == 0	-0.101332	0.024898	-4.07	<0.01
sim.18.*X* - san.18.2 == 0	-0.005208	0.024898	-0.209	1
san.24.*X* - san.18.2 == 0	-0.056232	0.024898	-2.258	0.5015
sim.24.*X* - san.18.2 == 0	0.034969	0.024898	1.404	0.9616
san.24.2 - sim.18.2 == 0	-0.053479	0.020595	-2.597	0.2778
sim.24.2 - sim.18.2 == 0	**0.031378**	**0.020595**	**1.524**	**0.9323**
san.18.3 - sim.18.2 == 0	-0.073561	0.0226	-3.255	0.0525
sim.18.3 - sim.18.2 == 0	-0.004694	0.0226	-0.208	1
san.24.3 - sim.18.2 == 0	-0.085171	0.0226	-3.769	<0.01
sim.24.3 - sim.18.2 == 0	0.002128	0.0226	0.094	1
san.18.*X* - sim.18.2 == 0	-0.095056	0.024898	-3.818	<0.01
sim.18.*X* - sim.18.2 == 0	0.001068	0.024898	0.043	1
san.24.*X* - sim.18.2 == 0	-0.049957	0.024898	-2.006	0.6844
sim.24.*X* - sim.18.2 == 0	0.041244	0.024898	1.657	0.8849

sim.24.2 - san.24.2 == 0	0.084857	0.020595	4.12	<0.01
san.18.3 - san.24.2 == 0	-0.020082	0.0226	-0.889	0.9992
sim.18.3 - san.24.2 == 0	0.048786	0.0226	2.159	0.5755
san.24.3 - san.24.2 == 0	-0.031692	0.0226	-1.402	0.962
sim.24.3 - san.24.2 == 0	0.055607	0.0226	2.46	0.3618
san.18.*X* - san.24.2 == 0	-0.041577	0.024898	-1.67	0.879
sim.18.*X* - san.24.2 == 0	0.054547	0.024898	2.191	0.552
san.24.*X* - san.24.2 == 0	0.003522	0.024898	0.141	1
sim.24.*X* - san.24.2 == 0	0.094724	0.024898	3.804	<0.01
san.18.3 - sim.24.2 == 0	-0.10494	0.0226	-4.643	<0.01
sim.18.3 - sim.24.2 == 0	-0.036072	0.0226	-1.596	0.9084
san.24.3 - sim.24.2 == 0	-0.116549	0.0226	-5.157	<0.01
sim.24.3 - sim.24.2 == 0	-0.02925	0.0226	-1.294	0.9792
san.18.*X* - sim.24.2 == 0	-0.126435	0.024898	-5.078	<0.01
sim.18.*X* - sim.24.2 == 0	-0.03031	0.024898	-1.217	0.9872
san.24.*X* - sim.24.2 == 0	-0.081335	0.024898	-3.267	0.0499
sim.24.*X* - sim.24.2 == 0	0.009866	0.024898	0.396	1
sim.18.3 - san.18.3 == 0	0.068868	0.024442	2.818	0.1712
san.24.3 - san.18.3 == 0	**-0.011609**	**0.024442**	**-0.475**	**1**
sim.24.3 - san.18.3 == 0	0.075689	0.024442	3.097	0.0836
san.18.*X* - san.18.3 == 0	-0.021495	0.026581	-0.809	0.9997
sim.18.*X* - san.18.3 == 0	0.074629	0.026581	2.808	0.1754
san.24.*X* - san.18.3 == 0	0.023605	0.026581	0.888	0.9992
sim.24.*X* - san.18.3 == 0	0.114806	0.026581	4.319	<0.01
san.24.3 - sim.18.3 == 0	-0.080477	0.024442	-3.293	0.047

sim.24.3 - sim.18.3 == 0	**0.006821**	**0.024442**	**0.279**	**1**
san.18.*X* - sim.18.3 == 0	-0.090363	0.026581	-3.4	0.0332
sim.18.*X* - sim.18.3 == 0	0.005761	0.026581	0.217	1
san.24.*X* - sim.18.3 == 0	-0.045263	0.026581	-1.703	0.8639
sim.24.*X* - sim.18.3 == 0	0.045938	0.026581	1.728	0.8519
sim.24.3 - san.24.3 == 0	0.087299	0.024442	3.572	0.019
san.18.*X* - san.24.3 == 0	-0.009886	0.026581	-0.372	1
sim.18.*X* - san.24.3 == 0	0.086239	0.026581	3.244	0.0538
san.24.*X* - san.24.3 == 0	0.035214	0.026581	1.325	0.9751
sim.24.*X* - san.24.3 == 0	0.126415	0.026581	4.756	<0.01
san.18.*X* - sim.24.3 == 0	-0.097184	0.026581	-3.656	0.0137
sim.18.*X* - sim.24.3 == 0	-0.00106	0.026581	-0.04	1
san.24.*X* - sim.24.3 == 0	-0.052085	0.026581	-1.959	0.7161
sim.24.*X* - sim.24.3 == 0	0.039117	0.026581	1.472	0.9466
sim.18.*X* - san.18.*X* == 0	0.096124	0.02856	3.366	0.0379
san.24.*X* - san.18.*X* == 0	**0.0451**	**0.02856**	**1.579**	**0.9144**
sim.24.*X* - san.18.*X* == 0	0.136301	0.02856	4.772	<0.01
san.24.*X* - sim.18.*X* == 0	-0.051025	0.02856	-1.787	0.8218
sim.24.*X* - sim.18.*X* == 0	**0.040177**	**0.02856**	**1.407**	**0.9612**
sim.24.*X* - san.24.*X* == 0	0.091201	0.02856	3.193	0.063

REFERENCES

1. Aruna S., Flores A., Barbash D. A. 2009. Reduced Fertility of *Drosophila* melanogaster Hybrid male rescue (*Hmr*) Mutant Females Is Partially Complemented by *Hmr* Orthologs From Sibling Species. *Genetics, 181*, 1437–1450.

2. Austin C. J., Moehring A. J. 2013. Optimal Temperature Range of a Plastic Species, *Drosophila simulans*. *Journal of Animal Ecology, 82*, 663–672.

3. Barbash D. A., Roote J., Ashburner M. 2000. The *Drosophila melanogaster* Hybrid male rescue Gene Causes Inviability in Male and Female Species Hybrids. *Genetics*, *154*, 1747–1771.

4. Bordenstein S. R., O'Hara P. F., Werren J. H. 2001. *Wolbachia*-induced Incompatibility Precedes Other Hybrid Incompatibilities in *Nasonia*. *Nature, 409,* 707–710.

5. Capy P., Gibert P. 2004. *Drosophila melanogaster, Drosophila simulans*: So Similar Yet So Different. *Genetica, 120,* 5–16.

6. Carrington L. B., Lipkowitz J. R., Hoffmann A. A., Turelli M. 2011. A Re-Examination of *Wolbachia*-Induced Cytoplasmic Incompatibility in California *Drosophila simulans*. *PLoS One, 6*, 1–12.

7. Cattani M. V., Presgraves D. C. 2012. Incompatibility Between *X* Chromosome Factor and Pericentric Heterochromatic Region Causes Lethality in Hybrids between *Drosophila melanogaster* and Its Sibling Species. *Genetics, 191*, 549–559.

8. Cenzer M.L., 2016. Adaptation to an invasive host is driving the loss of a native ecotype. *Evolution*, 70, 2296-2307.

9. Comeault A. A., Flaxman S. M., Riesch R., Schwander T., Slate J. et al. 2015. Selection on a Genetic Polymorphism Counteracts Ecological Speciation in a Stick Insect. *Current Biology, 25*, 1975–1981.

10. Cooper J. C., Phadnis N. 2016. A Genomic Approach to Identify Hybrid Incompatibility Genes. *Fly, 10*, 142–148.

11. Coyne J. A., Orr H. A. 1989. Patterns of Speciation in *Drosophila*. *Evolution, 43*, 362–381.

12. Coyne J. A., Simeonidis S., Rooney P. 1998. Relative Paucity of Genes Causing Inviability in Hybrids Between *Drosophila melanogaster* and *D. simulans*. *Genetics*, *150*, 1091–1103.

13. Coyne J. A., Orr H. A. 2004. *Speciation*. Vol. 37. Sinauer Associates. Sunderland, MA.

14. Delph L. F., Demuth J. P. 2016. Haldane' s Rule: Genetic Bases and Their Empirical Support. *Journal Of Heredity*, 1–9.

15. Dobzhansky T., Holz A. M., Spassky B. 1942. Genetics of Natural Populations. VIII. Concealed Variability in the Second and the Fourth Chromosomes of *Drosophila pseudoobscura* and its Bearing on the Problem of Heterosis. *Genetics*, *27*, 463–490.

16. Dowling D. K., Abiega K. C., Arnqvist G. 2007. Temperature-Specific Outcomes of Cytoplasmic-Nuclear Interactions on Egg-to-Adult Development Time in Seed Beetles. *Evolution*, *61*, 194–201.

17. Edmands S. 2007. Between a Rock and a Hard Place: Evaluating the Relative Risks of Inbreeding and Outbreeding for Conservation and Management. *Molecular Ecology*, *16*, 463–475.

18. Hopkins R. 2013. Reinforcement in Plants. *New Phytologist*, *197*, 1095–1103.

19. Hothorn T., Bretz F., Westfall P. 2008. Simultaneous Inference in General Parametric Models. *Biometrical Journal*, *50*, 346–363.

20. Hudson E. J., Price T. D. 2014. Pervasive Reinforcement and the Role of Sexual Selection in Biological Speciation. *Journal of Heredity*, *105*, 821–833.

21. Hutter P. and Ashburner M. 1987. Genetic Rescue of Inviable Hybrids Between *Drosophila melanogaster* and its sibling species. *Nature*, *327*, 331-333.

22. Inoue Y., Watanabe T. K. 1979. Inversion Polymorphisms in Japanese Natural Populations of *Drosophila melanogaster*. *Japanese Journal of Genetics*, *54* 69–82.

23. Koevoets T., van de Zande L., Beukeboom L. W. 2012. Temperature Stress Increases Hybrid Incompatibilities in the Parasitic Wasp Genus *Nasonia*. *Journal of Evolutionary Biology*, *25*, 304–316.

24. Lachaise D., Harry M., Solignac M., Lemeunier F., Benassi V. et al.. 2000. Evolutionary Novelties in Islands: *Drosophila santomea*, a New melanogaster Sister Species from São Tomé. *Proceedings of the Royal Society of London*, *267*, 1487–1495.

25. Llopart A., Lachaise D., Coyne J. A. 2005a. An Anomalous Hybrid Zone in *Drosophila*. *Evolution*, *59*, 2602–2607.

26. Llopart A., Lachaise D., Coyne J. A. 2005b. Multilocus Analysis of Introgression Between Two Sympatric Sister Species of *Drosophila*: *Drosophila yakuba* and *D. santomea*. *Genetics*, *171*, 197–210.

27. Maheshwari S., Barbash D. A. 2011. The Genetics of Hybrid Incompatibilities. *Annual Review of Genetics*, *45*, 331–355.

28. Maheshwari S., Barbash D. A. 2012. An Indel Polymorphism in the Hybrid Incompatibility Gene Lethal Hybrid Rescue of *Drosophila* Is Functionally Relevant. *Genetics*, *192*, 683–691.

29. Masly J. P., Presgraves D. C. 2007. High-Resolution Genome-Wide Dissection of the Two Rules of Speciation in *Drosophila*. *PLoS Biology*, *5*, 1890–1898.

30. Matute D. R., Butler I. A., Coyne J. A. 2009a Little Effect of the *tan* Locus on Pigmentation in Female Hybrids between *Drosophila santomea* and *D. melanogaster*. *Cell*, *139*, 1180–1188.

31. Matute D. R., Novak C. J., Coyne J. A. 2009b. Temperature-Based Extrinsic Reproductive Isolation in Two Species of *Drosophila*. *Evolution*, *63*, 595–612

32. Matute D. R., Butler I. A., Turissini D. A., Coyne J. A. 2010. A Test of the Snowball Theory for the Rate of Evolution of Hybrid Incompatibilities. *Science*, *329*, 1518–1522.

33. Matute D. R., Gavin-Smyth J. 2014. Fine Mapping of Dominant X-Linked Incompatibility Alleles in *Drosophila* Hybrids. *PLoS Genetics*, *10*, 1–15.

34. Mendelson T. C. 2003. Sexual Isolation Evolves Faster than Hybrid Inviability in a Diverse and Sexually Dimorphic Genus of Fish (Percidae: *Etheostoma*). *Evolution*, *57*, 317–327.

35. Moyle L. C., Nakazato T. 2010. Hybrid Incompatibility "Snowballs" Between *Solanum* Species. *Science*, *329*, 1521–1524.

36. Muller H. J., Raffel D., Gershenson S. M., Prokofveya-Belgovskaya B. A. A. 1937. A Further Analysis of Loci in the So-Called "Inert Region" of the X Chromosome of *Drosophila*. *Genetics*, *22*, 87–93.

37. Muralidharan S., Box M. S., Sedivy E. L., Wigge P. A., Weigel D. et al. 2014. Different Mechanisms for Arabidopsis thaliana Hybrid Necrosis Cases Inferred from Temperature Responses. *Plant Biology*, *16*, 1033–1041.

38. Noor M. A. F., Feder J. L. 2006. Speciation Genetics: Evolving Approaches. *Nature Reviews: Genetics*, *7*, 851–861.

39. Nosil P., Schluter D. 2011. The Genes Underlying the Process of Speciation. *Trends in Ecology & Evolution*, *26*, 160–167.

40. Orr H. A. 1995. The Population Genetics of Speciation: the Evolution of Hybrid Incompatibilities. *Genetics*, *139*, 1805–1813.

41. Orr H. A., Madden L. D., Coyne J. A., Goodwin R., Hawley R. S. 1997. The Developmental Genetics of Hybrid Inviability: A Mitotic Defect in *Drosophila* Hybrids. *Genetics*, *145*, 1031–1040.

42. Orr H. A., Turelli M. 2001. The Evolution of Postzygotic Isolation: Accumulating Dobzhansky-Muller Incompatibilities. *Evolution*, *55*, 1085–1094.

43. Orr H. A., Masly J. P., Phadnis N. 2007. Speciation in *Drosophila*: From Phenotypes to Molecules. *Journal of Heredity*, *98*, 103–110.

44. Phadnis N., Baker E.P., Cooper J.C., Frizzell K., Hseih E. et al. 2015. An essential cell cycle regulation gene causes hybrid inviability in *Drosophila*. *Science*, *350*, 1552-1555.

45. Presgraves D. C. 2003. A Fine-Scale Genetic Analysis of Hybrid Incompatibilities in *Drosophila*. *Genetics*, *163*, 955–972.

46. Presgraves D. C., Balagopalan L., Abmayr S. M., Orr H. A. 2003. Adaptive Evolution Drives Divergence of a Hybrid Inviability Gene Between Two Species of *Drosophila*. *Nature*, *423*, 715–719.

47. R Foundation for Statistical Computing. 2014. R: A Languge and Environment for Statistical Computing.

48. Rabosky D. L., Matute D. R. 2013. Macroevolutionary Speciation Rates Are Decoupled from the Evolution of Intrinsic Reproductive Isolation in *Drosophila* and Birds. *PNAS*, *110*, 15354–15359.

49. Rosenblum B. E., Sarver B. A. J., Brown J. W., Roches S. D., Hardwick K. M. et al. 2012. Goldilocks Meets Santa Rosalia: An Ephemeral Speciation Model Explains Patterns of Diversification Across Time Scales. *Evolutionary Biology*, *39*, 255–261.

50. Sanchez L., Santamaria P. 1997. Reproductive Isolation and Morphogenetic Evolution in *Drosophila* Analyzed by Breakage of Ethological Barriers. *Genetics*, *147*: 231-242

51. Sawamura K. 2016. Genome-Wide Analyses of Hybrid Incompatibility in Drosophila. *Advanced Techniques in Biology and Medicine*, *4*, 10–12.

52. Servedio M. R., Noor M. A. F. 2003. The Role of Reinforcement in Speciation: Theory and Data. *Annual Review of Ecology, Evolution, and Systematics*, *34*, 339–364.

53. Sobel J.M., Chen G.F., Watt L.R., Schemske D.W. 2010. The biology of speciation. *Evolution*, *64*, 295-315.

54. Stanley S. M., Parsons P. A., Spence G. E., Weber L. 1980. Resistance of Species of the *Drosophila melanogaster* Subgroup to Environmental Extremes. *Australian Journal of Zoology*, *28*, 413–421.

55. Sturtevant H. A. 1920. Genetic Studies on *Drosophila simulans*. I. Introduction. Hybrids with *Drosophila melanogaster*. *Genetics*, *5*, 488–500.

56. Tang S., Presgraves D. C. 2009. Evolution of the *Drosophila* Nuclear Pore Complex Results in Multiple Hybrid Incompatibilities. *Science*, *323*, 779–782.

57. Turissini D. et al. 2016. *Drosophila* Hybrids Have Trouble Finding Food. *Submitted*.

58. Wade M. J., Johnson N. A. 1994. Reproductive Isolation Between Two Species of Flour Beetles, *Tribolium castaneum* and *T. freemani*: Variation Within and Among Geographical Populations of *T. castaneum*. *Heredity*, *72*, 155–162.

59. Wade M. J., Johnson N. A., Toquenaga Y. 1999. Temperature Effects and Genotype-by-Environment Interactions in Hybrids: Haldane' s Rule in Flour Beetles. *Evolution*, *53*, 855–865.

60. Wang R. J., White M. A., Payseur B. A. 2015. The Pace of Hybrid Incompatibility Evolution in House Mice. *Genetics*, *201*, 229–242.

61. Yamada T., Marubashi W., Niwa M. 2000. Apoptotic Cell Death Induces Temperature-Sensitive Lethality in Hybrid Seedlings and Calli Dervied from the cross of *Nicotiana suaveolens* × *N. tabacum*. *Planta*, *211*, 614–622.

CHAPTER THREE : THE RIM LOCUS AFFECTS MATE CHOICE BEHAVIOR IN *DROSOPHILA MELANOGASTER*[2]

Introduction

Reproductive isolation can be broken down into three main categories of reproductively isolating barriers: pre-mating, post-mating pre-zygotic, and post-mating post-zygotic (Dobzhansky, 1937, Coyne and Orr, 2004). Pre-mating barriers describe the set of barriers which cause two species not to mate. These can be as simple as physical differences or involved complex behavioral differences (Arthur and Dyer, 2015). Pre-mating barriers are crucial to our study of speciation, as they are thought to evolve more rapidly than either type of post-mating barrier (Coyne and Orr, 1997, Rabosky and Matute, 2014). However, despite their rapid evolution, pre-mating barriers are often not strong enough to prevent the fusion of nascent species (Comeault et al., 2015, Cenzer, 2016). For this reason, they are thought to be crucial for the initiation of speciation but less important for its maintenance (Rosenblum et al., 2012).

In animals, post-mating barriers are slower to evolve than pre-mating barriers in comparative analyses (Coyne and Orr, 1989, Coyne and Orr, 1997), but are important to prevent fusion of nascent species after divergence has occurred (Rosenblum et al., 2012). They can be sorted into two main categories, pre-zygotic and post-zygotic. Pre-zygotic barriers occur when species mate successfully, but are unable to successfully form a zygote. These barriers typically involve incompatibilities between the male and female reproductive

[2] A manuscript currently in preparation

tracts (Ahmed-Braimah, 2016). Post-zygotic barriers occur when a zygote forms but there is a fitness consequence for the hybrid offspring. This frequently involves reductions in hybrid fertility or total lethality of the hybrid organism (Coyne and Orr, 2004, Maheshwari and Barbash, 2011). It is not uncommon for one gender of offspring to be fertile while the other is sterile, or one gender to be viable but sterile while the other is inviable. Almost always, the heterogametic sex is the more severely affected in these cases, a pattern termed Haldane's Rule (Turelli and Orr, 1995, Delph, 2016). While these post-zygotic barriers are important for maintenance of diverged species, pre-mating barriers are likely more important for initiation of speciation (Mendelson, 2003, Rabosky and Matute, 2013). The strong cosmopolitan/Zimbabwe mate choice phenotype in *Drosophila melanogaster* may provide an example of incipient speciation.

Drosophila melanogaster collected from a region in northern Zimbabwe/southern Zambia in the early 90's were found to have higher levels of divergence in several nuclear genes (Begun and Aquadro, 1993). This was an unusual finding, as previous collections of globally distributed *D. melanogaster* had found the average level of nuclear DNA divergence to be about half that found in the flies collected from Zimbabwe (Kreitman, 1983, Kreitman and Aguade, 1986). It was found that female *D. melanogaster* from these divergent Zimbabwe lines display an extreme mate choice phenotype: they will overwhelmingly (40-50:1) choose Zimbabwe *D. melanogaster* males over cosmopolitan males when offered a choice (Wu et al., 1995). Cosmopolitan females display a slight preference for cosmopolitan males (2:1), but nothing anywhere near as strong as the Zimbabwe females. This pattern persisted across many Zimbabwe lines and the discrimination is found in single female/two male mate choice trials, single female/one male no choice trials, and many females/many

males mating cage trials (Wu et al., 1995). Interestingly, F1 females from crosses between Zimbabwe and cosmopolitan *D. melanogaster* behave like cosmopolitan females (true for F1 females generated from both male Zimbabwe/female cosmopolitan and vice versa), suggesting that the Zimbabwe mate choice phenotype is recessive in nature (Hollocher et al., 1997a).

Initial research suggested that the females may be responding to differences in the cuticular hydrocarbon profiles of the males (Grillet et al., 2006). However, studies mapping and modifying the genes involved have been unable to show that the particular genes underlying the observed changes in cuticular hydrocarbons has an effect on mate choice, suggesting that the cuticular hydrocarbons may be indicative of a larger difference in the males, rather than causative. The mate choice does not seem attributable to differences in song between the males (Colegrave et al., 2000) or to visual cues. In short, it is still not entirely understood by which metric the females are discriminating between the males. While it is possible to get F1s from both directions even with Zimbabwe lines displaying the most extreme mate choice preference, some evidence suggests that some post-mating isolation has begun to arise. Zimbabwe females mated to cosmopolitan males have lower egg hatch rates than when mated to Zimbabwe males (but not higher rates of embryonic lethality, suggesting lower fertilization rate of eggs rather than zygotic mortality), an indication of the onset of post-mating pre-zygotic isolation between Zimbabwe and cosmopolitan *D. melanogaster* (Alipaz et al., 2001).

The majority of the research performed in this system has been in attempt to map the genetic basis of the mate choice phenotype. Some reports found that polymorphism in the Zimbabwe Y chromosome is substantially reduced compared to polymorphism in the

cosmopolitan Y chromosome (Larracuente and Clark, 2013). Similarly, cuticular hydrocarbons originally seemed to be a promising avenue to pursue, and differences in cuticular hydrocarbon production between cosmopolitan and Zimbabwe lines were mapped to the *desat2* gene (Greenberg et al., 2003). Modification of the *desat2* gene showed that altering the cuticular hydrocarbon profile of *D. melanogaster* males has a statistically significant but relatively small effect on female preference (Coyne and Elwyn, 2006). Introgressions of Zimbabwe chromosome arms into cosmopolitan females has shown that the trait is highly polygenic nature, but maps most strongly to chromosome arm 3R (Hollocher et al., 1997b). Chromosome 2 also shows strong association with Zimbabwe mate preference. The X chromosome has some weak association with the trait. (Hollocher et al., 1997b).

We performed deficiency mapping across the whole of chromosome arm 3R in order to map the gene(s) involved in the Zimbabwe female mate preference phenotype. We discovered a region spanning three loci (*couch potato, tincar*, and *rim),* and ultimately mapped the phenotype to the *rim* locus. Sequence analysis revealed that Zimbabwe and cosmopolitan flies differ by two SNPs in the coding sequence of *rim* with high Fst between Zimbabwe and cosmopolitan lines. Using the CRISPR/Cas9 system, we performed precise gene replacement of these two SNPs, swapping the Zimbabwe SNPs in Zimbabwe flies to the cosmopolitan equivalent, and vice versa. Cosmopolitan *D. melanogaster* females bearing the two Zimbabwe SNPs of *rim* display near total (>90%) Zimbabwe mate preference, suggesting that these two SNPs are important for Zimbabwe mate choice.

Results and Discussion

Z female preference is weakly driven by visual cues. To test whether females were discriminating between Z and M males based on visual cues, we performed mate choice trials with Z or M females in well-lit or entirely dark areas (Table 3.1). Z females mated in the

light chose Z males 98.6% of the time, while Z females mated in the dark chose M males 82.6% of the time ($p < 0.0001$, Fisher's exact test). M females mated in the light chose M males 75.8% of the time, while M females mated in the dark chose M males 55.4% of the time ($p = 0.0275$, Fisher's exact test). We conclude from these date that visual cues inform mate choice in *Drosophila melanogaster* females but do not entirely drive it.

Table 3.1 Z Female preference is not strongly affected by visual cues or male song

Genotype + Condition	**M males chosen**	**Z males chosen**	**p**
Z females, light	1	127	
Z females, dark	22	108	< 0.0001
M females, light	94	43	
M females, dark	81	65	0.0275
Z females, winged males	1	64	
Z females, wingless males	5	39	0.0383
M females, winged males	58	13	
M females, wingless males	45	33	0.0024

Mating experiments in the dark show a statistically significant but small loss of Z preference in Z females, and a significant and moderate loss of M preference in M females. Z females given wingless males show a statistically significant but marginal loss of Z preference; M females show a significant and substantial loss of M preference

Z female preference is weakly driven by male song

Previous studies have characterized differences in song between Z and M males (Colegrave et al., 2000), so we were interested in testing the effect this has on discrimination between Z and M males by females. We compared mate choice experiments with Z or M females given males with or without wings; males without wings cannot perform their song (Table 3.1). Z females given males with wings chose Z males 98.7% of the time, while Z

females given wingless males chose Z males 88.8% of the time (p = 0.0383, Fisher's exact test). M females given males with wings chose M males 81.9% of the time, while M females given wingless males chose M males 60.1% of the time (p = 0.00204, Fisher's exact test). We conclude from these data that male song influences mate choice in *Drosophila melanogaster* females but does not entirely determine it.

Table 3.2 Z preference is moderately affected by cuticular hydrocarbons

Genotype + Condition	**M males chosen**	**Z males chosen**	**p**
Z females, normal males	0	25	
Z females, normal Z, M with Z CHCs	8	27	0.0162
Z females, normal M, Z with M CHCs	3	8	0.0231
Z females, M with Z CHCs, Z with M CHCs	2	17	0.1808
M females, normal males	15	4	
M females, normal Z, M with Z CHCs	28	8	1
M females, normal M, Z with M CHCs	15	13	0.122
M females, M with Z CHCs, Z with M CHCs	26	18	0.1588

Z females presented with a Z male and an M male treated with Z CHCs show a significant but moderate loss of Z preference. This is also true when females are presented with an M male and a Z male treated with M CHCs. Swapping the CHC profile of the males has no effected. M preference is not significantly affected by alterations to CHC profiles.

Previous work has shown that Z and M males differ in their CHC profiles, and that this may affect female preference (Greenberg et al., 2003, Elwyn and Coyne, 2006). To test this, we performed mating experiments with Z or M females and various types of male: normal Z or M males, normal Z males and M males with Z CHCs, normal M males and Z males with M CHCs, and Z males with M CHCs and M males with Z CHCs (Table 3.2). Z females given normal Z or M males chose Z males 100% of the time. When given normal Z males and M males with Z CHCs, they chose Z males 77% of the time ($p = 0.0162$, Fisher's exact test); when given normal M males and Z males with M CHCs, they chose Z males 73% of the time ($p = 0.0231$, Fisher's exact test); when given Z males with M CHCs and M males with Z CHCs, they chose Z males 89.5% of the time ($p = 0.1808$, Fisher's exact test). None of the CHC changes had significant effects on M female preference. We conclude from these data that CHC profile plays a moderate role in Z female preference, but is not solely responsible for it.

Table 3.3 Deficiency mapping across 3R reveals that the *rim, couch potato, and tincar* genes are associated with the Zimbabwe mate choice phenotype

Stock	**M males chosen**	**Z males chosen**	**p**
90C6-91A2 Bal/Z	70	24	
90C6-91A2 df/Z	19	89	< 0.0001
90B6-90E2 Bal/Z	58	19	
90B6-90E2 df/Z	22	46	< 0.0001
90C2-90D1 Bal/Z	99	33	
90C2-90D1 df/Z	12	70	< 0.0001

Name of the stock is the cytological bands deleted on the cosmopolitan chromosome in the df-carrying stock; these stocks express only the Zimbabwe allele of any loci contained within those cytological bands.

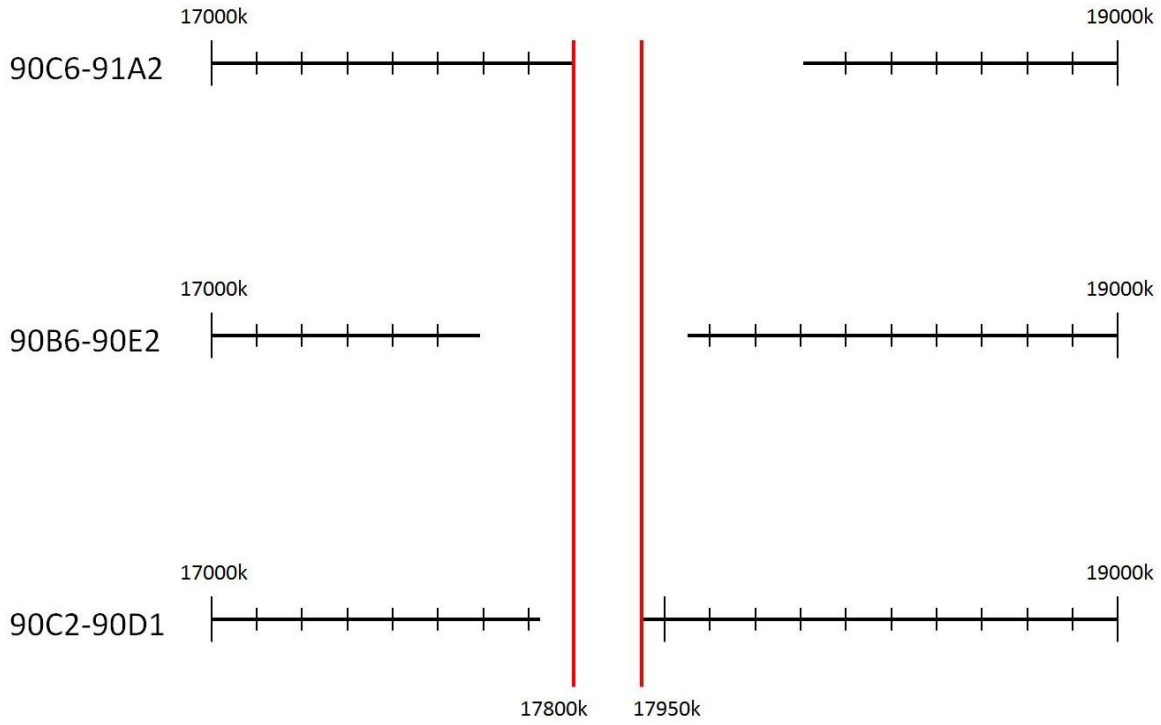

Figure 3.1 Overlap of the deficiency stocks which reveal a recessive Z choice phenotype covers a 150 kbp region of 3R containing the genes *couch potato, tincar, and rim*
Labels on the lefthand side are the cytological bands on 3R deleted in each deficiency.

Deficiency mapping across 3R reveals three genes strongly associated with Zimbabwe preference

Taking advantage of Zimbabwe alleles being recessive to corresponding cosmopolitan alleles, we used deficiency stocks tiling the whole of 3R to uncover recessive Zimbabwe mate choice alleles. The total list of stocks is found in Table S3.2. Flies bearing a deficiency over a Zimbabwe chromosome display cosmopolitan behavior unless the deficiency they carry uncovers a recessive Zimbabwe allele which causes Zimbabwe mate choice preference, in which case they display a shift towards Zimbabwe behavior. We found 3 sets of cytological bands that, when deleted to expose recessive Zimbabwe alleles, trigger a switch in mate choice from cosmopolitan to Zimbabwe (Fisher's exact test $p < 0.0001$ for each cytological band) (Table 3.3). The overlap of these 3 deficiencies covers a 150 kbp

region of 3R containing three genes: *couch potato*, *tincar*, and *rim* (Figure 3.1). This significant association with the Zimbabwe phenotype suggests that the loci contained within this genomic region are strongly associated with the Zimbabwe mate choice phenotype.

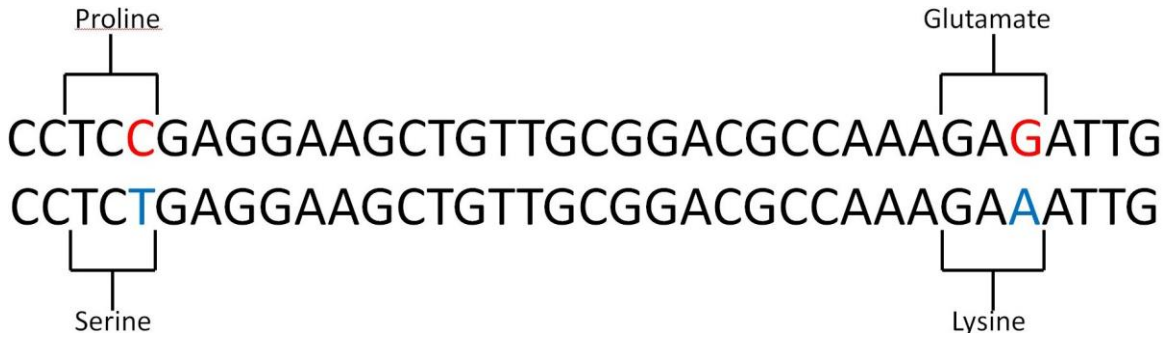

Figure 3.2 High Fst sequence divergence in *rim* between cosmopolitan and Zimbabwe lines

Top row: cosmopolitan sequence. Bottom: Zimbabwe sequence. Divergent sites are marked in red for cosmopolitan lines and blue for Zimbabwe lines. The glutamate to lysine substitution has an Fst of > 0.8; proline to serine, Fst > 0.5.

The *rim* gene affects female mate choice

The overlapping region of the cytological bands associated with Zimbabwe identified in Table 3 is a 150 kbp region of 3R. 3 genes are contained fully within this region: *rim*¸ *couch potato*, and *tincar*. To dissect which genes affected the phenotype, we obtained p-element insertion null stocks of each gene from Bloomington Stock Center. Null stocks for *tincar* and *couch potato* do not recapitulate the observed Zimbabwe mate choice phenotype (Fisher's exact test p = 0.4583 for *couch potato*, p = 0.5424 for *tincar*), while null stocks for *rim* do (Fisher's exact test p = 0.0396) (Table 3.4). To confirm this, we obtained flies with a large segment of the coding region of *rim* deleted, and crossed these *rim* null flies to Zimbabwe flies. These F1 females are heterozygous throughout the genome, except at the *rim* locus, where they express only the Zimbabwe allele. As with the deficiency mapping, these flies display a strong Zimbabwe preference compared to cosmopolitan/Zimbabwe flies

(Fisher's exact test $p < 0.0001$) (Table 3.4), showing that the *rim* allele a female *D. melanogaster* carries strongly affects her mate choice.

Table 3.4 Zimbabwe alleles of *rim* cause the Zimbabwe mate choice phenotype

Stock	**M males chosen**	**Z males chosen**	**p**
couch potato Bal/Z	16	15	
couch potato p-element null/Z	14	21	0.4583
tincar Bal/Z	7	8	
tincar p-element null/Z	21	15	0.5424
rim Bal/Z	17	8	
rim p-element null/Z	14	21	0.0396
rim103b Bal/Z	55	28	
rim103b rim103b/Z	22	74	< 0.0001

p-element insertion nulls of *tincar* and *couch potato* do not display the Zimbabwe mate choice phenotype, while p-element insertion nulls of *rim* do. Flies which are *rim* null/Zimbabwe *rim* also display a Zimbabwe mate preference.

Zimbabwe or cosmopolitan mate choice is controlled by two SNPs in *rim*

Sequencing and genetic comparison of Zimbabwe and cosmopolitan lines revealed that *rim* differs between Zimbabwe and cosmopolitan lines by two conserved SNPs in its coding region, both of which cause non-synonymous amino acid substitutions (Figure 3.2). To determine if these SNPs were indeed the causative variants, we used the CRISPR-Cas9 system to generate flies bearing the SNPs of the opposite mating type (i.e., cosmopolitan flies carrying the two Zimbabwe SNPs in *rim*, and vice versa). We found that cosmopolitan female flies bearing the Zimbabwe *rim* SNPs display an overwhelmingly strong preference for Zimbabwe males (Table 3.5), suggesting that these two SNPs are strongly involved with the observed mating preference in Zimbabwe females ($p < 0.0001$, Fisher's exact test).

Surprisingly, we found that the cosmopolitan SNPs of *rim* are homozygous lethal on a Zimbabwe background. We maintained the Zimbabwe chromosome bearing the cosmopolitan allele of *rim* against a balancer chromosome; the X chromosome and chromosome 2 were homozygous Zimbabwe chromosomes. Mating experiments were performed with the *rim*M offspring of these *rim*M/Balancer flies mated to pure Zimbabwe flies (Table 3.5). We found that flies which are Zimbabwe throughout the genome but carry one copy of the cosmopolitan SNPs of *rim* display a strong cosmopolitan preference ($p < 0.0001$, Fisher's exact test), showing that the dominant cosmopolitan phenotype observed in previous F1s can be caused by one copy of the cosmopolitan SNPs of *rim*.

Table 3.5 The identified SNPs in *rim* strongly affect female mating preference

Stock	**M males chosen**	**Z males chosen**	**p**
Cosmopolitan	53	24	
Cosmopolitan with Zimbabwe *rim*	4	78	< 0.0001
Zimbabwe	1	72	
Zimbabwe heterozygous for cosmopolitan *rim*	12	8	< 0.0001

Cosmopolitan females bearing the Zimbabwe SNPs display Zimbabwe preference. Zimbabwe females heterozygous for the cosmopolitan SNP display cosmopolitan preference.

We found that Zimbabwe mate choice is not driven by a single cue. Visual cues, audio cues, and olfactory cues all have a significant but modest impact on Z female preference, suggesting that no one feature of *D. melanogaster* males is being discriminated against or for. While no researchers have succeeded in quantifying visual differences between M and Z males, differences in song and CHC profile have been shown in the literature (Colegrave et al., 2000, Greenberg et al., 2003, Elwyn and Coyne, 2006). If one of

these variables were the main cue females were responding to, we would expect that normalizing it between Zimbabwe and cosmopolitan males should produce a 50/50 choice split for both types of female (e.g., if it were mostly a visual cue, females in the dark would lose their discrimination criteria and should display substantially less strong preference). We did not observe this for any of the variables tested. Rather, each produced a modest but statistically significant shift in female preference, suggesting that Zimbabwe females are discriminating against multiple characteristics in the male flies.

These data show that one locus is able to strongly explain the variance observed between Zimbabwe and cosmopolitan females. The *rim* gene, which is most highly expressed in the central nervous system, is capable of producing large shifts in mate choice behavior by itself. Otherwise cosmopolitan flies expressing a Zimbabwe allele of *rim* choose Zimbabwe males over cosmopolitan males. This supports the theory that Zimbabwe mate choice is driven primarily by one or several causative alleles, with a larger number of Zimbabwe alleles having arisen after the initial population split occurred and gene flow was reduced. We would not expect to see so strong a signal from a single locus had the trait initially evolved as highly complex multi-locus trait. Even more interestingly, Zimbabwe and cosmopolitan differ in *rim* by only two highly conserved SNPs which segregate in those populations, rather than large changes to the gene itself, and RNAseq data from previous work suggests that the expression level of *rim* is not significantly different between Zimbabwe and cosmopolitan flies (Gelbert and Emmert, 2013). It is likely that one or two amino acid changes caused a functional change in the activity and/or targets of the RIM protein, which had the consequence of affecting mate choice.

Our results show that Zimbabwe *rim* SNPs are sufficient to cause Zimbabwe-like behavior. Transforming cosmopolitan flies to carry the Zimbabwe allele of *rim* causes these flies to display an overwhelmingly strong preference for Zimbabwe males, even more strongly than null/Zimbabwe flies (which may indicate a dosage effect of the RIM protein on Zimbabwe preference). That two SNP changes are sufficient to drive such a strong shift in phenotype lends support to our theory that the Zimbabwe preference is largely controlled by a small number of loci (or perhaps even a single locus), and that the phenotype arose quite suddenly within the populations, causing the initial split. These results suggest that pre-mating barriers, which are important for the initiation of speciation, can be driven by changes at the level of single base pairs, rather than requiring sweeping genomic changes. This is important for the process of speciation. That changes in one or two SNPs can induce such extreme alterations in mating behavior strongly suggests that speciation can be driven by recessive changes at the base-pair level.

We also found that the cosmopolitan *rim* SNPs are sufficient to cause cosmopolitan-like behavior, even when heterozygous. Genetically Zimbabwe flies which are homozygous for the cosmopolitan allele of *rim* do not survive to pupation, and so transformant flies must be maintained as heterozygotes over a balancer chromosome. This likely indicates a recessive DMI between the cosmopolitan allele of *rim* and the Zimbabwe allele of another locus. This may suggest that the Zimbabwe mate choice phenotype is a response to this putatitive DMI, providing strong selective pressure for the fixation of the Zimbabwe alleles of *rim* and thus the pre-mating isolation observed in Zimbabwe lines (alternatively, this may indicate the beginning of the evolution of post-zygotic incompatibilities between Zimbabwe and cosmopolitan *Drosophila melanogaster*). As might be expected given the cosmopolitan

preference of F1 females, genetically Zimbabwe flies bearing one copy of the cosmopolitan allele of *rim* over a Zimbabwe allele of *rim* display a cosmopolitan preference, suggesting that the dominance of the cosmopolitan phenotype may be attributable to the cosmopolitan allele of *rim* borne by the F1s. Overall, our data strongly suggest that the alleles of *rim* a female bears have a strong influence on her mate preference.

Overall, our data suggest that the Zimbabwe/cosmopolitan split provides evidence that one or two SNPs can drive speciation. The cosmopolitan and Zimbabwe populations of *Drosophila melanogaster* appear to be undergoing a speciation event. Their genomes have begun to diverge substantially (Begun and Aquadro, 1993, Ting et al., 2001, Kauer et al., 2002, Kauer and Schlotterer, 2004, Larracuente and Clark, 2013), and post-mating isolation has begun to arise between the two populations (Alipaz et al., 2001). These data suggest that this is the onset of speciation between these two populations, likely driven by the cessation of gene flow due to the very strong pre-mating isolation between the two populations. Our data show that simply modifying two SNPs in the *rim* gene is sufficient to cause cosmopolitan *D. melanogaster* females to almost entirely prefer Zimbabwe males, suggesting that one or two SNPs are sufficient to drive strong pre-mating isolation and ultimately induce speciation between two populations.

Materials and Methods

We mapped the loci involved in the cosmopolitan/Zimbabwe mate choice in female *Drosophila melanogaster* using single female mate choice assays and identified candidate SNPs, which we modified using the CRISPR/Cas9 system. Modified females were then tested in the same single female mate choice experiments. We describe each step as follows.

Virgin collection and stocks

Table S1 lists all the stocks used in this report. For mating experiments and modification of flies with CRISPR/Cas9 we used one isofemale line of cosmopolitan *D. melanogaster*, Ral-528, obtained from Bloomington Stock Center, and one isofemale line of Zimbabwe *D. melanogaster*, ZS2, obtained from collections in Zimbabwe in 1990 (Begun and Aquadro, 1990). We let females oviposit; when larvae were observed in the bottles, they were monitored daily for black pupae. All lines were reared on standard cornmeal/Karo/agar medium at 24ºC under a 12 h light/dark cycle in 100mL bottles.

Drosophila melanogaster deficiency stocks tiling the whole of chromosome arm 3R were purchased from Bloomington Stock Center (*rim103b* stocks were obtained from Graeme Davis's lab in California). Once quarantined, stocks were expanded in 200mL plastic bottles containing cornmeal food. We let females oviposit; when larvae were observed in the bottles, they were monitored daily for black pupae. All flies were kept at 24ºC under a 12 hour light/dark cycle.

To cross cosmopolitan *D. melanogaster* deficiency stocks to male Zimbabwe *D. melanogaster*, as well as to perform single female mate choice assays, we needed virgin females. Stocks were kept in 300 mL plastic bottles with cornmeal fly food. Flies were allowed to mate and oviposit for 1 week, and then cleared. Once dark pupae were observed, bottles were cleared every 12 hours and monitored for eclosion. Females were collected within 8 hours of eclosion under CO_2 anesthesia and kept for three days in single-sex groups of 20 flies in 30 mL vials containing corn meal food. Zimbabwe males were collected as well but were not necessarily virgins, and kept in all-male vials containing 20 individuals per vial. On day four, we assessed whether there were larvae in the media in both the female and male

vials. If the inspection revealed any progeny, the vial was discarded. If the vials had no larvae, the virgin individuals were used for crosses.

Deficiency mapping

Because Zimbabwe mate preference is recessive to cosmopolitan mate preference, we used deficiency mapping to detect regions of the genome containing recessive Zimbabwe alleles with an influence on female mate preference. Our crossing design detects recessive Zimbabwe-choice regions by uncovering regions containing a recessive Zimbabwe-choice allele; deficiency/Zimbabwe stocks lack the cosmopolitan allele (which is dominant) in those genomic regions, and so express only the Zimbabwe alleles. Any deficiencies which show a change in mate choice have uncovered a Zimbabwe allele affecting the mate choice. The approach involves crossing females from *Drosophila melanogaster* (*mel*) stocks containing known genomic deletions, or "deficiencies" (*df*, Bloomington Drosophila Fly Stock Center), maintained as heterozygotes against a balancer (*Bal*) chromosome carrying a dominant homozygous lethal mutation, to Zimbabwe *D. melanogaster* males. On day 4 after virgin collection, males and females were mixed in a 30mL plastic vial with cornmeal fly food. Females were crossed to males at a 1:1 ratio with at least 10 females used per cross. Vials were inspected every five days to check for progeny. We transferred the parents to a new vial when we observed larvae.

Mate choice assays

We used single female mate choice assays to test the mate choice preference of cosmopolitan, Zimbabwe, Balancer/Zimbabwe, and deficiency/Zimbabwe *D. melanogaster* females. First, virgin females of each stock or Bal/Z or df/Z being tested were collected as per the virgin collection protocol noted above. The night before the experiment, we placed cosmopolitan and Zimbabwe males food dyed with red or blue food coloring, which marks

them without affecting mate preference (Table S3.1). We placed one virgin female of each stock or genotype being tested in a vial of food (50 replicates), and added one cosmopolitan and one Zimbabwe male to each vial containing a virgin female. Flies were allowed to mate for 2 hours, during which time vials were checked every 5 minutes. When a copulation is observed, the non-chosen male was aspirated out and disposed of. After two hours, all vials in which matings were observed were scored for the type of male chosen by the female.

Mate choice experiments performed in the dark were performed in a completely dark internal room with no source of visible light. For mate choice experiments involving males with or without wings, wings were removed from males by surgery 24 hours prior to mating. For CHC experiments, males were co-habitated with males of the opposite race, causing them to acquire the CHCs of the opposite race males through continued close contact.

CRISPR/Cas9

Through deficiency mapping and genetic null flies, the *rim* gene was positively identified as being involved with Zimbabwe mate preference. Analysis of existing sequencing data shows that cosmopolitan and Zimbabwe lines differ by two SNPs in *rim,* one with an Fst of 0.8 and one with an Fst of 0.5. We identified these as candidate SNPs for modification with CRISPR/Cas9.

Guide RNA oligos were designed using the "Optimal Target Finder" tool at http://flycrispr.molbio.wisc.edu/tools, and ordered from Bioneer. The chimeric RNA backbone plasmid pBbs1-chiRNA was provided by the Jeff Sekelsky group at UNC Chapel Hill. The backbone construct was cut with Bbs1 purchased from New England Biolabs and gel purified on a 1% agarose gel using the QiaQuick gel extraction kit; guide RNA oligos were added and ligated into the construct backbone with New England Biolabs T7 DNA ligase. Ligated guide RNA constructs were transformed into Top10 *E. coli* purchased from

Invitrogen and colonies grown at 37 C overnight. Selected colonies were sequenced to check for proper ligation of the guide RNA oligos into the construct. Cas9 was expressed from a Heat Shock Promoter 70-Cas9 vector, obtained from the Sekelsky group. Repair templates were synthesized using the GeneArt synthesis system by Thermo Fisher. All constructs were transformed into Top10 *E. coli* and maintained as glycerol/LB stocks at -80 C; aliquots were grown in 250 mL of LB broth and purified by Qiagen Maxi kit prep to prepare for injection.

Ral-528 and ZS2 flies were sent to Model Injection Systems to prepare for injections, and injections used a construct mix with these concentrations: 300 ng/uL Hsp70-Cas9, 600 ng/uL repair template, 150 ng/uL for each guide RNA construct. Injected progeny were collected as virgins and mated singly to Zimbabwe or Cosmopolitan flies containing a TM3 balancer chromosome with the Stubbly phenotypic marker; when progeny were observed, transformants were harvested by digestion with Proteinase K in squishing buffer [put concentrations here]. A 450 bp band containing the sites of interested was amplified through PCR using Invitrogen Taq polymerase and dNTPs and purified using the Thermo Fisher GeneJet PCR cleanup system, before being sent for Sanger sequencing by Eurofins. TM3-bearing progeny of positive transformants were collected as virgins and mated 1:1 with their siblings, and then sequenced themselves. Homogeneous modified lines were derived from matings occurring between two flies bearing a transformant chromosome/TM3. Virgin progeny not bearing TM3 were collected from these crosses and kept in 300 mL bottles containing cornmeal food under a 12h light/dark cycle to create true-breeding modified lines (Zimbabwe flies with cosmopolitan *rim* were maintained as heterozygotes over TM3; cosmopolitan *rim* was found to be recessive lethal on a Zimbabwe background).

Supplemental Data

Supplemental Table 3.1 Color of the food dye given to males does not significantly affect female preference

Males were both the same race as the female.

Female type	Yellow males chosen	Green males chosen	χ^2	p
Cosmopolitan	32	28	0.267	0.6056
Zimbabwe	20	23	0.209	0.6473

Supplemental Table 3.2 List of all deficiency stocks and mating results used to map Zimbabwe preference on 3R

Bolded selections are included in Table 3.3. Name of the stock is the cytological bands deleted on the cosmopolitan chromosome in the df-carrying stock; these stocks express only the Zimbabwe allele of any loci contained within those cytological bands.

Stock	M males chosen	Z males chosen
81F-81F Bal/Z	75	24
81F-81F df/Z	68	42
82F8-83A4 Bal/Z	96	32
82F8-83A4 df/Z	84	28
83A6-83B6 Bal/Z	102	37
83A6-83B6 df/Z	95	26
83B4-83B6 Bal/Z	122	47
83B4-83B6 df/Z	125	47
83B7-83D1 Bal/Z	94	49
83B7-83D1 df/Z	88	40
83B7-83E1 Bal/Z	58	18
83B7-83E1 df/Z	53	31
83C1-84B2 Bal/Z	98	46
83C1-84B2 df/Z	78	38

83E2-83E5 Bal/Z	99	46
83E2-83E5 df/Z	86	29
85A5-85D1 Bal/Z	103	31
85A5-85D1 df/Z	57	49
85D16-85D24 Bal/Z	90	29
85D16-85D24 df/Z	92	22
85D1-85D11 Bal/Z	113	41
85D1-85D11 df/Z	64	44
85D6-85D15 Bal/Z	71	25
85D6-85D15 df/Z	80	31
85E9-85F1 Bal/Z	96	48
85E9-85F1 df/Z	93	47
85F11-86B1 Bal/Z	57	14
85F11-86B1 df/Z	46	22
85F5-85F14 Bal/Z	65	22
85F5-85F14 df/Z	51	24
86C7-86D7 Bal/Z	58	24
86C7-86D7 df/Z	52	27
86C7-86E13 Bal/Z	111	48
86C7-86E13 df/Z	101	32
86D8-87A2 Bal/Z	66	43
86D8-87A2 df/Z	75	19
86F9-87B13 Bal/Z	107	46
86F9-87B13 df/Z	81	32
87B10-87E9 Bal/Z	64	32

87B10-87E9 df/Z	64	19
87E3-88A4 Bal/Z	81	28
87E3-88A4 df/Z	67	34
88A4-88C9 Bal/Z	81	44
88A4-88C9 df/Z	105	53
88F6-89A8 Bal/Z	66	21
88F6-89A8 df/Z	50	12
89A8-89B2 Bal/Z	79	20
89A8-89B2 df/Z	80	20
89B17-89D5 Bal/Z	74	35
89B17-89D5 df/Z	63	10
89B18-89D8 Bal/Z	124	54
89B18-89D8 df/Z	133	49
89B7-89E7 Bal/Z	64	29
89B7-89E7 df/Z	51	12
89E11-90C1 Bal/Z	83	32
89E11-90C1 df/Z	74	13
89E1-89E2 Bal/Z	87	33
89E1-89E2 df/Z	66	20
89E5-89E11 Bal/Z	52	14
89E5-89E11 df/Z	48	35
90B6-90E2 Bal/Z	**58**	**19**
90B6-90E2 df/Z	**22**	**46**
90C2-90D1 Bal/Z	**99**	**33**
90C2-90D1 df/Z	**12**	**70**

90C6-91A2 Bal/Z	**70**	**24**
90C6-91A2 df/Z	**19**	**89**
90F4-91B8 Bal/Z	64	21
90F4-91B8 df/Z	73	31
91A5-91F1 Bal/Z	64	25
91A5-91F1 df/Z	86	30
91D4-92A11 Bal/Z	84	27
91D4-92A11 df/Z	55	29
92A11-92E2 Bal/Z	100	44
92A11-92E2 df/Z	123	68
92F7-93B6 Bal/Z	99	44
92F7-93B6 df/Z	90	50
93A2-93B8 Bal/Z	71	31
93A2-93B8 df/Z	68	46
93A4-93B13 Bal/Z	84	42
93A4-93B13 df/Z	78	40
93B9-93D4 Bal/Z	83	40
93B9-93D4 df/Z	87	27
94F1-95A4 Bal/Z	66	33
94F1-95A4 df/Z	79	16
95C12-95D8 Bal/Z	85	35
95C12-95D8 df/Z	62	28
95D10-96A7 Bal/Z	75	24
95D10-96A7 df/Z	102	57
95E7-96A18 Bal/Z	129	54

95E7-96A18 df/Z	113	52
96A7-96C3 Bal/Z	70	19
96A7-96C3 Bal/Z	71	41
96A7-96C3 df/Z	63	50
96A7-96C3 df/Z	58	14
96C8-96D1 Bal/Z	69	44
96C8-96D1 df/Z	81	29
96E2-96E6 Bal/Z	73	31
96E2-96E6 df/Z	79	35
96F10-97D2 Bal/Z	79	27
96F10-97D2 df/Z	75	26
98B6-98B6 Bal/Z	65	25
98B6-98B6 df/Z	68	15
98B6-98E5 Bal/Z	86	24
98B6-98E5 df/Z	94	24
98F10-99B9 Bal/Z	89	28
98F10-99B9 df/Z	77	47
99B5-99C2 Bal/Z	95	29
99B5-99C2 df/Z	89	37
99B5-99F1 Bal/Z	88	42
99B5-99F1 df/Z	83	35
99D1-99E1 Bal/Z	55	33
99D1-99E1 df/Z	53	32
99D3-99D8 Bal/Z	106	38
99D3-99D8 df/Z	98	19

99D5-99E2 Bal/Z	67	26
99D5-99E2 df/Z	82	22
99E1-3Rt Bal/Z	97	31
99E1-3Rt df/Z	79	48
99E3-99F6 Bal/Z	90	14
99E3-99F6 df/Z	64	46
99F8-100A5 Bal/Z	85	20
99F8-100A5 df/Z	90	43
100A-100F Bal/Z	101	31
100A-100F df/Z	91	33
100A5-100B1 Bal/Z	73	33
100A5-100B1 df/Z	56	30
100B1-100C1 Bal/Z	96	51
100B1-100C1 df/Z	99	41
100C7-100E3 Bal/Z	83	31
100C7-100E3 df/Z	70	18
100E1-100E3 Bal/Z	109	43
100E1-100E3 df/Z	119	43

REFERENCES

1. Ahmed-Braimah, Y. H. (2016). Multiple genes cause postmating prezygotic reproductive isolation in the Drosophila virilis group. *Genes Genomes Genetics*, *X*, 1–17. http://doi.org/10.1534/g3.116.033340

2. Alipaz, J. A., Wu, C., & Karr, T. L. (2001). Gametic incompatibilities between races of Drosophila melanogaster. *Proceedings of the Royal Society of London*, *268*(July 2000), 789–795. http://doi.org/10.1098/rspb.2000.1420

3. Arthur, N. J., & Dyer, K. A. (2015). Asymmetrical sexual isolation but no postmating isolation between the closely related species Drosophila suboccidentalis and Drosophila occidentalis. *BMC Evolutionary Biology*, *15*, 1–9. http://doi.org/10.1186/s12862-015-0328-y

4. Begun, D. J., & Aquadro, C. F. (1993). African and North American populations of Drosophila melanogaster are very different at the DNA level. *Nature*, *365*, 548–550.

5. Cenzer, M. L. (2016). Adaptation to an invasive host is driving the loss of a native ecotype. *Evolution*, (70), 2296–2307. http://doi.org/10.1111/evo.13023.This

6. Colegrave, N., Hollocher, H., Hinton, K., & Ritchie, M. G. (2000). The courtship song of African Drosophila melanogaster. *Journal of Evolutionary Biology*, *13*, 143–150.

7. Comeault, A. A., Flaxman, S. M., Riesch, R., Schwander, T., Slate, J., & Nosil, P. (2015). Selection on a genetic polymorphism counteracts ecological speciation in a stick insect. *Current Biology*, *25*, 1975–1981. http://doi.org/10.1016/j.cub.2015.05.058

8. Coyne, J. A., & Elwyn, S. (2006). Does the desaturase-2 locus in Drosophila melanogaster cause adaptation and sexual isolation? *Evolution*, *60*, 279–291.

9. Coyne, J. A., & Orr, H. A. (1989). Patterns of Speciation in Drosophila. *Evolution*, *43*(2), 362–381.

10. Coyne, J. A., & Orr, H. A. (1997). "Patterns of speciation in Drosophila" revisited. *Evolution*, *51*, 295–303.

11. Coyne, J. A., & Orr, H. A. (2004). *Speciation* (First Edit). Sinauer Associates.

12. Delph, L. F., & Jeffery, P. (2016). Haldane's Rule: genetic bases and their support. *Journal of Heredity*, 383–391. http://doi.org/10.1093/jhered/esw026

13. Dobzhansky, T. (1937). Further data on the variation of the Y chromosome in Drosophila pseudoobscura. *Genetics*, *340*, 340–346.

14. Gelbert, W. M., & Emmert, D. B. (2013). FlyBase High Throughput Expression Pattern Data. *Flybase*

15. Greenberg, A. J., Moran, J. R., & Coyne, J. A. (2003). Ecological adaptation during incipient speciation revealed by precise gene replacement. *Science*, *111*, 1754–1757.

16. Grillet, M., Dartevelle, L., & Ferveur, J.-F. (2006). A Drosophila male pheromone affects female sexual receptivity. *Proceedings of the Royal Society B*, *273*, 315–323. http://doi.org/10.1098/rspb.2005.3332

17. Hollocher, H., Ting, C.-T., Pollack, F., & Wu, C.-I. (1997a). Incipient speciation by sexual isolation in Drosophila melanogaster: variation in mating preference and correlation between sexes. *Evolution*, *51*(4), 1175–1181.

18. Hollocher, H., Ting, C.-T., Wu, M.-L., & Wu, C.-I. (1997b). Incipient speciation by sexual isolation in Drosophila melanogaster: extensive genetic divergence without reinforcement. *Genetics*, *147*, 1191–1201.

19. Kauer, M., Zangerl, B., Dieringer, D., & Schlotterer, C. (2002). Chromosomal patterns of microsatellite variability contrast sharply in African and non-African populations of Drosophila melanogaster. *Genetics*, *256*, 247–256.

20. Kauer, M. O., & Schlotterer, C. (2004). An analysis of genetic differentiation among assortatively mating Drosophila melanogaster in Zimbabwe. *Journal of Evolutionary Biology*, *17*, 493–500. http://doi.org/10.1111/j.1420-9101.2004.00709.x

21. Kreitman, M. (1983). Nucleotide polymorphism at the alcohol dehydrogenase locus of Drosophila melanogaster. *Nature*, *304*, 412-417

22. Kreitman, M. E., & Aguade, M. (1986). Excess polymorphism at the adh locus in Drosophila melanogaster. *Genetics*, *114*, 93–110.

23. Larracuente, A. M., & Clark, A. G. (2013). Surprising differences in the variability of Y chromosomes in African and cosmopolitan populations of Drosophila melanogaster. *Genetics*, *193*, 201–214. http://doi.org/10.1534/genetics.112.146167

24. Maheshwari, S., & Barbash, D. A. (2011). The genetics of hybrid incompatibilities. *Annual Review of Genetics*, *45*, 331–355. http://doi.org/10.1146/annurev-genet-110410-132514

25. Mendelson, T. C. (2003). Sexual isolation evolves faster than hybrid inviability in a diverse and sexually dimorphic genus of fish (Percidae: etheostoma). *Evolution*, *57*, 317–327.

26. Rabosky, D. L., & Matute, D. R. (2013). Macroevolutionary speciation rates are decoupled from the evolution of intrinsic reproductive isolation in Drosophila and birds. *PNAS*, *110*, 15354–15359. http://doi.org/10.1073/pnas.1305529110/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1305529110

27. Rosenblum, E. B., Sarver, B. A. J., Brown, J. W., Des Roches, S., Hardwick, K. M., Hether, T. D., … Harmon, L. J. (2012). Goldilocks meets Santa Rosalia: an ephemeral speciation model explains patterns of diversification across time scales. *Evolutionary Biology*, *39*, 255–261. http://doi.org/10.1007/s11692-012-9171-x

28. Ting, C., Takahashi, A., & Wu, C. (2001). Incipient speciation by sexual isolation in Drosophila: concurrent evolution at multiple loci. *PNAS*, *98*(12), 6709–6713.

29. Turelli, M., & Orr, H. A. (1995). The dominance theory of Haldane's Rule. *Genetics*, *140*, 389–402.

30. Wu, C., Hollocher, H., Begun, D. J., Aquadro, C. F., Xu, Y., & Wu, M.-L. (1995). Sexual isolation in Drosophila melanogaster : A possible case of incipient speciation. *Proceedings of the National Academy of Sciences*, *92*, 2519–2523.

CHAPTER FOUR : DISCUSSION AND FUTURE WORK

Introduction

Understanding reproductive isolation is crucial for understanding the diversity of life we observe on the planet. Without pre-mating isolation, species cannot diverge and specialized into niches or ecological roles (Butlin et al., 2008). Without post-mating isolation, species which have diverged are likely to merge back into one species through introgression and gene flow (Rosenblum et al., 2012).

Pre-mating isolation is important and necessary for initiation of speciation. Speciation occurs when two populations of a species experience a cessation of gene flow, such that allele frequencies in the two populations are no longer equal and can diverge over time through drift or differential selection on the two populations. Pre-mating isolation initiates this process by causing mating to cease, which halts gene flow and leads to speciation. Pre-mating isolation is a side effect of speciation in allopatry, but is entirely necessary for speciation in sympatry or peripatry (Johannesson, 2010). Our data clearly show that extreme pre-mating isolation can be caused by even small genomic changes: one or two SNPs are sufficient to cause extraordinarily strong pre-mating isolation between two sympatric populations. This suggests that the genetic mechanisms of speciation need not be complex or polygenic, but may arise as a consequence of a single mutation.

Post-mating isolation evolves much more slowly, but is equally as important for speciation (Comeault et al., 2015). While it likely cannot be an initiating factor in speciation (as it is unlikely that two populations producing viable progeny would somehow manage to

evolve post-mating isolation whilst gene flow is occurring), it is critical for the maintenance of species by lowering rates of gene flow between diverged species and creating negative fitness consequences for hybrids (Cenzer, 2016). Studying post-mating isolation may not help us understand the factors which initiate speciation, but it is key for understanding how species diverge over time and what these genetic mechanisms tell us about evolution.

Post-Mating Isolation

We found that hybrids between *D. melanogaster* and *D. simulans* are both more temperature sensitive than those between *D. melanogaster* and *D. santomea*, and that they are more viable at a higher temperature. This finding disagrees with previous work in insect systems (Koevoets et al., 2012), and is fairly unintuitive, since *D. simulans* is not a temperature specialist species. Our finding that sensitivity to extrinsic defects decreases with divergence time between species suggests that after many DMIs have accumulated, an environment less harsh on the progeny will not reduce stress enough for hybrids with many DMs to survive. While environmental effects clearly do play an important role in post-zygotic isolation, they are more impactful in hybrids with several major incompatibilities, rather than many. These data suggest that post-zygotic isolation is complex and can be affected by a myriad of factors; as such, we cannot simply consider it independent of environmental variables.

We found no effect of chromosomal location on hybrid inviability, which is a surprising and unexpected result. Previous mapping studies have found that the X chromosome frequently plays a large role in *Drosophila* hybrid fitness (Moehring et al., 2006), and a number of hypotheses have been put forth to explain the X chromosome's disproportionate impact on hybrid fitness (e.g. faster X hypothesis). That we find no substantial effect of the X chromosome on hybrid fitness suggests that after species are

sufficiently diverged, the disproportionate effect of the X may be lessened; a possible conclusion is that the X is most influential during early speciation. It is also significant that we find no substantial interaction between chromosomal location and temperature. The effect of temperature on hybrid viability seems not to be locus-specific, and instead has broad genome-wide effects. Future studies of post-mating isolation should consider that it may not involve intuitive loci or factors, and that seemingly anything may be involved in the process. Environmental factors can no longer be disregarded in studies of post-zygotic isolation between diverged species.

Future Studies of Post-Mating Isolation

One limitation of the data presented in Chapter 2 is that we examined only two types of hybrid, *D. melanogaster*/*D. simulans* and *D. melanogaster*/*D. santomea*. It would be interesting and informative to expand our study to other species pairs, and compare the results. For example, comparing hybrids of *D. melanogaster* with either *D. yakuba* or *D. teissieri* to the hybrids of *D. melanogaster* with *D. santomea* would be revealing, as *D. teissieri* and *D. yakuba* are sister species of *D. santomea* with different divergence times to *D. santomea* (2.4 million and 1 million years, respectively), but an identical divergence time to *D. melanogaster* (15 million years). Similarly, comparing hybrids of *D. melanogaster* and *D. mauritiana* with the *D. simulans* hybrids would be most informative, as *D. mauritiana* and *D. simulans* are more similar to each other than they are to *D. melanogaster*.

Our results also present an opportunity to map the loci involved in DMIs in the species crosses we examined. While we corrected for overlapping deficiencies to determine which regions contained unique hybrid incompatibilities, discarding regions which do not overlap could present a narrow range of cytological bands containing loc involved in hybrid incompatibility. Targeted knockout with p-element insertion could theoretically allow

targeted examination of the loci contained within these regions, allowing mapping of at least one of the partners involved in the DMIs in these crosses. Mapping the second partner would be more difficult, but possible, and would involve more molecular work than we performed here (e.g. CoIP). As few examples of confirmed DMIs exist (Cooper and Phadnis, 2016), this work would be valuable in and of itself.

Pre-Mating Isolation

While it has always been thought that pre-mating isolation evolves rapidly (Orr et al., 1997), the study of the cosmopolitan and Zimbabwe races of *Drosophila melanogaster* has shown that it can evolve extraordinarily rapidly and with little genomic change. Our data show that with as few as two SNP changes, nearly total pre-mating isolation can evolve, and need not involve geographic or post-mating barriers to gene flow. Additionally, Chapter 3 shows that even in a highly complex trait such as mate choice, changes at the base pair level can have drastic effects, which can in turn have enormous effects on speciation and species divergence. While speciation is thought of as a complex process involving sweeping genetic changes, these data show that minute genetic alterations could, in theory, drive speciation.

Our finding that the cosmopolitan allele of *rim* is recessive lethal on an otherwise Zimbabwe background is quite interesting, and equally unexpected. While it is unclear why this is the case, it does suggest that there may be selective pressure against this allele of *rim* in otherwise genetically Zimbabwe flies. This could suggest a selective mechanism for how the Zimbabwe allele of *rim* spread to prevalence so rapidly, and why there is such extreme mating isolation between these two populations. If one population carries a recessive lethal allele, it is likely interbreeding would eventually cease due to selection against the deleterious allele. Most likely this is the result of a DMI between the cosmopolitan allele of *rim* and the Zimbabwe allele of another gene.

Future Studies of Pre-Mating Isolation

Chapter 3 shows that initiation of speciation in animals can occur rapidly, and with small genomic changes. Examining diverged species for the signatures of this may be revealing. For example, *D. santomea* and *D. yakuba* are thought to have diverged on the island of Sao Tome roughly 1 million years ago (Turissini et al., 2015), and share a region of peripatry in which hybrids are regularly collected (Llopart et al., 2005).

Our work also suggests exciting future potential for study in the Zimbabwe/cosmopolitan mate choice system, including determining fully what cues the Zimbabwe females are basing their mate decision on. Additionally, mapping the other loci associated with the mate choice would be informative - while 3R mapped most strongly to the trait, 3L and all of chromosome 2 also had associations with mate choice, suggesting that there are other loci involved with the trait to be uncovered (Hollocher et al., 1997). Studies of these loci would be informative for determining which was potentially the first to evolve in the Zimbabwe population, and might suggest a mechanism for how this split occurred.

The prospect of mapping the recessive DMI between the cosmopolitan allele of *rim* and an as yet unidentified Zimbabwe allele at another locus is also very exciting. While the work would be relatively arduous and time consuming, small introgressions of cosmopolitan chromosomes into the Zimbabwe genome would map the genomic location of the interacting partner down to a smaller region, at which point finer mapping techniques could be used to dissect the precise genomic location. This may reveal the genetic changes which drove the initial Zimbabwe choice phenotype and explain the apparent incipient speciation between these two races of *Drosophila melanogaster*.

Summary

Reproductive isolation is a hugely important topic, and one well worth studying. Studies of pre-mating and post-mating isolation both further our understanding of the processes involved in speciation and evolution, how species split and how this split is maintained over millions of years. Studying these processes can inform our understanding of all biological processes: why organisms behave the way they do, why complex genomic interactions occur in the manner we observe, why seemingly miniscule changes in the genome can have enormous consequences for a living thing. Reproductive isolation is a field of great breadth and depth; I look forward to following future work in it.

REFERENCES

1. Butlin, R. K., Galindo, J., & Grahame, J. W. (2008). Sympatric, parapatric or allopatric: the most important way to classify speciation? *Philosophical Transactions of the Royal Society B*, *363*(=), 2997–3007. http://doi.org/10.1098/rstb.2008.0076

2. Cenzer, M. L. (2016). Adaptation to an invasive host is driving the loss of a native ecotype. *Evolution*, (70), 2296–2307. http://doi.org/10.1111/evo.13023

3. Comeault, A. A., Flaxman, S. M., Riesch, R., Schwander, T., Slate, J., & Nosil, P. (2015). Selection on a genetic polymorphism counteracts ecological speciation in a stick insect. *Current Biology*, *25*, 1975–1981. http://doi.org/10.1016/j.cub.2015.05.058

4. Cooper, J. C., & Phadnis, N. (2016). A genomic approach to identify hybrid incompatibility genes. *Fly*, *10*, 142–148. http://doi.org/10.1080/19336934.2016.1193657

5. Coyne, J. A., & Orr, H. A. (2004). *Speciation* (First Edit). Sinauer Associates.

6. Fitzpatrick, B. M., Fordyce, J. A., & Gavrilets, S. (2008). What, if anything, is sympatric speciation? *Journal of Evolutionary Biology*, *21*, 1452–1459. http://doi.org/10.1111/j.1420-9101.2008.01611.x

7. Hollocher, H., Ting, C.-T., Pollack, F., & Wu, C.-I. (1997). Incipient speciation by sexual isolation in Drosophila melanogaster: variation in mating preference and correlation between sexes. *Evolution*, *51*(4), 1175–1181.

8. Johannesson, K. (2010). Are we analyzing speciation without prejudice? *Annals of the New York Academy of Sciences*, *1206*, 143–149. http://doi.org/10.1111/j.1749-6632.2010.05701.x

9. Koevoets, T., van de Zande, L., & Beukeboom, L. W. (2012). Temperature stress increases hybrid incompatibilities in the parasitic wasp genus Nasonia. *Journal of Evolutionary Biology*, *25*, 304–316. http://doi.org/10.1111/j.1420-9101.2011.02424.x

10. Llopart, A., Lachaise, D., & Coyne, J. A. (2005). An anomalous hybrid zone in Drosophila. *Evolution*, *59*, 2602–2607.

11. Matute, D. R., Butler, I. A., Turissini, D. A., & Coyne, J. A. (2010). A test of the snowball theory for the rate of evolution of hybrid incompatibilities. *Science*, *329*, 1518–1522.

12. Orr, H. A., Madden, L. D., Coyne, J. A., Goodwin, R., & Hawley, R. S. (1997). The Developmental Genetics of Hybrid Inviability: A Mitotic Defect in Drosophila Hybrids. *Genetics*, *145*, 1031–1040.

13. Rosenblum, E. B., Sarver, B. A. J., Brown, J. W., Des Roches, S., Hardwick, K. M., Hether, T. D., … Harmon, L. J. (2012). Goldilocks meets Santa Rosalia: an ephemeral speciation model explains patterns of diversification across time scales. *Evolutionary Biology*, *39*, 255–261. http://doi.org/10.1007/s11692-012-9171-x

14. Turissini, D. A., Liu, G., David, J. R., & Matute, D. R. (2015). The evolution of reproductive isolation in the Drosophila yakuba complex of species. *Journal of Evolutionary Biology*, *28*, 557–575. http://doi.org/10.1111/jeb.12588