THE GENETICS OF HYBRID INCOMPATIBILITY EARLY IN THE 'SPECIATION CONTINUUM'

by

ERIC T. WATSON

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

May 2014

Copyright © by Eric T. Watson 2014

All Rights Reserved

Acknowledgements

I would like to express my gratitude to my advisor, Dr. Jeffery Demuth, for his guidance and caring, for patiently correcting my writing and financially supporting my research, and providing me with an excellent atmosphere for doing research. I would also like to thank my committee members, Dr. Esther Betran, Dr. Cedric Feschotte, Dr. John "Trey" Fondon III, and Dr. Andre Pires Da Silva for guiding my research for the past several years and helping me to develop my background in molecular and population genetics, evolutionary biology, and genome biology. Special thanks goes to Dr. Todd Castoe, who was willing to participate in my final defense committee at the last moment.

I would also like to thank my parents, and sister, and my girlfriend Laura Drzich who all gave me their loving support. Many thanks to Nick Long, Diego Martin-Perez, Emmanuela Mujica, and other members of Dr. Demuth's laboratory for helping me with beetle maintenance and bench-work. My research would not have been possible without their kind help. Thanks to Heath Blackmon, Jacobo Reyes Velasco, and David Sanchez, who entertained many questions with patience and provided loving friendship. Very special thanks goes to Dr. Mohamed A.F. Noor, who was the catalyst for this journey. In my first semester as an undergraduate, he entertained my question "…so then what the heck is a species" when I described to him my attempt to "merge" two species of African cichlid fish in my home aquarium. To this day, I remain captivated by our first conversations and continue to devote my life to these fascinations

April 18, 2014

Abstract

THE GENETICS OF HYBRID INCOMPATIBILITY

EARLY IN THE 'SPECIATION CONTINUUM'

Eric T. Watson, PhD

The University of Texas at Arlington, 2014

Supervising Professor: Jeffery P. Demuth

Species are discrete groups of organisms that are reproductively isolated from other related groups of organisms. Variation in the strength of reproductive isolation between closely related taxa falls along a 'speciation continuum', from weakly isolated structured populations to fully isolated sister species. Efforts to understand the genetics of speciation have focused primarily on developmental defects in hybrids (intrinsic postzygotic isolation) leading to sterility or inviability. Once speciation is complete, species continue to diverge as a result of mutation, genetic drift, and selection, and new forms of reproductive isolation may continue to accumulate. Therefore, at later stages in the 'speciation continuum', it is unclear whether a given isolating barrier is a cause, or a consequence of speciation. In my dissertation, I focus on the genetics of intrinsic postzygotic isolation occurring early in the 'speciation continuum' at the onset of speciation. In chapter 1, I highlight where the assumptions of dominance theory are particularly problematic in marsupials, where X inactivation uniformly results in silencing the paternal X. I then present evidence of Haldane's rule for sterility but not for viability in marsupials, as well as the first violations of Haldane's rule for these traits among all mammals. Marsupials represent a large taxonomic group possessing heteromorphic sex

chromosomes, where the dominance theory cannot explain Haldane's rule. In this light, I evaluate alternative explanations for the pre- ponderance of male sterility in interspecific hybrids, including faster male evolution, X–Y interactions, and genomic conflict hypotheses. In chapter 2, I revisit three mechanisms highlighted by Rose and Doolittle (1983) as a convenient conceptual scaffold for understanding the variety of ways TEs might directly, or indirectly, cause reproductive incompatibility. In chapter 3, I describe an example of hybrid incompatibility (called "still") segregating in F_1 hybrids between populations of *T. castaneum*, whereby affected offspring exhibit a suite of maladaptive traits upon eclosion from the pupal stage. To investigate the genetic cause of the still phenotype, I sequenced the genomes of still and normal siblings using pooled-DNA and employed a genome scan approach that compares allele frequencies between extremely discordant sib pairs (still vs normal) to identify discordant alleles. In total, I identified 97 genes with significantly discordant non-synonymous SNPs between still and normal siblings. An additional 355 genes possess nucleotide changes that are either synonymous, or non-coding (i.e. occur in introns or within 1000kb upstream or downstream). Interestingly, a set of 19 candidate loci were recently identified as candidate phosphine resistance genes. Phosphine is an insecticidal fumigant which acts as a metabolic toxin by targeting redox reactions, and is used worldwide in grain storage and processing facilities. The Chicago population was collected over 7 decades aco, predating the use of phosphine, while Tanzania populations were potentially subjected to 30 years, or roughly 330 generations of routine phosphine exposure before it was collected and kept in the laboratory. I discuss this observation in light of the role of genetic conflict in generating hybrid incompatibilities, especially where they are still segregating early in the 'speciation continuum'.

Table of Contents

Acknowledgements iii

List of Illustrations ix

List of Tables xi

Chapter 1 Haldane's Rule In Marsupials: What Happens When Both Sexes Are Functionally Hemizygous? 1

Haldane's Rule Is A Pattern Based On Sex Chromosomes, Not Sex 1

Dominance Theory And The Role Of X Chromosomes In The Expression Of Haldane's Rule 2

Dosage Compensation, X Inactivation, And The Role Of Dominance In Haldane's Rule 4

Haldane's Rule For Sterility And Inviability In Hybrid Marsupials 7

Alternative Hypotheses For Haldane's Rule In Marsupials 8

Prospects For Future Research 11

Acknowledgements 12

Chapter 2 The Role Of Transposable Elements In Speciation 13

Introduction 13

Mechanism I: Genomic Disease 18

TE Suppression by Small RNAs: A Molecular Mechanism for Speciation 19

Drosophila Hybrid Dysgenesis 22

TE Derepression In Interspecific Animal Hybrids 24

TE Derepression In Plant Hybrids 26

Mechanism II: Mechanical Incompatibility 28

Evidence For Mechanism II 30

Structural Variation Caused By TE Activity 30

Structural Variations Associated With Reproductive Isolation 31
A TE Induced Structural Variant That Causes Isolation...................................... 32
Mechanism III: Genomic Resetting.. 33
Conclusion 35
Acknowledgements 35
Chapter 3 Dissecting The Genetic Basis Of Hybrid Incompatibility Within Species: A Neuromuscular Disorder At The Onset Of Speciation In *Tribolium castaneum* 36
Introduction.. 37
Methods... 40
Study Organism.. 40
Populations And Crosses 41
Genome Sequencing Of Discordant Siblings 42
Genome Sequencing Of Parent Populations 43
Extreme Discordant Sib-Pair Analysis (Pool-GWAS).. 43
Detecting The Signature Of Selection In Candidate Regions 44
Detection Of Structural Variation... 45
Results... 46
The ‘Still’ Hybrid Incompatibility Phenotype 46
Genome Scan For Discordant Alleles (Pool-GWAS) 47
Tests For Neutrality 50
Structural Variation... 61
Discussion 65
A Potential Role For Pesticide Resistance In Tribolium Speciation 65
Other Candidate Genes 68

The Potential General Role Of Genetic Conflict In Driving Hybrid Incompatibility Genes 69

Future Prospects For Speciation Research 71

Appendix A List Of Published Cases Of Marsupial Hybrids 73

Appendix B Potential Function Of Candidate Genes 77

Appendix C Distribution Of Discordant Snps Across Genomic Feature 81

Appendix D Mitochondrial Divergence Between Chicago And Tanzania Populations 83

References 85

Biographical Information 118

List of Illustrations

Figure 1-1 Expression of X-linked genes in therian mammals. Dosage inequality between the sexes is compensated in mammals by the inactivation of one X chromosome in females. Placental mammals inactivate either the maternal or paternal X chromosome at random, forming a mosaic of X-linked expression. Marsupial mammals inactivate only the paternal X chromosome, so that males and females only express the maternal X chromosome. Asterisk (*) indicates leaky expression... 6

Figure 2-1 Sense and anti-sense orientation of matching TEs are indicated by black and red arrows. Sense (A locus) and antisense (B locus) transcription are indicated by squiggle lines. Here the B locus produces antisense transcripys which initiate small RNA production and target silencing of the A locus. For simplicity only the haploid genome is illustrated except in hybrids.. 21

Figure 3-1 Distribution of the still phenotype in 48 reciprocal crosses with quantile plots. A total of 24 crosses were constructed with Chicago dams and Tanzania sires (a) and also with Tanzania dams and Chicago sires (b) with both replicates for each temperature pooled. Quantile plots indicate a range of 0-100% still in panel a, with a mean of 53% and median of 51% still, whereas panel (b) ranged from 0-27% with a mean of 7.8% and median of 7.6% still offspring per family. 49

Figure 3-2 Genome-wide scan of divergent SNPs between extremely discordant sib-pairs. 83,422 significantly differentiated SNPs mapped to the 10 assembled *T. castaneum* reference chromosomes, with 14,669 SNPs showing $F_{ST} \geq 0.29$. Candidate genes are located on the X chromosome (1 gene), chromosome 3 (11 genes), chromosome 8 (1 gene), chromosome 9, (45 genes), chromosome 10 (1 gene), and unassembled contigs (30 genes, not shown).. 52

Figure 3-3 Divergent SNPs mapping to Uns-7 scaffold, containing 19 predicted genes shown below the x-axis. 53

Figure 3-4 Structural variation across the genomes of hybrid siblings. Percent difference in proportional deviation (d_{pd}) was found by subtracting the measurement from still siblings from the normal siblings giving a positive value where fewer reads mapped in still, and negative values where fewer reads mapped to normal (see methods for proportional deviation calculation). 63

Figure 3-5 Models of hybrid incompatibility in a genomic conflict scenario. Circles represent derived alleles. In co-evolutionary models, DMIs are continually fixed either at the same loci (that is, two-locus, two-allele) or at different loci (that is, four-locus, two-allele) (see the table). In all examples with two substitutions in a lineage, the selfish locus (left) drives the evolution of the restorer locus (right). Grey arrows indicate negative epistatic interactions between complementary loci. In all models, the ancestral state is wild type except for the two-locus two-allele co-evolutionary model. In this model, the ancestral state is a co-evolving selfish element-restorer system; numbers represent the lineages in which the derived alleles originated. (from (Seehausen et al. 2014) 64

List of Tables

Table 1-1 Summary of studies of Haldane's rule in mammals. Data from eutherian mammals are from Schilthuizen et *al.* (2011). For a full list, see Appendix A.................... 7

Table 2-1 Mobile genetic elements and speciation in geologic time. Adapted from: Rebollo et al. 2010 and Kim et al. 2004.Transposable element activity in the following examples is concordant with phylogenetic activity of their hosts. Additional examples describe intraspecific diversity in numbers of TE families... 16

Table 3-1 List of candidate genes for 'still' hybrid incompatibility. Genes with asterisk (*) are on the Uns-7 scaffold recently mapped to chromosome 9. (POO: parent of origin [C=Chicago, T=Tanzania]).. 54

Chapter 1

Haldane's Rule In Marsupials:

What Happens When Both Sexes Are Functionally Hemizygous?

During the process of speciation, diverging taxa often hybridize and produce offspring wherein the heterogametic sex (i.e., XY or ZW) is unfit (Haldane's rule). Dominance theory seeks to explain Haldane's rule in terms of the difference in X-linked dominance regimes experienced by the sexes. However, X inactivation in female mammals extends the effects of hemizygosity to both sexes. Here, we highlight where the assumptions of dominance theory are particularly problematic in marsupials, where X inactivation uniformly results in silencing the paternal X. We then present evidence of Haldane's rule for sterility but not for viability in marsupials, as well as the first violations of Haldane's rule for these traits among all mammals. Marsupials represent a large taxonomic group possessing heteromorphic sex chromosomes, where the dominance theory cannot explain Haldane's rule. In this light, we evaluate alternative explanations for the preponderance of male sterility in interspecific hybrids, including faster male evolution, X–Y interactions, and genomic conflict hypotheses.

Haldane's Rule Is A Pattern Based On Sex Chromosomes, Not Sex

Of the few major patterns recognized in speciation, perhaps none has enjoyed as much attention as Haldane's rule. First noted by Haldane (1922), Haldane's rule describes the phenomenon that whenever divergent taxa produce hybrid offspring, the heterogametic (XY or ZW) sex suffers a reduction in fitness more often than the homogametic (XX or WW) sex. Several studies now demonstrate that Haldane's rule is a

preliminary, though perhaps not requisite, stage of speciation (Coyne and Orr 1989a; Sasa et al. 1998; Presgraves 2002; Demuth and Wade 2007; Malone and Fontenot 2008) and has come to be known as one of the ‘‘rules of speciation’’ (Coyne and Orr 2004; Demuth and Wade 2007; Turelli and Moyle 2007).

The study of Haldane’s rule has shaped our understanding of speciation by providing a broad pattern in which to study the mechanisms of population divergence and fixation of reproductive isolating barriers. Importantly, this pattern holds across the majority of taxa studied, with examples of male and female heterogamety both conforming to Haldane’s rule (Haldane 1922, 1932; Gray 1954, 1958; Hillis and Green 1990; Schilthuizen et al. 2011). For this reason, it is often noted that Haldane’s rule cannot simply be explained by the general sensitivity of one sex over the other (Coyne and Orr 1989b). Instead, Haldane’s rule is thought to be the result of genetic incompatibilities that are exaggerated in the genome of the heterogametic sex.

Dominance Theory And The Role Of X Chromosomes In The Expression Of Haldane’s Rule

Among the early genetic explanations for Haldane’s rule, Muller (1940, 1942) put forth the X-autosome interaction hypothesis noting that, owing to hemizygosity, hybrid males suffer from both dominant and recessive X-linked incompatibilities, whereas females only suffer from dominant incompatibilities (for convenience, we use male for the heterogametic sex and female for the homogametic sex). Orr (1993a) formalized Muller’s theory mathematically and added that hemizygosity is a ‘‘double-edged sword’’: although males express every X-linked incompatibility, on average females contain twice as many because they possess 2 X chromosomes. If dominant and recessive incompati- bilities

are equally likely, these 2 factors cancel each other and cannot explain the consistently lower fitness of males (Orr 1993a). However, if recessive alleles are more likely to produce severe incompatibilities, Haldane's rule will result (Muller and Pontecarvo 1942). Thus, somewhat ironically, the ''dominance theory'' relies on genetic incompatibilities being recessive on average. To further complicate matters, ''dominance,'' in the context of Haldane's rule, refers only to the X-chromosome component of what is more precisely an epistatic interaction between X-linked and other (perhaps multiple) loci (Demuth and Wade 2007).

Beginning in the 1980s, studies extending the dynamics of hemizygosity to females began to suggest forces in addition to dominance might contribute to Haldane's rule. For instance, when females carrying both X chromosomes from one parent in an otherwise hybrid genome (unbalanced females) were made from *Drosophila* species pairs that normally obey Haldane's rule for sterility, the unbalanced female hybrids remained fertile. However, if the species pair normally obey Haldane's rule for viability, unbalanced females became inviable (Coyne 1985; Orr 1993b). Later observations exploring Aedes mosquitoes that lack hemizygous sex chromosomes found related results. Hybrids follow Haldane's rule for sterility, but not viability (Presgraves and Orr 1998). Although the mechanisms by which dominance effects in females are made equivalent to males is different in the unbalance female and *Aedes* studies, the conclusions are the same; when both sexes have the same dominance effects, fertility conforms to Haldane's rule, but viability does not.

The unbalanced female and *Aedes* studies are instructive to the situation in mammals because female X-chromosome inactivation (XCI) results in only one X chromosome being expressed. Genetic explanations invoking dominance assume that chromosomal hemizygosity is equivalent to functional hemizygosity in terms of gene

expression (Turelli and Orr 1995). Although this assumption is valid for *Drosophila*, where X-chromosome dosage compensation is achieved by the hypertranscription of the hemizygous X in males to equal the dosage expected of diploid autosomes and/or female Xs (Lucchesi 1973), it is clear that dosage compensation is not similarly achieved in other taxa (e.g., *Caenorhabditis elegans* and therian mammals—Xiong et al. 2010; *Anopheles*—Hahn and Lanzaro 2005; birds—Itoh et al. 2010; *Lepidoptera*—Zha et al. 2009; stickleback—Leder et al. 2010; platypus—Deakin et al. 2009; *Tribolium*—Prince et al. 2010; reviewed in Mank et al. 2011). Indeed, in their original formulation of dominance theory, Turelli and Orr (1995) asked, ‘‘Does the dominance theory work given mammalian dosage compensation?’’ Dosage compensation in mammals is particularly problematic for dominance theory because it involves XCI in females wherein one X chromosome is transcriptionally silenced (i.e., females are functionally hemizygous). Conse- quently, the average transcript ratio from XX:AA females is approximately 0.5—equal to the ratio in X:AA males (Gupta et al. 2006; but see Nguyen and Disteche 2006; Xiong et al. 2010). Following, we revisit Turelli and Orr’s question highlighting data from marsupial hybrids that have not previously been appreciated for what they may tell us about the genetic mechanisms underlying Haldane’s rule.

Dosage Compensation, X Inactivation, And The Role Of Dominance In Haldane’s Rule

XCI is achieved by different means in metatherian (marsupial) and eutherian (placental) mammals. In placental mammals, one copy of the X is randomly inactivated, forming a mosaic of maternal and paternal X-chromosome expression (Lyon 1961). In marsupial cells, males and females, both experience functional hemizygosity of the same set of alleles because it is always the paternal X chromosome that is inactivated (Cooper

et al. 1971; Richardson et al. 1971; Al Nadaf et al. 2010). The consequences of mosaic XCI for dominance theory and Haldane's rule in placentals depend on the degree of autonomy among cells, which is largely unknown in mammals. However, the consistent hemizygous expression of only the maternal X chromosome in marsupials has clear implications, providing a situation where, if Haldane's rule is observed, dominance theory cannot be the explanation.

First, and perhaps more importantly, dominance theory assumes that all loci have diploid expression in F_1 females, as is the case in *Drosophila* (Turelli and Orr 1995; Orr and Turelli 1996; Turelli and Orr 2000). Because hemizygous expression of the maternal X chromosome is shared in both sexes in marsupials (Figure 1-1), dominance effects are the same in males and females. Hence, under dominance theory, hybrid males and females are expected to suffer the same expected reduction of fitness due to X-linked incompatibilities, and Haldane's rule is not expected to consistently result. Additionally, the idea that females should suffer twice the average number of X-linked incompatibilities of males (Orr 1993a) is moot, if both sexes express the same X-chromosome complement. In sum, because X-linked alleles with strict paternal XCI are never functionally heterozygous, there is no X-linked dominance in either sex, and dominance theory predicts that Haldane's rule should not hold in marsupial hybrids.

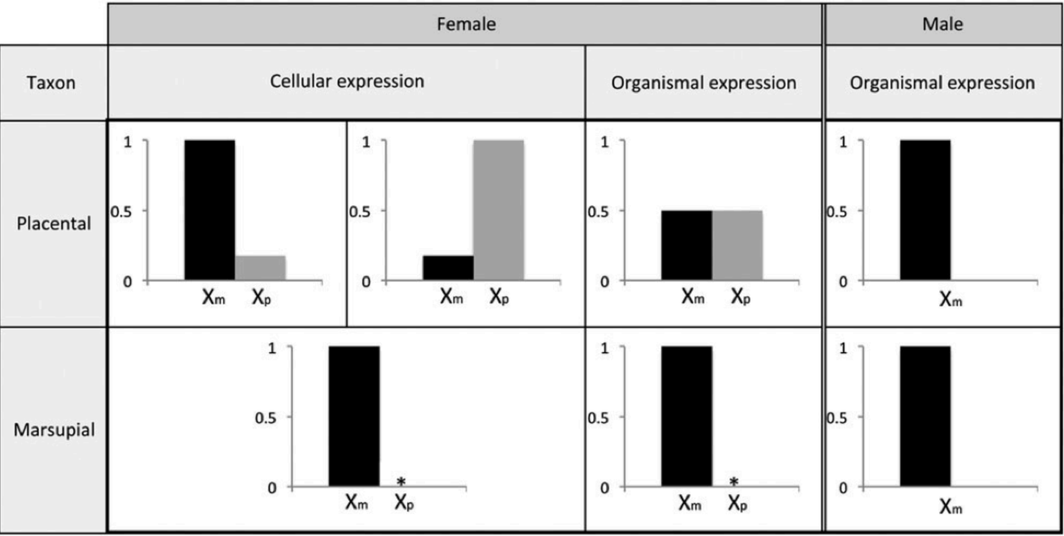


Figure 1-1 Expression of X-linked genes in therian mammals. Dosage inequality between the sexes is compensated in mammals by the inactivation of one X chromosome in females. Placental mammals inactivate either the maternal or paternal X chromosome at random, forming a mosaic of X-linked expression. Marsupial mammals inactivate only the paternal X chromosome, so that males and females only express the maternal X chromosome. Asterisk (*) indicates leaky expression.

Table 1-1 Summary of studies of Haldane's rule in mammals. Data from eutherian mammals are from Schilthuizen et *al.* (2011). For a full list, see Appendix A

Group	Asymmetric phenotype	Crosses with asymmetric effects	Crosses obeying Haldane's rule	Percentage of obeying
Eutheria	Sterility	34	34	100.0
	Inviability	5	5	100.0
Metatheria	Sterility	7	6	85.7
	Inviability	2	1	50
Combined	Sterility	41	40	97.6
	Inviability	7	6	85.7

Haldane's Rule For Sterility And Inviability In Hybrid Marsupials

Despite the prediction from dominance theory, published accounts show that Kangaroos (Macropus), Rock Wallabies (*Petrogale*), and Pademelons (*Thyogale*) obey Haldane's rule for fertility in most cases (Close and Lowry 1990 and references therein; Sharman et al. 1990; Eldridge and Close 1993). In 10 of 11 species pairs, males are sterile while females are fertile. However, in the remaining species pair, females are sterile while males remain fertile, representing the only reported exception to Haldane's rule for fertility among mammals (Table 1-1). When corrected for nonindependence among species pairs, the number obeying Haldane's rule becomes 6 of 7 for sterility. Incomplete data for an additional 12 species pairs also suggest that fertility is more frequently disturbed in males than in females (Appendix A).

Despite the paucity of the female viability data reported, it remains possible to draw conclusions about Haldane's rule for inviability. The probability of having exactly k offspring of a particular sex is (n k)p^k, where n is the total number of offspring observed

and $p = 0.5$. Two species pairs exhibit a >95% chance that one sex is rare or absent (*Petrogale assimilis* x *P. penicillata*, and *P. i. mareeba* x *P. assimilis*). The first species pair produced 9 females and 1 male, conforming to Haldane's rule for inviability, whereas the second species pair produced 8 males and no females—a violation of Haldane's rule for inviability (Sharman et al. 1990). Interestingly, the reciprocal to this cross produced 6 females and 1 male, following Haldane's rule, although the result is nonsignificant ($p = 0.05469$). The majority of other interspecific crosses do not show significant asymmetry in viability between sexes ($0.5 \leq p \leq 0.05469$).

In placental mammals, Haldane's rule is obeyed in all 34 species pairs that produce sterile hybrids and in all 5 species pairs that produce only viable female hybrids (Laurie 1997; Schilthuizen et al. 2011). Marsupial crosses (reviewed in Close and Lowry 1990; Sharman et al. 1990; Eldridge and Close 1993) bring the overall therian mammal total for hybrid sterility to 40 of 41 species pairs and potentially add support for an additional 12 species pairs where data are limited (Appendix A). Although marsupials contain the only exceptions to Haldane's rule for both sterility and inviability among mammals, 52 of 53 hybridizations obeying the rule is still overwhelming support. If, as we propose, dominance theory cannot explain Haldane's rule in marsupials, why does the pattern still hold so regularly?

Alternative Hypotheses For Haldane's Rule In Marsupials

A key insight following the unbalanced female experiments pointed out that the genetic basis for viability is likely to involve the same set of loci in both sexes, whereas the loci governing fertility are probably different in males and females (Wu and Davis 1993). Since then, many evolutionary biologists have viewed Haldane's rule as a

composite phenomenon (Coyne 1992; Johnson et al. 1992; Orr 1993b; Wu and Davis 1993). This recognition, along with the observation that male sterility evolves faster than female sterility and faster than inviability in both sexes despite male fertility being less sensitive than viability to mutagenic disruption, led Wu and colleagues to propose that either sexual selection may drive rapid evolution of genes that contribute to male sterility or spermatogenesis may be inherently more sensitive than oogenesis to perturbation (Wu and Davis 1993, Wu et al. 1996). This so called ‘‘faster male’’ hypothesis has since been supported by diverse lines of evidence in plants and animals (Brothers and Delph 2010; Schilthuizen et al. 2011), including rapid evolution of male reproductive proteins by positive selection in placental mammals (Torgerson et al. 2002; Swanson et al. 2003; Clark and Swanson 2005; Good and Nachman 2005; Khaitovich et al. 2005). The data for marsupials is thus far consistent with the composite view of Haldane’s rule. Evidence for Haldane’s rule for viability is lacking, as expected under dominance theory with strict paternal XCI. Evidence for Haldane’s rule for sterility is abundant, which is consistent with faster male evolution. Future studies of marsupial reproductive protein evolution may provide additional support for faster male evolution.

Importantly, dominance and faster male theories need not be mutually exclusive. Indeed, Turelli and Orr (2000) discuss their relative roles under the same mathematical framework, where the influence of dominance scales with the proportion of the genome that is X linked, and the role of faster male evolution scales as the relative number and severity of male versus female incompatibilities. In marsupials, this interplay between dominance and faster male theories may remain, depending on the degree to which paternal X inactivation is leaky. Unfortunately, detailed mechanistic understanding of paternal XCI is still poorly understood. The most detailed study to date shows that paternal alleles escape XCI in 5–65% of cell lines (Al Nadaf et al. 2010). However, it

remains unclear what proportion of transcripts at the surveyed loci belonged to the paternal X (i.e., it is unknown whether paternal alleles ever attain full expression) and furthermore, escape from inactivation may be stochastic (Al Nadaf et al. 2010).

An additional source of incompatibilities that may explain Haldane’s rule in marsupials includes X–Y interactions, which have been suggested as a possible cause of male sterility in marsupials (Graves 1996; Graves and ONeill 1997). In placental mammals, proper meiotic pairing of the X and Y is facilitated by pseudoautosomal regions (PARs). Disruption of pairing in the PAR blocks meiosis and results in male infertility due to abnormal sperm development (Burgoyne et al. 1992) and is a suggested explanation for Haldane’s rule in placental mammals (Graves 1996). However, because marsupials do not possess a PAR region (X and Y pair at the tips in the absence of homology), X–Y interactions are suggested to be genic (Sharp 1982; Graves and ONeill 1997). In hybrid males, the X and Y chromosomes are derived from different species, and heterospecific interactions or loss of gene complement may contribute to Haldane’s rule for fertility in marsupials.

Additionally, genomic conflict, in the form of competition among oötids for inclusion into the pronucleus, potentially plays a role in Haldane’s rule in marsupials (reviewed in McDermott and Noor 2010). Centromeric sequences involved in spindle fiber attachment have been shown to be involved in such competition in mammals (Henikoff et al. 2001; Pardo-Manuel de Villena and Sapienza 2001). Separated by 1–2 My (Gifford et al. 2005), 2 wallaby sister species*, Macropus rufogriseus* and *M. eugenii*, differ greatly in the repeat content of their centromeric sequences (O’Neill et al. 1998; Metcalfe et al. 2007). Interspecific crosses between these species produce infertile male and female hybrids that display extensive chromosomal remodeling and genomic instability, for example, changes in chromatin structure and the amplification of satellite repeats and

transposable elements (Metcalfe et al. 2007). The effects of genomic instability may contribute to Haldane's rule in marsupials if centromeric misalignment of the X and Y chromosomes during metaphase in hybrids leads to the failure of spermatogenesis (McKee 1997; Zwick et al. 1999; Henikoff et al. 2001). Furthermore, meiotic inactivation of sex chromosomes in male hybrids, a process crucial for male fertility in mammals (Royo et al. 2010), could be delayed or derailed by the decondensation or amplification of centro- meric regions. In marsupials, meiotic sex chromosome inactivation occurs before the X–Y associations that lead to the formation of the sex chromatin beginning at mid-pachytene (Namekawa et al. 2007). A delay in the formation of the sex chromatin may trigger the late-pachytene meiotic checkpoint and lead to spermatocyte apoptosis and reduced fertility, a phenomenon attributable to chromosomal asynapsis in placental mammals (Luan et al. 2001). If true, this mechanism would be consistent with more general ‘‘faster heterogametic sex’’ hypotheses that propose the XY sex evolves faster because of the conflicting pressures that the X and Y chromosomes experience to distort the sex ratio (Frank 1991; Hurst and Pomiankowski 1991; Tao and Hartl 2003).

Prospects For Future Research

In marsupial mammals, nature provides us with a system analogous to the unbalanced female experiments in *Drosophila* (Coyne 1985; Orr 1993b) and *Aedes* mosquito lacking hemizygous sex chromsomes (Presgraves and Orr 1998). In each case, the X-chromosome contribution to reproductive isolation is the same in males and females. In most species, dominance theory and faster male theory cannot be disentangled to reveal the cause of hybrid male sterility where they act simultaneously (Wu et al. 1996; Coyne and Orr 2004). However, because dominance theory cannot

explain Haldane's rule in marsupials, understanding the genetics of hybrid male sterility and the evolutionary dynamics of sex chromosomes in this large group of diverse organisms will provide useful insight into one of the most sweeping empirical observations in evolutionary biology.

A major current limitation is that no records exist for hybridizations in non-macropodid marsupials such as opossums and possums (Didelphidae and Caenolestidae), gliders (Petauridae), bandicoots (Peramelemorphidae), and marsupial moles (Notoryctidae). While macropods, such as kangaroos, typically produce one offspring per season, many species in these families are highly fecund and produce anywhere from 4 to 10 offspring in a litter making them highly amenable to studying biases in sex ratio. Future research involving marsupials with high fecundity will potentially provide excellent candidates for studying the evolution of hybrid male sterility in nonplacental mammals.

Acknowledgements

We thank Maeli Mellotto, Esther Betrán, and Heath Blackmon for their discussion and assistance in improving the manuscript.

Chapter 2

The Role Of Transposable Elements In Speciation

With innovative DNA sequencing technologies has come a new appreciation for the content of animal and plant genomes. Overwhelmingly, a picture has been painted in which mobile and repetitive elements dominate the genomic landscape. Mobile genetic elements have recently been shown to contribute coding and regulatory sequences during their proliferation, leading to functional and regulatory novelty as well as element-mediated rearrangements coinciding with speciation events. Additionally, dormant elements occasionally erupt in bouts of excision and transposition in interspecific hybrids, resulting in a suite of maladaptive traits. The potential for mobile elements as key players in the evolution and diversification of genomes and species is immense yet in many respects transposable elements still remain the “dark matter” in the genome. This is particularly true of their role in speciation, and in order to fully appreciate their role, much work is still needed. In this chapter, we investigate the evidence for transposable elements as drivers of diversification and speciation.

Introduction

Transposable elements (TEs) are a diverse group of genetic sequences that share the common ability to move within a genome. The broadest distinction among TEs classifies them based on whether or not transposition involves an RNA intermediate (Wicker et *al*., 2007). Retrotransposons (class I TEs) transpose through a replicative “copy and paste” mechanism, which involves the production of a processed mRNA transcript that becomes re-inserted into the host genome after being reverse transcribed into complimentary DNA by an element-encoded reverse transcriptase. In contrast, DNA

transposons (class II TEs) rely on a non-replicative “cut and paste” mechanism, involving a diversity of element-encoded enzymes, such as transposase, C-integrase, and tyrosine recombinase. Both class I and class II TEs are prone to proliferative bursts and exist as either autonomous elements, which encode the proteins necessary to catalyze there own transposition; or non-autonomous elements, which use the replication machinery of autonomous TEs.

TEs are often considered “parasitic” and “selfish” due to their ability to invade a genome despite imposing a fitness cost to their hosts in the process. By inserting themselves into coding or regulatory regions, TEs often have deleterious consequences, but their impact varies among taxa. For instance, TE mobilization is associated with over 100 human diseases (Goodier and Kazazian 2008; Belancio, Hedges, and Deininger) but this represents a relatively low percentage of pathogenic mutations (~0.3%; (Callinan and Batzer 2006; Kazazian 1998). In mouse and *Drosophila*, TE insertions constitute a much larger proportion of deleterious mutations (10% and 50% respectively; Maksakova et *al.* 2006; Finnegan 1992). The variation in the fitness consequences of TEs likely reflects an interaction between host genome defense and variation in the composition of TE types (Eickbush and Furano 2002).

While transposable elements (TEs) are perhaps best known for their capacity to disrupt host gene function, speculation about their potential evolutionary benefits as drivers of diversity have been around almost since their discovery in the 1940s (McClintock 1950; 1956; Britten and Davidson 1971). However, it was not until the genomic era that the ubiquity, abundance, and diversity of TEs has been fully appreciated (Kidwell and Lisch 2001; Fedoroff 2012). For example, variation in TE abundance between the genomes of different species reveals lineage specific dynamics in the composition and abundance of different classes of TEs. Recent large-scale

sequencing projects and advances in the screening of genomic sequences for hallmarks of repetitive elements reveal that TEs comprise 10-90% of plant nuclear genomes (Li et *al*. 2004; Haas et *al.* 2005), up to 20% of fungal genomes (Wicker et *al*. 2007), 9% of the chicken genome (International Chicken Genome Sequencing Consortium 2004), 15-22% of the *Drosophila melanogaster* genome (Kapitonov and Jurka 2003;Biemont and Vieira 2005), 6-17% of the *Tribolium castaneum* genome (Wang et *al*. 2008), and 50-66% of the human genome (International Human Genome Sequencing Consortium 2001; Koning et *al*. 2011). However, given that identifying TEs and their copy number from genomic sequence is fraught with computational difficulties, some of these values likely remain underestimates (e,g., Koning et *al.* 2011). Across eukaryotes, TEs are a much better predictor of genome size than protein coding gene content (Feschotte and Pritham 2007a). In addition to their abundance, a handful of well-annotated genomes illustrate that the TEs present in eukaryotes are very diverse (Wicker et *al*. 2007; Mandal and Kazazian 2008; Venner et *al*. 2009; Jurka et *al*. 2011; Chénais et *al*. 2012).

The diversity of genetic novelty infused by TE invasion and proliferation has been a fundamental argument for their role in driving organismal diversity (reviewed in Fedoroff 1999; Kazazian 2004; Biémont and Vieira 2006; Oliver and Greene 2009; 2011) and there are a growing number of specific examples where TE domestication has contributed coding and/or regulatory sequences implicated in adaptive novelty (Brandt et *al*. 2005; Feschotte and Pritham 2007b; Feschotte 2008; Butter et *al*. 2010). Furthermore, the burst dynamics of TEs wherein they invade, rapidly proliferate, and then are silenced by host defense has been hypothesized to explain broad evolutionary patterns from geological time scales to contemporary biodiversity (Zeh et *al*. 2009; Oliver and Greene 2011; Table 2-1).

Table 2-1 Mobile genetic elements and speciation in geologic time. Adapted from: Rebollo et al. 2010 and Kim et al. 2004.Transposable element activity in the following examples is concordant with phylogenetic activity of their hosts. Additional examples describe intraspecific diversity in numbers of TE families

TE event/Species history	**Reference**
Reduced L1 and SINE accumulation during radiation of African apes (14–15 Mya).	Intl Hum Gen Seq Consort (2002)
Expansion of L1 subfamilies parallels intense speciation in *Rattus sensu stricto*.	Verneau et *al*. (1998)
Lx family amplification is concomitant to the radiation of murine mammals.	Pascale et *al*. (1990)
Rapid speciation in the genus *Taterillus* (gerbil) occurred along with intense activity of TEs in nascent lineages.	Dobigny et *al*. (2004)
Intense DNA elements transposition during the *Myotis* radiation.	Ray et *al*. (2008)
DNA transposon bursts parallel speciation events in pseudotetraploid salmonids and occurred after genome duplication.	de Boer et *al*. (2007)
Acquisition and consequent transposition of an endogenous retrovirus (ERV) element and lineage specific enrichment of TEs in *Entamoeba histolytica*.	Lorenzi et *al*. (2008)
Repeated bouts of Haplochromine-specific SINE insertions followed by extensive radiations found in all inhabited lakes.	Shedlock et *al*. (2004)
Peak L2 and MIR activity coincides with marsupial-placental split 120-150 MYA.	Kim et *al*. 2004
Peak L1 activity corresponds to the eutherian radiation 100MYA.	Kim et *al*. 2004
Unprecedented LTR activity on the Y-chromosome corresponds to the K-T ecological disturbance 70 MYA	Kim et *al*. 2004
Alu and young L1 activity is restricted to the radiation of Old World and New World monkeys 40 and 25 MYA, respectively.	Kim et *al*. 2004
Diverse species exhibit different numbers of TE families between subpopulations with relatively low amounts of sequence divergence (0-561 families at < 1% and 5-1093 families at < 5% divergence)	Jurka et *al*. 2011
Tourist-like MITE, miniature Ping (mPing), present in 14 copies in *Oryza indica* and 70 copies in *O. japonica*.	Jiang et *al*. 2003
Ty3/*gypsy*-like retrotransposon expansion in hybrid species *H. anomalus*, *H. deserticola*, and *H. paradoxus* occurred near the time of their origin 0.5 to 1 Mya.	Ungerer et *al*. 2009

Despite evidence suggesting a role for TEs as drivers of organismal diversity, it does not necessarily follow that TE sare important to the process of speciation. In sexually reproducing species, speciation is the process of converting segregating variation within a species to fixed differences between species through the evolution of reproductive isolation (Dobzhansky 1937; Mayr 1942). Darwin's theory of natural selection, while providing an elegant mechanism for adaptation, did not fully explain how speciation *per se* occurs. The logical difficulty that arises is that selection should weed out variants causing reduced fitness in conspecific matings (Orr 1996). For this reason, even if we assume that TEs are broadly important to biodiversity, it is also worthwhile to address whether they might play a special role in generating isolation.

Thirty years ago Rose and Doolittle (1983) authored a paper in *Science* titled "Molecular Biological Mechanisms of Speciation" in which they examined the empirical evidence for repetitive DNA's role in the formation of reproductive isolation. It marked an early synthetic effort to relate TEs to the process of speciation. In their review, the mechanisms by which TEs might facilitate the origin of species were subdivided into three general categories: I) Genomic Disease, II) Mechanical Genome Incompatibility, and III) Genome Resetting. Based primarily on what was then recent data demonstrating hybrid dysgenesis between P and M strains of *Drosophila melanogaster*, Rose and Doolittle concluded that Genomic Disease, while seemingly least plausible, had more empirical support than the other two categories. We have learned much in the intervening three decades and in this chapter we will revisit the categories of Rose and Doolittle as a convenient conceptual scaffold for understanding the variety of ways TEs might directly, or indirectly, cause reproductive incompatibility.

Mechanism I: Genomic Disease

The view that TEs are genomic parasites lends itself naturally to the hypothesis that antagonistic coevolution between disease (the TEs) and immunity (the host's genomic defenses) might promote speciation. Much like how B and T cells of the vertebrate adaptive immune system recognize and remember specific pathogens, a host's need to suppress the deleterious effects of selfishly proliferating TEs may drive specificity of host genomic defenses against their particular compliment of TEs (Ironically, TE domestication contributes to the V(D) J recombination system that makes adaptive immunity to pathogens possible; Market and Papavasiliou 2003; Zhou et *al*. 2004; Kapitonov and Jurka 2005; reviewed in Litman et *al*. 2010). Consequently, populations evolving in allopatry that are exposed to different TE pressures may diverge in genomic defense as well.

There are a few ways this genomic disease model could result in reduced fitness for F_1 hybrids. First, since F_1s are haploid for both parental genomes, to the extent that defense mechanisms are haploinsufficient, TEs from both parental types may escape suppression. Second, if the contribution to defense is inherited primarily from one parent but not the other, TEs from the non-contributing parent might be freed from suppression. Uniparental inheritance of defense is particularly intriguing because it predicts asymmetries in hybrid breakdown. For instance, it is straightforward to imagine that uniparental defense could result in "Darwin's corollary", which observes that hybridization barriers are often asymmetric (i.e. the degree of isolation depends on which population is the maternal vs paternal parent; see Turelli and Moyle 2007). If a population that has adapted to invasion by a novel TE hybridizes with a naïve population, hybrids in only one direction will suffer. Additionally, since sex-limited chromosomes (Y or W) tend to

accumulate TEs disproportionately relative to the rest of the genome (Charlesworth et *al*. 1994; Abe et *al*. 2000; Bachtrog 2005; Charlesworth et *al*. 2005) the heterogametic sex may be more likely to suffer in the F_1. For example, if XX mothers contribute to defense, XY hybrid sons might be affected disproportionately because they will not be guarded against TEs originating from their father's Y chromosome. The expected pattern of heterogametic F_1 hybrids suffering disproportionately is consistent with a "rule of speciation", Haldane's rule (Haldane 1922).

TE Suppression by Small RNAs: A Molecular Mechanism for Speciation

While it is not the only mechanism by which TE activity can be controlled, the small RNA "immune response" to TEs represents an ancient, pan-eukaryotic, genomic defense against TE mobilization, and provides the kind of host-pathogen specificity necessary for antagonistic coevolution (reviewed in Aravin et *al*. 2007; Girard and Hannon 2008; Malone and Hannon 2009; Michalak 2009; Bourc'his and Voinnet 2010; Castillo and Moyle 2012). Small RNA-mediated TE control happens in three basic phases: detection, amplification, and repression (Girard and Hannon 2008). First, the detection phase relies on TE's propensity to produce anti-sense, double-stranded, or aberrant RNAs, which are recognized by core RNAi machinery (e.g. Dicer RNase III family proteins) and are cleaved into small RNAs such as endogenous siRNAs (siRNAs). During the amplification phase, the primary pool of siRNAs is copied by an RNA dependent RNA polymerase and associated with Argonaute proteins (Ghildiyal et *al*. 2008). These siRNA - Argonaute complexes then recognize complimentary RNA sequences, guiding cleavage of additional transcripts. Depending on the taxon and tissue, TE repression is ultimately achieved by a combination of transcript degradation

(post-transcriptional silencing) and/or siRNA target directed methylation and/or histone modifi- cation (transcriptional silencing).

In animals, there is an additional detection mechanism that provides protection to the germline and takes advantage of TE's unique ability to move within the genome, the piwiRNA pathway. Rather than relying on Dicer dependent dsRNA recognition as above, detection begins with antisense transcription of TEs that have transposed into special RNA gene clusters (e.g. the flamenco locus in *Drosophila* contains sequences that repress Gypsy, Idefix and ZAM family TEs; Pelisson et *al*. 1994; Prud'homme et *al*. 1995; Sarot et *al*. 2004). Processed transcripts from these clusters, called piRNAs, associate with members of the Piwi subfamily of Argonaute proteins (e.g. *Drosophila*: Piwi, Aubergine, AGO3; Mouse: Miwi, Mili, Miwi2). Amplification occurs by a "ping pong" cycle wherein the antisense piRNAs direct cleavage of sense strand TE RNAs which also associate with Piwi proteins that then target cleavage of additional antisense piRNAs…and so on. Post-transcriptional and transcriptional repression is ultimately achieved similarly to that of the siRNA pathway above.

While we have provided only the coarsest overview, it should be clear that the genomic disease. Immune response analogy is an apt one despite predating discovery of small RNA mediated TE control by more than a decade (Bingham et *al*. 1982; Rose and Doolittle 1983; Ginzburg et *al*. 1984). Breakdown at any of the three stages of small RNA mediated TE control could result in hybrid problems. At detection, uniparental inheritance of small RNA is predicted to be particularly problematic, and haploinsufficiency could follow from problems in the amplification phase. Inadequacy of either of these phases would be expected to result in breakdown of repression.

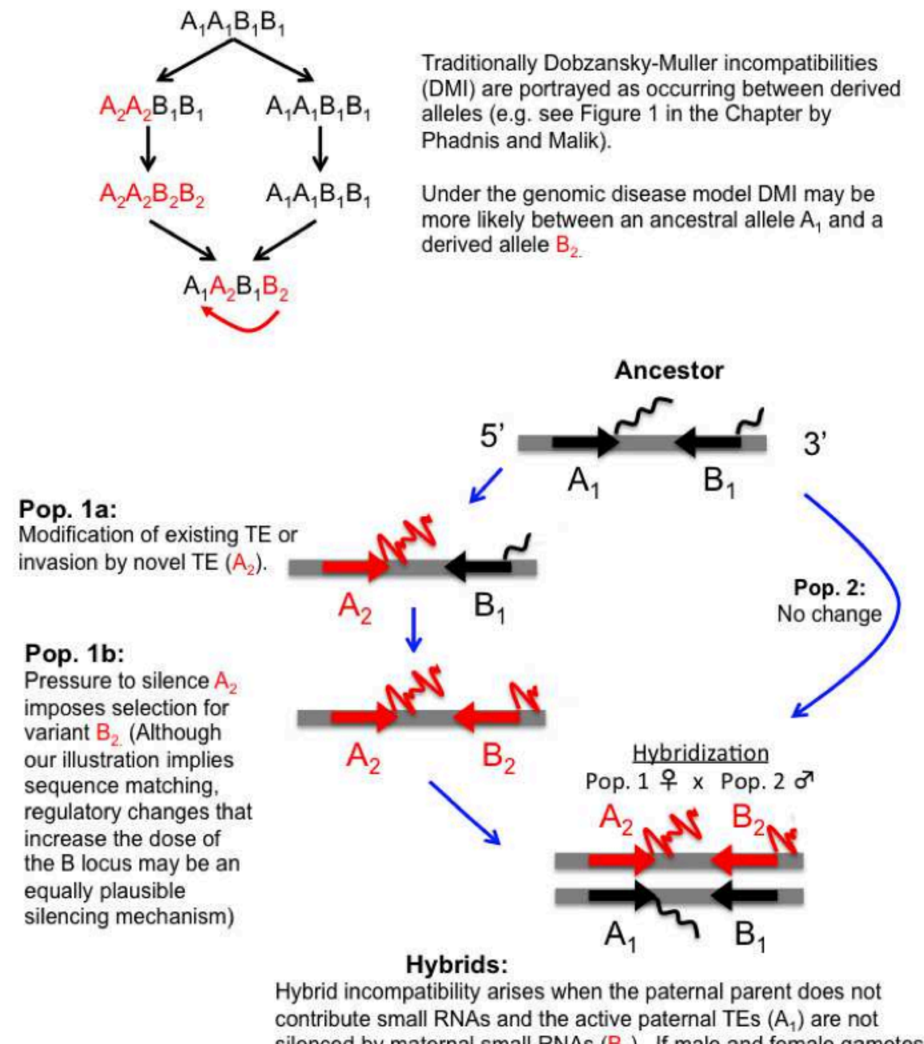


Figure 2-1 Sense and anti-sense orientation of matching TEs are indicated by black and red arrows. Sense (A locus) and antisense (B locus) transcription are indicated by squiggle lines. Here the B locus produces antisense transcripys which initiate small RNA production and target silencing of the A locus. For simplicity only the haploid genome is illustrated except in hybrids.

If we further abstract the genomic disease model to a two locus genetic interaction model wherein A_1 and A_2 represent TE variants, and B_1 and B_2 represent matching variants at a host defense locus, the history of work surrounding the Dobzhansky-Muller (DM) model of post- zygotic isolation become immediately relevant (Figure 49). While a complete reconciliation of the genomic disease model with DM is outside our scope, it is a potentially interesting avenue for future work. The mathematical framework surrounding DM is well developed (e.g. Turelli and Orr 2000; Orr and Turelli 2001; Demuth and Wade 2005) as is the population genetic theory for TE proliferation (Charlesworth and Charlesworth 1983; Charlesworth and Langley 1989; Ribeiro and Kidwell 1994; Brookfield and Badge 1997), reconciling the two may facilitate more specific predictions about the strength of asymmetry and consequences of lineage specific divergence.

Evidence for Mechanism I

Drosophila Hybrid Dysgenesis

Early evidence (in fact motivation) for the genomic disease model came from two independent systems in *Drosophila* (P-M and I-R) where crossing different strains of *D. melanogaster* resulted in hybrids with a suite of maladaptive traits such as sterility, gonad hypertrophy, extensive chromosomal aberrations, male recombination, and elevated germline mutation (Picard and L'Héritier 1971; Kidwell and Kidwell 1975; Picard 1976; Engels and Preston 1979; Schaefer et *al.* 1979; Hiraizumi et *al*. 1973; Yamaguchi, 1976; Kidwell et *al*. 1977; Woodruff and Thompson, 1977; Thompson et *al*. 1978) These studies also revealed that age and rearing temperature impact the degree of hybrid phenotypes

(Picard 1976; Bucheton et *al*. 1976) and collectively the consequences of these crosses was termed “hybrid dysgenesis” (Kidwell et *al*. 1977).

The underlying cause of hybrid dysgenesis was eventually traced to TE activity in offspring of P x M crosses and I x R crosses (Pelisson 1981; Rubin et *al*. 1982; Bingham et *al*. 1982; Kidwell 1983; Bucheton et *al*. 1984). In both cases strains more recently collected from the wild had active TEs that were not present in long-time lab strains. Hybrid dysgenesis only occurred in these crosses when the paternal parent carried the autonomous TE and the maternal parent did not. In the I-R system dysgenesis only arises when inducer (I) strain males are crossed to reactive (R) strain females. In the P-M system dysgenesis only occurs when paternal (P) strain males are crossed with maternal (M) strain females (Kidwell et *al*. 1977; Bingham et *al*. 1982; Kidwell 1983). The asymmetry in consequences of these crosses, particularly I-R sterile hybrid females, suggested that a maternal factor is responsible for offspring’s ability to control inherited TEs (Bregliano et *al*. 1980). Discovery that members of the Piwi subfamily of Argonaute proteins are maternally loaded into the pole plasm that will give rise to the future germline, and are necessary for TE silencing (Reiss et *al*. 2004; Sarot et *al*. 2004), added to early evidence for maternal inheritance of TE repression (Jensen et *al*. 1999). Later, small RNAs were also shown to be inherited maternally (Blumenstiel and Hartl 2005) and the protein-RNA complexes are responsible for TE silencing through females. Since the cytoplasm is mostly discarded from sperm there is little opportunity for male transmission of small RNAs. Thus, the hybrid dysgenesis observed in P-M and I-R hybrids appears to result from uniparental (maternal) inheritance of piRNA based TE control and the consequent inability to silence paternally derived TEs (Brennecke et *al*. 2008). Furthermore, maternal piRNA mediated suppression of TEs in the germline appears to be widely conserved among animals (reviewed in O'Donnell and Boeke 2007).

TE Derepression In Interspecific Animal Hybrids

Even before the mechanisms of small RNA based TE silencing were well worked out, researchers sought to test predictions of TE-mediated speciation in a comparative framework by looking for similar phenotypes in additional *Drosophila* hybrids (Coyne 1986; Hey 1988; Coyne 1989). These experiments assayed hybrids from isofemale lines derived from natural populations, as well as hybrids between natural and laboratory mutant lines. Enthusiasm over the role of TEs in speciation subsided, when it appeared that hybrid dysgenesis was not a ubiquitous feature of interspecific hybrids from the *D. melanogaster* or *D. affinis* subgroups (Coyne 1985; 1986; 1989; Eanes et *al*. 1988).

Additionally, even in *Drosophila* systems where hybrid dysgenesis occurs, sterility often varies with temperature and age. This provides an opportunity for previously naïve populations to adapt to new elements; perhaps mirroring the invasion of novel TEs within populations. For example, by repeatedly crossing males from an I strain to older dysgenic females that had regained some fertility, Pélisson and Bregliano (1987) were able to introduce repression of I element proliferation in a previously reactive genome within 15 generations. Based primarily on the *Drosophila* findings, Coyne (1986) suggested that to demonstrate a convincing case for TEs causing speciation requires:

- a difference in the distribution of element families among species,
- that these differences cause reproductive isolation (temperature-sensitive gonadal dysgenesis or elevated mutation and recombination rates do not themselves result in isolation)

- the existence of reciprocal sterility effects of the elements in cases where sterility occurs in reciprocal crosses between species.

Despite early pessimism, new findings from *D. melanogaster* x *D. simulans* hybrids have helped fuel renewed interest in TEs role in speciation. For example, Kelleher et *al*. (2012) recently showed that in contrast to intraspecific *D. melanogaster* crosses where maternal transmission of piRNAs fails to silence active paternal TEs, the interspecific hybrids have widespread activation of both maternally and paternally derived TEs. The study also showed that hybrid offspring are phenotypically most similar to flies with mutations in piRNA pathway genes (e.g. *Piwi*, *Aubergine* and *Argonaut3*). Ten proteins in the piRNA pathway showed excess amino acid changes between *D. melanogaster* and *D. simulans* suggesting a model wherein divergence in piRNA effector proteins between species is responsible for TE derepression in hybrids (Kelleher et *al*. 2012).

Several other studies also now provide evidence for both lineage-specific evolution of TE families (Table 18) and elevated activity of TEs in interspecific animal hybrids. For example, hybrids between *Drosophila koepferae* and *D. buzzati* experience increased *Osvaldo* transposition, a TE present in a repressed state in both parent genomes (Labrador et *al*. 1999). In mammals, hybrids between the kangaroo species *Macropus rufogriseus* and *M. agilis* exhibit unstable centromeres due to the expansion of KERV-1 TEs (Metcalfe et *al*. 2007). Hybrids between the Wallaby species *M. eugenii* and *Wallabia bicolor* also experience elevated TE activity, leading to the expansion of centromeric heterochromatin (O'Neill et *al*. 1998). In lake whitefish (*Coregonus spp*.), evidence for an increase in TE activity was revealed by sequencing the transcriptomes of hybrids between normal and dwarf species (Renaut et *al*. 2010).

TE Derepression In Plant Hybrids

There are also several examples of TE proliferation in hybrid plants that are consistent with the genomic disease model. Indeed, Mclintock's original observations of mosaic maize kernels that lead to the discovery of TEs and motivated her genomic shock hypothesis (McClintock 1950; 1984), are broadly compatible with the genomic disease model. Furthermore, the first mechanistic example of genomic defense induced by antisense TE transcription was also found in maize; silencing of the *MuDR* element involves transcription of *Mu killer,* which is the inverted duplicate of a partially deleted *MuDR* and induces heritable silencing through the small RNA pathway (Slotkin et *al*. 2003; 2005).

The best-studied and most suggestive link between TE derepression and speciation involves crosses between *Arabidopsis thaliana* and *A. arenosa,* which show postzygotic isolation mediated at least in part by proliferation of the pericentromeric *ATHILA* retrotransposon. Typically, *A. thaliana* ovules fertilized by *A. arenosa* pollen result in 95-100% seed abortion, and the reciprocal cross is impossible because *A. thaliana* pollen do not germinate on the *A. arenosa* stigma (Comai et *al*. 2000). Josefsson et *al*. (2006) demonstrated that seed abortion is significantly reduced when the *A. thaliana* female parent has higher ploidy than the *A. arenosa* pollen parent. In this cross, hybrid viability is correlated with the expression of only paternally derived *ATHILA* elements, which are more abundant in the *A. arenosa* (the paternal parent) genome than in *A. thaliana* (Josefsson et *al*. 2006). To explain this Josefsson et *al.* (2006) proposed the dosage dependent induction (DDI) model suggesting that increased relative maternal ploidy (i.e. dose) increases protection of the embryo and endosperm by balancing maternal suppressive factors with the paternal *ATHILA* copy number.

Since not all TE families are mobilized in the *A. thaliana* x *A. arenosa* cross, it remains unclear to what extent quantity and or specificity of repressors are important to hybrid inviability (Josefsson et *al*. 2006; Michalak 2009; Calarco and Martienssen 2011; Castillo and Moyle 2012). Furthermore, Martienssen and colleagues also suggest that both paternal and maternal factors may be involved (Slotkin et *al*. 2009; Martienssen 2010; Calarco and Martienssen 2011). Unlike animals, which only contribute small RNAs maternally, in *Arabidopsis* (and other angiosperms) both sexes contribute small RNAs that impact TE silencing. However, the two sexes produce distinct classes of small RNAs. Females produce 24-bp RNAs that interact with ARGONAUT9 in a manner reminiscent of the Piwi family - piRNA pathway in animals (Olmedo-Monfil et *al*. 2010). In males, pollen consists of three haploid cells - two identical sperm cells within an encompassing vegetative cell (McCormick 1993). TEs are derepressed via a regulated loss of DNA methylation in the vegetative nucleus and give rise to a distinct class of 21-bp small RNAs that are then passed to the sperm cells where they contribute to silencing their cognate elements (Slotkin et *al*. 2009). If it is necessary to have appropriate constituents from both the 21 and 24-bp small RNA pathways to silence TEs in hybrid offspring, the inviability of *A. thaliana* x *A. arenosa* hybrids may result from mismatches in suppression from either or both parental genomes.

Other examples that circumstantially tie TE derepression in plant hybrids to speciation include sunflower and rice. In sunflowers, three diploid species have arisen by ancient hybridization between *Heliantus annuus* and *H. petiolaris*. The hybrid species *H. anomalus*, *H. deserticola*, and *H. paradoxus* independently experienced large increases in retrotransposon content that are broadly consistent with the time of the species' origins (Welch and Rieseberg 2002; Schwarzbach and Rieseberg 2002; Gross et *al*. 2003; Ungerer et *al*. 2006; 2009). Interestingly, in five current natural hybrid zones between *H.*

annuus and *H. peteolaris*, TE copy number in hybrid plants does not exceed the parental species values despite active transcription of the same TE families that are expanded in the three species derived from ancient hybridization (Kawakami et *al*. 2011). This finding suggests that post-transcriptional regulation currently limits TE proliferation in *Helianthus* hybrids. The fact that the three hybrid species are each adapted to extreme environmental conditions is consistent with the 'genomic shock' and 'epi-transposon' ideas where TE proliferation in hybrids is promoted by a combination of biotic and abiotic stressors (in this case hybridization and harsh environmental conditions) upsetting the epigenetic defenses that normally keep TE activity suppressed (McClintock 1984; Wessler 1996; Lisch 2009; Zeh et *al*. 2009).

In rice, the genomic distribution and abundance of the miniature inverted-repeat transposable element (MITE) *mPing* differs within and between *indica* and *japonica* subspecies of *Oryza sativa* (Jiang et *al*. 2003)*,* reflecting multiple rounds of differential amplification (Lu et *al*. 2012). Additionally, it has been shown in laboratory crosses that a hybridization signal provided by pollination with wild rice, *Zizania litifolia* remobilizes *mPing,* its supposed autonomous partner *Pong*, and at least two other TE families (*Tos17* and *Dart*; Shan et *al*. 2005; Wang et *al*. 2010). This suggests that bursts of TE proliferation in rice may be a consequence of hybridization; however, it remains far from illustrating a direct role in speciation.

Mechanism II: Mechanical Incompatibility

Rose and Doolittle's (1983) focus on Mechanical Incompatibility dealt with two potential roles that TEs might play in hybrid infertility and inviability. First, rapid sequence turnover, particularly in heterochromatic repeats, could result in meiotic nondisjunction

due to failure of homologous chromosomes to recognize each other during meiosis. Second, differential proliferation among chromosomes could disrupt necessary spatial relationships among nonhomologous chromosomes that are based on relative chromosome arm length (see Rose and Doolittle 1983). They concluded that there was little evidence for these types of mechanical isolation despite ample evidence for sequence turnover and proliferation. In fact they noted how little sequence homology was actually necessary for chromosomes to segregate properly. Somewhat ironically, the first "speciation gene" identified in animals, *Odysseus*, is now known to be a coevolutionary result of genomic conflict driven by the need to bind with rapidly changing heterochromatic repeats (see chapter by Phadnis and Malik in this volume); and this interplay between heterochromatin associated factors is emerging as a common cause of *Drosophila* incompatibilities (e.g. Brideau et *al*. 2006, Bayes and Malik 2009, Ferree and Barbash 2009, Cattani and Presgraves 2012).

Another way that TE activity might result in mechanical incompatibility is by generating structural variation (e.g. inversions, translocations, duplications, or deletions). Isolation due to chromosomal rearrangements could arise as a direct consequence of underdominance (i.e. lower heterozygote fitness than either homozygote) in heterokaryotypic hybrids. For instance, crossovers within inversion heterozygotesmay result in gametes that do not contain a complete gene complement. If such crossovers were common, hybrids between populations or species with different inversions would be expected to show underdominance for fertility. A major difficulty with the potential for rearrangements to cause isolation directly is that fixation of an underdominant rearrangement is unlikely except in situations where genetic drift is strong (Walsh 1982; Lande 1985). If, on the other hand, there is no underdominance so that rearrangements are more likely to fix, then they are expected to have little effect on isolation.

Although direct responsibility for isolation seems unlikely, structural variants (particularly inversions) may indirectly facilitate isolation by reducing or eliminating regional recombination near breakpoints in heterokaryotypic individuals (Rieseberg 2001; Noor et *al*. 2001; Hoffmann and Rieseberg 2008; Brown and O'Neill 2010; McGaugh and Noor 2012). Recombination plays a critical role in the shaping of integrated systems within species, forming adaptive peaks in the potential field of gene combinations and holding the species as a cohesive unit (Wright 1931; 1932; Gavrilets 2004). Reduced recombination facilitates the maintenance of linkage among genes involved in adaptive divergence and reproductive isolation (Rieseberg 2001; Noor et *al*. 2001; Navarro and Barton 2003). Under this scenario, mechanical isolation is only indirectly responsible for what is more accurately a genic model of speciation. The so-called “genomic islands of divergence” that arise in regions of low recombination only contribute to speciation when they contain factors contributing to reproductive isolation (Feder and Nosil 2009). For TEs to be implicated in speciation via this mechanism requires: 1) identifying TEs as the cause of inversion, and 2) showing that loci within the inversion cause isolation.

Evidence For Mechanism II

Structural Variation Caused By TE Activity

Besides normal transposition events, which depending on the type of TE, may or may not involve duplication, TE mediated structural variants arise when: 1) homologous TEs at different genome locations recombine (ectopic recombination), 2) when TE excision results in ectopic sequences being incorporated during double strand break repair (non-homologous end joining), or 3) when the ends from two TEs synapse and engage in complete or partial simultaneous transposition (alternative transposition;

reviewed in Gray 2000; Feschotte and Pritham 2007b). There is abundant evidence that each of these are common causes of structural variation (Collins and Rubin 1984; Engels and Preston 1984; Lister et *al*. 1993; Lim and Simmons 1994; E L Walker 1995; Hua-Van et *al*. 2002; Zhang et *al*. 2009; Xing et *al*. 2009; Quinlan et *al*. 2010; Guillén and Ruiz 2012).

In some cases TEs may occur near breakpoints as a consequence of inversion rather than a cause because they tend to accumulate in regions of low recombination. However, TEs are clearly causative in other cases; perhaps the best examples coming from *Drosophila buzzatii*, where Ruiz and colleagues have mapped members of the class II TE families, *FoldBack* (*FB*) and *hAT*, to breakpoints of several inversions that occur in natural populations (Caceres et *al*. 1999; Casals et *al*. 2003; Delprat et *al*. 2009; Guillén and Ruiz 2012).

Structural Variations Associated With Reproductive Isolation

Studies in *Drosophila* spp. (Noor et *al*. 2001; Machado et *al*. 2007a), *Sorex* shrews (Basset et *al*. 2008), *Anopheles* mosquitos (Michel 2006), *Rhagoletis* flies (Feder et *al*. 2003), and *Mimulus* monkeyflower ecotypes (Lowry and Willis 2010) all demonstrate that rearranged regions diverge more quickly than collinear ones, or maintain greater divergence in the presence of gene flow between closely related races or species. However, the association between islands of divergence and speciation is not necessarily straightforward, as patterns of nucleotide differentiation are not sufficient evidence in and of themselves to infer a causal link with speciation (Noor and Bennett 2009; Turner and Hahn 2010). Additionally, regional differentiation may be maintained by selection even without chromosomal rearrangements (e.g. Turner et *al*. 2005; Harr 2006; Feder and Nosil 2009; Nadeau et *al*. 2012).

There are at least two examples where divergence in inversions has been tied to reproductive isolation. First, Noor and colleagues have mapped isolating barriers to inversion differences between *D. pseudoobscura* and *D. persimilis* (Ortiz-Barrientos et *al.* 2004; Ortiz-Barrientos and Noor 2005). These two species diverged within the last 500,000 years and are fixed differently for two inversions, but have experienced extensive introgression in other parts of the genome. Differences in cuticular hydrocarbons, mating preference, hybrid male sterility and inviability, and hybrid courtship dysfunction all map wholly or in part to the two fixed inversions (reviewed in Noor et *al*. 2001; Machado et *al*. 2007b; McGaugh and Noor 2012).

Second, Lowry and Willis (2010) discovered a geographically widespread inversion polymorphism responsible for local adaptation and prezygotic isolation in *Mimulus guttatus*. In a reciprocal transplant experiment wherein they isolated the effect of the inversion in nature by reciprocally introgressing into the genetic background of alternate ecotypes (perennial vs. annual) they showed that the inversion affects morphology, flowering time, survivorship, and prezygotic isolation.

A TE Induced Structural Variant That Causes Isolation

Clearly there is independent evidence for both TE's role in generating structural variation, and structural variation being tied to isolation. There is also at least one case where a TE induced variant is at least partially responsible for originating postzygotic isolation; the maternal effect dominant embryonic arrest (Medea) system in the red flour beetle *Tribolium castaneum*. Four Medea factors have been isolated from nature, but only 2 are stable in lab culture (M1 and M4; Beeman et *al*. 1992; Beeman and Friesen 1999). M1 is found in most regions of the world and has a selfish advantage when

invading naïve populations because offspring of heterozygous mothers who do not receive a copy of M1 all die early in development (Wade and Beeman 1994). M4 has an even broader geographic range, and produces a very similar phenotype, but M1 and M4 do not cross rescue. The causative locus for M1 has been mapped to a 21.5 kb composite Tc-1 like transposon insertion containing several defective *Tribolium* gene duplicates (Lorenzen et *al*. 2008). The insertion occurs in the ~700 bp intergenic region between the 3' UTR of two functional genes and while the mechanistic cause of the maternal phenotype remains unknown preliminary data show that the insertion may disrupt cis-natural antisense transcription that occurs in wildtype beetles (Demuth et *al*. unpublished data).

What makes Medea relevant to the present discussion of speciation is that offspring from crosses between M1 (or M4) and a second, hybrid incompatibility factor (H-factor), do not fully develop. H-factor is found widely distributed in India, and represents a strong postzygotic isolating barrier despite Medea and H-factor each producing viable, fertile offspring in combination with wildtype populations (Thomson et *al*. 1995; Thomson and Beeman 1999). Recently, H-factor was mapped to variation within the introns of an ecdysone receptor homolog, but the functional mechanisms underlying interaction between Medea and H-factor remain a mystery (Drury et *al*. 2011).

Mechanism III: Genomic Resetting

The last potential mechanism by which Rose and Doolittle suggested TEs might play a role in speciation involves their capacity to contribute regulatory sequences. Autonomous TEs encode the proteins and regulatory information necessary to catalyze their own transcription. There is now abundant evidence that proteins and regulatory

sequences derived from TEs, have been exapted to perform functions for their host (Britten 1997; Feschotte 2008). Transposases in particular have been a recurrent source of "domesticated" genes. Because they are DNA binding proteins, typically with self-specificity, as a TE disperses throughout a genome it has the potential to "rewire" regulatory circuits if the transposase becomes domesticated (Cordeaux et *al*. 2006; Feschotte 2008). In addition to abundant evidence for isolated recruitment of functional coding and non-coding sequences originating from TEs, there is growing evidence for their role in large scale rewiring of transcriptional modules (Bourque et *al*. 2008; Kunarso et *al*. 2010; Lynch et *al*. 2011; Schmidt et *al*. 2012), but no evidence that we are aware of for this mechanism in speciation.

Despite a lack of evidence for this mechanism in its original formulation, the mechanistic details of genome defense against TE activity discussed above have several facets that could easily fall under the heading "genomic resetting" though it begins to blur the line with the genomic disease model. For instance, growing evidence shows that in the germline of plants and animals, safeguarding gametes and embryonic cells involves wholesale changes in the heterochromatin of "companion" cells such that TEs are derepressed in order to provide templates for small RNA that will go on to re-establish TE silencing in cells that form the offspring (Chambeyron et *al*., 2008; Slotkin et *al*. 2009; Calarco and Martienssen, 2011). This is genome resetting, or reprogramming, in a literal sense. Indeed the newfound interplay between small RNAs, TEs, methylation and histone modification has prompted a call for a chromatin model of speciation (Michalak 2009).

Conclusion

As we stated in introducing this chapter, we have learned much in the 30 years since Rose and Doolittle's "Molecular Biological Mechanisms of Speciation". However, in many respects TEs belong to the "dark matter" of genomes and their role in speciation remain obscure. Because of their repetitive nature they are typically viewed as a nuisance to genome sequencing projects, remaining in bins of unassembled "junk" reads. Surveying them individually poses its own challenges, so getting good estimates of variety and copy number among taxa requires significant effort and purposeful investigation. Given that the decades old search for "speciation genes" has provided few candidates and even fewer where early causation can be inferred, it would be premature to draw strong positive or negative conclusions about TEs role as causative agents in speciation. Additionally, since our ability to assay epigenetic mechanisms of genome defense (e.g. small RNA, methylation, histone modification) on a genome-wide scale in non-model organisms are still nascent technological advances, perhaps it is not surprising that we lack many strong examples. As our mechanistic understanding of TEs and ability to survey them improves, we fully expect to find additional evidence for the role of TEs in speciation.

Acknowledgements

We would like to thank E. Betrán, and C. Feschotte for helpful comments and discussions about the manuscript. This work was supported by funding from National Institutes of Health grant #2R01GM065414-05A1.

Chapter 3

Dissecting The Genetic Basis Of Hybrid Incompatibility Within Species: A Neuromuscular Disorder At The Onset Of Speciation In *Tribolium castaneum*

Species are discrete groups of organisms that are reproductively isolated from other related groups of organisms. Thus, in sexually reproducing species, speciation is the process of splitting one species into two or more, and occurs through the origin of reproductive isolating barriers which permit the maintenance of genetic and phenotypic distinctiveness in geographical proximity. Reproductive isolation, including Haldane's rule, in *Tribolium castaneum* is among the earliest reported with respect to genetic divergence, making it an excellent model for speciation research. Here, we describe an example of hybrid incompatibility (called "still") segregating in F_1 hybrids between populations of *T. castaneum*, whereby affected offspring exhibit a suite of maladaptive traits upon eclosion from the pupal stage. Still hybrids exhibit a neuromuscular disorder resulting in paralysis, ataxia, and severely reduced life span. To investigate the genetic cause of the still phenotype, we sequenced the genomes of still and normal siblings using pooled-DNA and employed a genome scan approach that compares allele frequencies between extremely discordant sib pairs (still vs normal) to identify discordant alleles. Our approach uncovered two genomic regions located on chromosomes 3 and 9, which harbor the majority of candidate genes underlying the still hybrid neuromuscular disorder. In total, we identified 97 genes with significantly discordant non-synonymous SNPs between still and normal siblings. Interestingly, a set of 19 candidate loci were recently identified as candidate phosphine resistance genes (Jagadeesan et al. 2013). Phosphine is an insecticidal fumigant which acts as a metabolic toxin by targeting redox reactions, and is used worldwide in grain storage and processing facilities. The Chicago population

was collected over 7 decades aco, predating the use of phosphine, while Tanzania populations were potentially subjected to 30 years, or roughly 330 generations of routine phosphine exposure before it was collected and kept in the laboratory. Strong selection acting on mitochondrial function imposed by phosphene fumigants have potentially driven changes in mitochondrial and nuclear-encoded metabolic enzyme complexes, such as the phosphine resistance mutant, resulting in mitonuclear coadaptation that is disrupted in hybrids possessing Chicago mitochondria and Tanzania nuclei. I discuss this observation in light of the role of genetic conflict in generating hybrid incompatibilities, especially where they are still segregating early in the 'speciation continuum'.

Introduction

Identifying the genetic changes and evolutionary forces underlying the reduction in gene exchange between populations is the central problem of speciation research, and has proven to be a challenging industry. Most cases can only be studied at a single stage in the speciation process, which operates on timescales beyond several human generations. Genetic analysis is also precluded once reproductive isolation is complete and in examples where speciation is ongoing, fitness reductions in hybrids make identifying the genes underlying speciation challenging. Further complicating matters, speciation results from a composite of reproductive isolating barriers, which act in concert to reduce or eliminate gene flow between incipient species. Efforts to understand the genetics of speciation have focused primarily on developmental defects in hybrids (intrinsic postzygotic isolation) leading to sterility or inviability (Coyne and Orr 1989; 1997; Presgraves 2010; Rabosky and Matute 2013). This is perhaps due to the unexpected pattern noted by J.B.S. Haldane that whenever divergent taxa produce hybrid offspring,

the heterogametic sex (XY males in mammals and *Drosophila*) is either absent, rare, or sterile (Haldane 1922). Subsequently, much of what we know about the genetics of speciation results from our efforts to explain the evolution of hybrid male sterility or inviability in the light of 'Haldane's rule'. The predominant genetic model of postzygotic hybrid incompatibility was independently developed by Bateson, Dobzhansky, and Muller (BDM model) (Bateson 1909; Dobzhansky 1937; Muller 1940; 1942). The BDM model predicts that hybrid incompatibilities (BDMIs; although hereafter we use the more common abbreviation, DMIs) result from negative epistasis between populations: alleles present in one population cause developmental defects in the genetic background of another population, leading to hybrid sterility and inviability. However, despite decades of mapping and massive sequencing projects, central questions about these epistatic interactions remain. For example, how many genes are required for speciation? What is the minimum contribution to isolation for a gene to be considered a "speciation gene"? Are changes more likely to be regulatory, or in coding regions? What is the role of particular classes of mutations, such as transposable element insertions, gene duplications, or structural rearrangements? Are certain functional classes more likely to contribute to speciation than others?

Elegant and detailed genetic mapping experiments have led to the discovery of 4 loci each for hybrid male sterility, and inviability in *Drosophila,* along with a roughly equal number in fish, mice, yeast, plants, and copepods (Presgraves 2010; Maheshwari and Barbash 2011; Nosil and Schluter 2011). Yet in every case, DMIs require interaction with several unidentified genes in order to produce hybrid sterility and inviability phenotypes (Presgraves 2010). For example, recent evidence suggests that hybrid male sterility in *Drosophila* may predominantly evolve as a polygenic threshold trait through the combined action of many interacting genes of small-effect (Moran and Fontdevila 2014).

Once speciation is complete, species continue to diverge as a result of mutation, genetic drift, and selection, and newly derived alleles may be involved in DMIs in the laboratory. Therefore, at later stages in the speciation process, it is unclear whether the fixation of many previously identified DMIs were a cause, or a consequence of speciation (Presgraves et al. 2003; Coyne and Orr 2004; Mallet 2006). Additionally, many regions of the genome that don't contribute to reproductive isolation may diverge between allopatric populations. To understand the genetics of hybrid incompatibilities, an appeal for increased focus on DMIs acting early in species divergence and in non-model organisms has been made (Johnson 2010; The Marie Curie SPECIATION Network 2011; Maheshwari and Barbash 2011; Nosil and Schluter 2011; Seehausen et al. 2014). Hybrid incompatibilities arising at early stages of the speciation process are of particular interest, since they likely involve fewer interacting loci, and may still be segregating. Although it is often assumed that DMIs are fixed between species, both Bateson and Muller recognized that hybrid sterility and inviability may remain segregating for some time, and that this variation could be leveraged to understand the genetics of speciation (Bateson 1909; Muller 1942; Orr 1995). The role of segregating incompatibilities in speciation is poorly understood, despite documented examples of variation in reproductive compatibility in a diverse array of taxa (reviewed in: Cutter 2012). Here, we examine a novel example of hybrid incompatibility in the flour beetle, *Tribolium castaneum*, and present a high-resolution genomic study of variation between extremely discordant sibling pairs, resulting in a set of candidate incompatibility genes putatively involved in incipient speciation.

Methods

Study Organism

The red flour beetle, *T. castaneum* Herbst (Coleoptera: Tenebrionidae), is a widespread human commensal and major pest of stored grain products. While presently most abundant in structures used to process and store grain, *T. castaneum* are also associated with rotting logs and leaf-cutter bee nests, feeding on plant and animal detritus (Linsley 1942; 1944; Sokoloff 1974). *Tribolium* is a model organism for ecological, genetic, and developmental studies with a wealth of genetic tools, such as classical mutation studies, whole-genome molecular maps, transcriptome studies, and RNA interference, as well as a recently published draft genome (Wang et al. 2007; Richards et al. 2008; Kim et al. 2010). *Tribolium* is also an excellent model for speciation research. Variation for the severity of reproductive isolation segregates within and among populations of *T. castaneum* when mated with its sister species, *T. freemani* as well as between populations of *T. castaneum* (Wade and Johnson 1994; Wade et al. 1997; Demuth and Wade 2007a,b; Drury et al. 2011). Haldane's rule observed between populations of *T. castaneum* in the form of temperature dependent male deformities is the earliest with respect to genetic divergence among well studied animals (Demuth and Wade 2007b). Here, we describe an example of hybrid incompatibility (called "still") segregating in F_1 hybrids between populations of *T. castaneum*, whereby affected offspring exhibit a suite of maladaptive traits upon eclosion from the pupal stage. Still hybrids exhibit a neuromuscular disorder resulting in paralysis, ataxia, and severely reduced life span.

Populations And Crosses

The beetles used in this study were derived from two populations. One population was derived from more than 50 adults collected from Dar es Salaam, Tanzania and has been maintained in laboratory culture since 1989 (Drury and Wade 2010). The second population, c-SM (Chicago), is an outbred laboratory population created in 1977 by mass mating four inbred strains, originally collected around the late 1950's from Brazil and Chicago (Park et al. 1961; Wade 1977). Previous studies show that the productivity of Tanzania sires is reduced by roughly 20% in crosses with Chicago dams compared to conspecific females, suggesting epistatic interactions as a component of hybrid fitness (Drury and Wade 2010). Still individuals are seen in roughly three out of every four crosses and not in the reciprocal cross indicating that the alleles responsible for hybrid incompatibility are segregating in one or both populations, and may result from epistatic interactions involving uniparentally inherited genetic factors (*i.e.* cytoplasmic organelles and maternal transcripts and proteins).

To assess parent-of-origin effects and larval rearing temperature on the expression of still, a total of 48 reciprocal crosses were made by placing a virgin male and a virgin female into eight dram vials containing about 8 grams of standard medium (20:1, flour: brewers yeast, by weight). Each pair mated and laid eggs for one week under standard environmental conditions for stock maintenance (24 h dark, 29˚C, approx. 70% relative humidity) after which, the mating pair was removed and placed in a fresh vial of medium. Each mating pair was maintained for four weeks in four separate vials. To mitigate potential confounding effects of age and mating duration, vials containing eggs and medium from the first and third weeks were immediately incubated at 29˚C, and the

second and fourth vials were incubated at 35˚C. Upon eclosion from the pupal stage, we scored the offspring as 'still' or 'normal' in a blind design by placing them on top of an overturned smooth porcelain bowl. Taxis is almost completely abolished in still hybrids, so they remain on the top of the bowl, whereas normal hybrids are able to right themselves if overturned and walk off the edge of the bowl into a catchment. We counted the number of each phenotype, their sex, larval rearing temperature, and individual sire/dam origin.

Genome Sequencing Of Discordant Siblings

To investigate the genetic cause of the still phenotype, we sequenced the genomes of still and normal siblings using pooled-DNA and employed a genome scan approach that compares allele frequencies between extremely discordant sib pairs (still vs normal) to identify discordant alleles. Crosses were conducted by placing a 1-2 week old virgin Tanzania male and a similar age virgin Chicago female, in an eight-dram vial with 8 grams of standard medium. Each pair mated and laid eggs for 2 weeks under standard environmental conditions for stock maintenance after which the parents were transferred to a clean vial and offspring were allowed to mature to adulthood. Parents continued to mate and lay eggs for up to 12 weeks (6 vials). We selected two families producing roughly equal proportions of still and normal F1 hybrids (one quarter of all families) for our mapping experiment, and all offspring from these families were frozen at 20˚C. DNA was purified from whole adult individuals by sodium acetate and ethanol precipitation. For each family, DNA extractions were ranked by quality score (260/280) and pooled in equimolar amounts by phenotype for genome sequencing. In family 1, 57 normal and 57 still individuals were sequenced, while 54 normal and 54 still individuals

were sequenced in family 2. Each sample was split into two technical replicates and libraries were constructed for sequencing on the Illumina Hi-seq 2000 platform using the 100 base pair, paired-end protocol. Samples were barcoded and run 4 per lane.

Genome Sequencing Of Parent Populations

To determine lineage specificity of discordant sites, we sequenced the genomes of pooled-DNA from both the Chicago and Tanzania parental populations. Virgin adults aged roughly 7-10 days were frozen at 20˚C and DNA was purified from whole individuals by sodium acetate and ethanol precipitation. The top 100 highest quality DNA extractions were pooled in equimolar amounts by sex for genome sequencing (50 individuals per sex). DNA pools were sequenced using the same protocol as for hybrids.

Extreme Discordant Sib-Pair Analysis (Pool-GWAS)

Paired-end reads were quality trimmed (Phred 20, min 50bp) and mapped to the *T.castaneum* reference genome (version 3.0 ;(Richards et al. 2008)) using the Burrows-Wheeler aligner (BWA) *v*0.7.5 with the "mem" option and the following parameters: -Mak 12 (Li and Durbin 2010). Sequence alignment/maps were converted to binary alignment/map (BAM) for all sites with nucleotide and read quality of 20 using SAMtools *v*0.1.19 (Li et al. 2009; Li and Durbin 2010). Indels, along with flanking nucleotides on either side, were masked from alignment if present in at least one populations and covered by at least two reads. Each BAM was converted to a multi-pileup file and synchronized for allele frequency calculations. Individual per-site allele frequencies from technical and biological replicates were compared using the Cochran-Mantel-Haenszel test as implemented in PoPoolation2 *v*1.201 (Kofler et al. 2011). Sites with significant

differences between technical and or biological replicates were discarded from downstream analysis. P-value cutoffs were adjusted to take into account the normal error rate of raw reads from an Illumina Hi-Seq 1000 sequencer (Margraf et al. 2010)) and corrected for false discovery rate at $p \leq (0.05 – 0.0125)$ (Benjamini and Hochberg 1995).

We used Fisher's exact test to reveal SNPs with significant allele frequency differences between still and normal genomes. P-value cutoffs were error rate and FDR adjusted as above. We calculated the fixation index (F_{ST}) for discordant sites using an approach adapted to digital data and allowing for differences in read depth between populations as implemented in PoPoolation2 *v*1.201 (Karlsson et al. 2007). Outlier loci (discordant sites) were identified to be at least four standard deviations from the average genome-wide Fst. Discordant sites were inspected for their potential effects on annotated genes, such as coding effects (synonymous or non-synonymous amino acid replacement, start codon gains or losses, stop codon gains or losses, or frame shifts) based on position (intronic, untranslated region, upstream, downstream, splice site, or intergenic regions) as implemented in SnpEff *v*3.4 (Cingolani et al. 2012). This produced a list of candidate genes involved in the 'still' hybrid incompatibility phenotype.

Detecting The Signature Of Selection In Candidate Regions

Genes containing significantly discordant SNPs in coding regions were investigated for patterns of DNA variation potentially caused by selection. Statistical tests follow the standard neutral model, where molecuar genetic variation is assumed to be selectively neutral in a panmictic population at equilibrium between mutation and genetic drift, and held at a constant size. First, we calculated Tajima's D as implemented in PoPoolation2 *v*1.201 (Karlsson et al. 2007) for each gene (including introns), which

compares the observed, sample size corrected number of polymorphisms with the observed nucleotide heterozygosity (Tajima 1989). The test statistic, D, takes on negative values when low-frequency mutations are overrepresented – consistent with the reduction of variation following positive directional selection (or weak negative selection) preceding the accumulation of low-frequency polymorphisms. Negative values of D are consistent with an excess of intermediate frequency polymorphisms, consistent with balancing selection. Alternative, demographic explanations for values of Tajima's D exist, but may be ruled out in genome-wide studies, since values for a given locus may be compared with the overall distribution of D across the genome. For example, negative values of D are also consistent with population expansion, where low frequency alleles are preserved at many loci across the genome.

We also explored patterns of molecular evolution between the Chicago and Tanzania populations at candidate genes. We counted the number of synonymous and nonsynonymous sites, which are expected to occur proportionally to the numbers of each kind of site. The ratio on nonsynonymous substitutions per nonsynonymous site (K_A) to synonymous substitutions per synonymous site (K_S),expressed as K_A/K_S, should therefore equal 1 if substitution patterns are selectively neutral. Values exceeding 1 are indicative of positive directional selection driving the fixation of alternative amino acids. Sites exceedingly lower than 1 have likely evolved under purifying selection to maintain amino acid composition.

Detection Of Structural Variation

The potential for structural variation leading to the still hybrid incompatibility was investigated using read depth analysis. Depth measures were averaged across each chromosome in sliding windows of 10,000 base pairs , with 200 base pair steps.

Structural variants inherited from either parent population are revealed as increases (duplication) or decreases (deletion) relative to the average of sliding windows across a chromosome. Depth of coverage measurements at each site were normalized as a proportional deviation from the chromosomal maximum depth. A proportional deviation of 1 means that a given site is equal to the maximum depth of coverage for that chromosome, representing a potential duplication across individuals, whereas a value of ~ 0 means that the site was not sequenced, due to a deletion in all individuals in a library.

Results

The 'Still' Hybrid Incompatibility Phenotype

Hybridization between geographically isolated populations of *T. castaneum* is known to result in negative phenotypes, ranging from inviability, developmental arrest, deformities, and Haldane's rule (Beeman et al. 1992; Thomson et al. 1995; Demuth and Wade 2007a,b; Drury et al. 2011). Hybrids produced by Chicago dams and Tanzania sires show no difference from wildtype parents in larval locomotion. However, upon eclosion from their pupal casing, some offspring show symptoms of a neuromuscular disorder that causes them to be paralyzed or ataxic leading to behavioral sterility. Others fail to completely separate from their pupal casing and die as young adults, while some appear to be asymptomatic and have wild type locomotion. To investigate the prevalence of affected, or 'still' offspring within and among families, sex-ratio effect, and temperature effects, we reared 24 sets of reciprocal crosses at 29˚C and 35˚C. While all 24 crosses with Tanzania dams and Chicago sires resulted in some premature deaths upon eclosion, the average proportion of prematurely dead offspring per family was low (mean= 0.08 ± 0.062 std. dev., n=94) compared to crosses with Chicago dams and

Tanzania sires (mean=0.53 ± 0.371 std. dev., n=96). No still offspring were seen in these families. The proportion of families producing still or prematurely dead offspring from Chicago dams and Tanzania sires was distinct. The median proportion of still offspring across families was 0.51, while the 90% quantile produced all still offspring, and the 10% quantile produced all normal offspring. Additionally, about 23% of families produced only normal offspring and about 30% of families produced mostly still offspring while the remaining 47% of families followed a normal distribution from 14% to 89% still offspring (Shapiro-Wilk Goodness-of-fit test, $p = 0.1942$) suggesting a simple, Mendelian architecture with the possibility of several modifiers. There was no effect of rearing temperature on the prevalence of still offspring (ANOVA, $F = 0.29$, $p = 0.59$) or any difference in prevalence between male and female hybrids (ANOVA, $F = 0.17$, $p = 0.689$). Attempts to generate backcrosses using still males failed due to death or physical inability to copulate with wild-type mates, and only one still female was able to produce a few viable offspring.

Genome Scan For Discordant Alleles (Pool-GWAS)

We sequenced pooled-DNA from discordant F1 siblings, as well as from the Chicago and Tanzania populations using Illumina paired-end libraries with HiSeq, 100-cycle. After quality trimming, a total of 1,718,518,410 out of 1,723,349,990 reads (99.27%) were mapped to the reference genome at an average depth of coverage ranging from 54x to 74x per library. Since the number of individuals sequenced per phenotype per library exceeds the targeted depth of coverage of 40x, we can expect one read per individual per site on average across the genome, resulting in unbiased allele frequency estimates (Futschik and Schlötterer 2010). Genetic distances between

Chicago and Tanzania were low (0.009 substitutions per site in the nucleus and 0.0025 in the mitochondrion) and genetic distances between normal and still siblings was about half of the parental genetic distance (0.004 substitutions per site in the nucleus). Genetic differentiation between Chicago and Tanzania (Fst = 0.372) is higher than the average between-population differentiation (Fst = 0.180) observed by Drury et al. (Drury et al. 2009).

Performing a Cochran-Mantel-Haenszel (CMH) test to remove SNPs with frequency differences between technical and or biological replicates resulted in 2,051,351 polymorphic sites within the still or normal pools. To uncover discordant alleles, we performed a CMH scan between normal and still pools, resulting in 83,422 significantly differentiated SNPs (mean –log p-value = 6.06, mean F_{ST}=0.19). Using an F_{ST} cutoff of 2 standard deviations ($F_{ST} \geq 0.29$) gives us the 5% most differentiated SNPs (n=14,669 within 1kb of a gene, mean depth of coverage per SNP= 95.9x). There were 5,596 shared SNPs between still and normal pools, while 5,119 SNPs were present only in the still pool, and 3,954 SNPs were present only in the normal pool. Of these significantly discordant SNPs, 99.8% are found in outlier regions on chromosome 3 (11.9%), chromosome 9 (48.3%) and in unassigned scaffolds (39.6%, Figure 3-1). An additional 200 kb region in unassigned genomic scaffold 7 (DS497671, Uns-7) was recently mapped to chromosome 9 at position 3,529,571 between scaffolds NW_001093382.1 and NW_001093536.1, and contains 20 genes recently identified as candidate insecticide resistance genes (Jagadeesan et al. 2013), Figure 3-2. In total, there are 89 genes with significantly discordant non-synonymous SNPs between still and normal siblings. An additional 355 genes possess nucleotide changes that are either synonymous, or non-coding (i.e. occur in introns or within 1000kb upstream or downstream) Table 3-1.

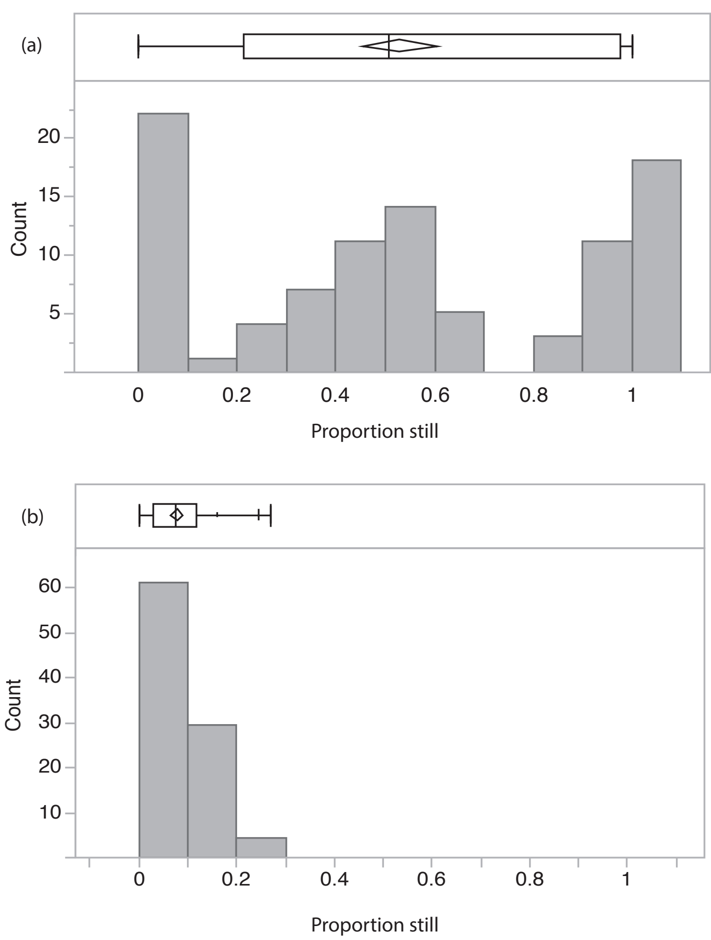


Figure 3-1 Distribution of the still phenotype in 48 reciprocal crosses with quantile plots. A total of 24 crosses were constructed with Chicago dams and Tanzania sires (a) and also with Tanzania dams and Chicago sires (b) with both replicates for each temperature pooled. Quantile plots indicate a range of 0-100% still in panel a, with a mean of 53% and median of 51% still, whereas panel (b) ranged from 0-27% with a mean of 7.8% and median of 7.6% still offspring per family.

Tests For Neutrality

The parent of origin (POO) for nonsynonymous changes in candidate genes were assigned by comparing shared alleles among the reference genome (GA2), Chicago, and Tanzania populations (e.g. the POO for an allele shared between Tanzania and Chicago is GA2; if shared between GA2 and Tanzania, then POO=Chicago; and if shared between GA2 and Chicago, then POO=Tanzania, Table 3-1). Average gene-wise FST between normal and still siblings ranged from 0.29 to 0.64. When we assayed the parental populations, several of our candidate genes show departures from neutrality, exhibiting nonzero values of the Tajima's D statistic. In Chicago, Tajima's D values for candidate genes ranged from -1.28 to 1.2 (mean = -0.28, stdev= 0.53), whereas candidate genes were more negatively skewed in Tanzania, ranging from -1.34 to 0.88 (mean = -0.26, stdev= 0.41) although the skew was nonsignificant (p=0.77, df=193). Candidate loci also showed nonneutral patterns of amino acid replacements in parental populations, with KA/KS ranging from 0.08 to 3.39 (mean= 1.02) for 25 genes with shared replacements between Chicago and Tanzania. Additionally, 28 genes had amino acid replacements in Tanzania, with KA/KS ranging from 0.09 to 2.69 (mean=0.87) and 10 genes had replacements in the Chicago population, ranging from 0.18 to 0.58 (mean=0.18, Table 3-1). Several additional genes had amino acid replacements, but no synonymous substitutions, precluding analysis of substitution rates. In summary, tests for neutrality suggest evidence for strong, directional selection for 12 genes shared between Tanzania and Chicago populations (differentiating them from the GA2 reference strain), and at loci found only in the Tanzania population for 11 genes, with no evidence for selection in the Chicago lineage.

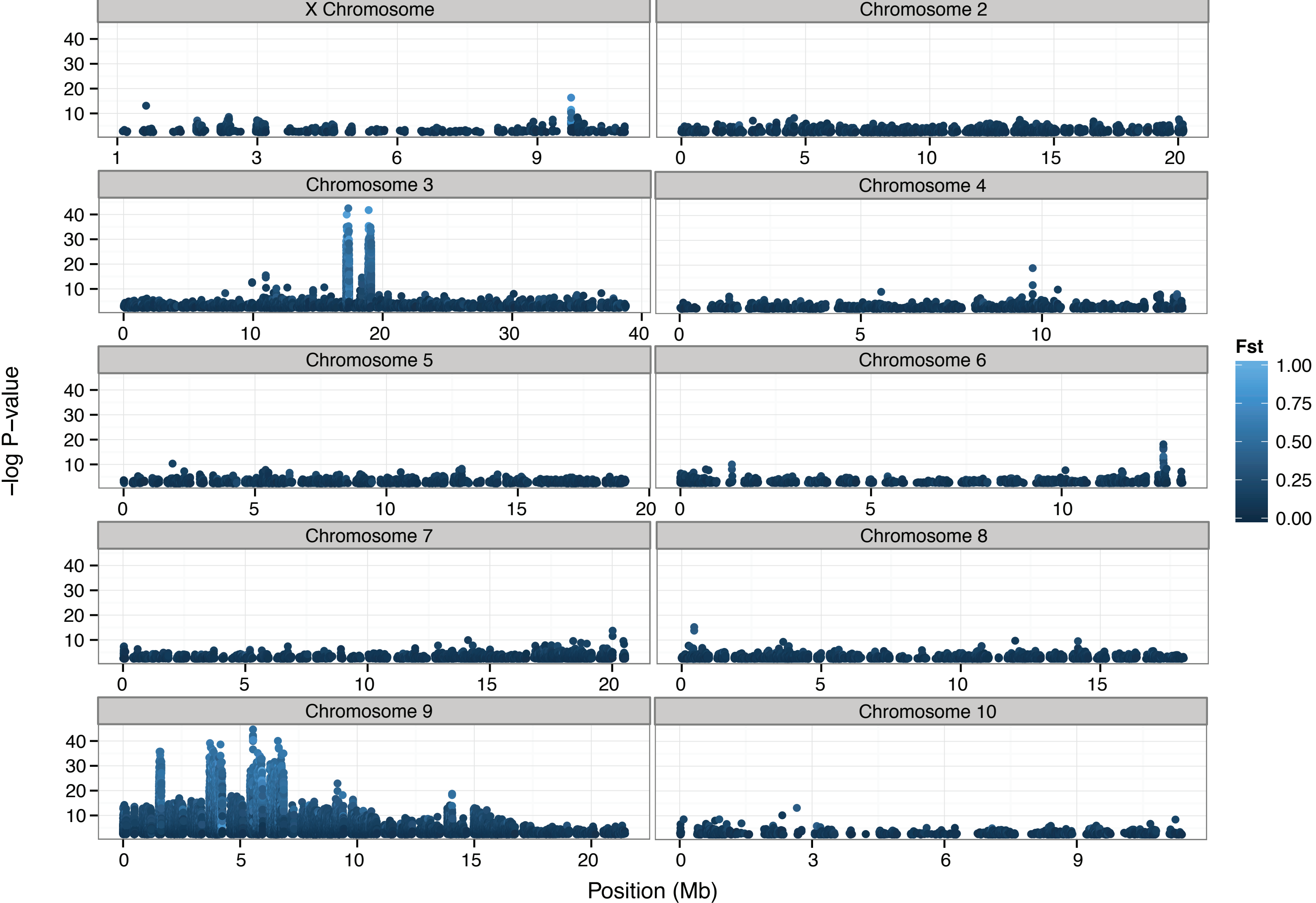

X Chromosome
Chromosome 2
Chromosome 3
Chromosome 4
Chromosome 5
Chromosome 6
Chromosome 7
Chromosome 8
Chromosome 9
Chromosome 10
−log P−value
Position (Mb)
Fst
1.00
0.75
0.50
0.25
0.00

Figure 3-2 Genome-wide scan of divergent SNPs between extremely discordant sib-pairs. 83,422 significantly differentiated SNPs mapped to the 10 assembled *T. castaneum* reference chromosomes, with 14,669 SNPs showing $F_{ST} \geq 0.29$. Candidate genes are located on the X chromosome (1 gene), chromosome 3 (11 genes), chromosome 8 (1 gene), chromosome 9, (45 genes), chromosome 10 (1 gene), and unassembled contigs (30 genes, not shown).

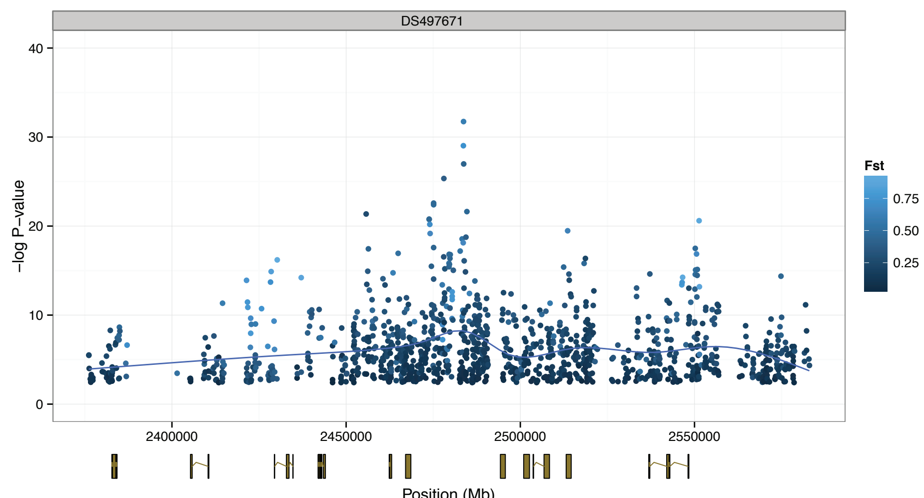


Figure 3-3 Divergent SNPs mapping to Uns-7 scaffold, containing 19 predicted genes shown below the x-axis.

Table 3-1 List of candidate genes for ‘still’ hybrid incompatibility. Genes with asterisk (*) are on the Uns-7 scaffold recently mapped to chromosome 9. (POO: parent of origin [C=Chicago, T=Tanzania])

		POO	POO			Ch	Ta	Shared			Chicago			Tanzania		
Ch	Gene	Nor.	Still	p	F_{ST}	D	D	K_a	K_s	K_a/K_s	K_a	K_s	K_a/K_s	K_a	K_s	K_a/K_s
1	TC001728	CCCC	TTTT	7.31	0.50	0.40	-0.62	0.0007			0.0003			0.0010		
3	TC002697	CC	TTTT	17.43	0.43	0.04	0.53	0.0018			0.0009			0.0009	0.0016	0.5668
3	TC002698	T	T	17.43	0.42	0.06	-0.32		0.0021					0.0011		
3	TC002699	CT	T	22.49	0.48	-0.22	-0.49		0.0093		0.0050			0.0050		
3	TC003729	T	C	23.64	0.43	-0.01	-0.18							0.0023		
3	TC003732	TT	T	16.26	0.41	-0.88	-0.99	0.0004	0.0020	0.1844				0.0008	0.0007	1.1063
3	TC003733	C	T	12.67	0.44	0.18	0.25	0.0017	0.0060	0.2802					0.0060	
3	TC010489	TTTTT	CCC	20.50	0.46	-0.56	-0.07	0.0014						0.0029	0.0026	1.1060
3	TC010495	TTTTT	CCTC	19.84	0.48	0.38	-0.14	0.0017			0.0009			0.0017		
3	TC010509	C	TT	13.45	0.45	-0.92	-0.73	0.0006			0.0006					
3	TC010557	TC	TT	17.17	0.36	0.05	0.02							0.0011		
3	TC010559	TTT	TT	13.94	0.59	-0.08	-0.17							0.0018	0.0016	1.1190

Table 3-1—*Continued*

8	TC002178	T	C	15.21	0.35	0.19	-0.71	0.0005								
9	TC001643	TTT	C	19.11	0.45	-0.50	-0.20	0.0025						0.0025		
9	TC001651	TTC	CCTT	18.00	0.51	-0.17	-0.27	0.0032							0.0012	
9	TC001654	T	TTT	20.57	0.36	-0.23	-0.55	0.0021			0.0011			0.0011		
9	TC001660	T	T	17.66	0.50	-1.06	-0.69							0.0015		
9	TC001664	T	T	12.74	0.38	-0.38	-1.01	0.0005				0.0010				
9	TC001667	CTTTT T	TTCTC	21.26	0.45	-0.80	-0.37	0.0008	0.0005	1.6135				0.0013	0.0005	2.6891
9	TC001668	TT	C	15.81	0.48	-1.27	0.00		0.0102					0.0006	0.0068	0.0923
9	TC001669	TTTTT	TTCC	16.53	0.43	0.14	0.04		0.0011					0.0019	0.0023	0.8344
9	TC001670	T	T	18.10	0.48	0.08	0.06							0.0040		
9	TC001672	TTT	TTTC	15.35	0.37	-0.41	-0.28		0.0014			0.0014		0.0008	0.0014	0.5625
9	TC001674	T	C	18.85	0.35	-0.99	0.18	0.0014	0.0184	0.0769		0.0053		0.0014	0.0105	0.1346
9	TC001677	TT	CC	15.86	0.41	-0.55	-0.24	0.0008	0.0014	0.5497	0.0008	0.0014	0.5497			
9	TC001914	C	T	16.29	0.59	-1.12	-0.61	0.0211	0.0204	1.0316						
9	TC001915	T	C	9.77	0.33	-1.05	-1.19							0.0030	0.0051	0.5804
9	TC001916	CCTCC	TTTTT	18.81	0.52	-0.22	-0.23	0.0064	0.0039	1.6445	0.0032			0.0032	0.0020	1.6445

Table 3-1—*Continued*

		C	CTTTT													
9	TC001918	TTTTTT TTTTTT	CCCCC CCC	19.95	0.55	-0.58	-0.32	0.0072	0.0021	3.3876	0.0012			0.0072		
9	TC001920	TTTTTT T	CC	19.40	0.47	-0.32	0.04	0.0007	0.0024	0.2794	0.0013			0.0027	0.0048	0.5587
9	TC001922	T	C	26.45	0.45	-0.21	-0.51		0.0016		0.0009					
9	TC001927	TTTTTT TCCCC CCT	CTTTT TCTTT TTT	12.22	0.46	-0.54	-0.15	0.0026	0.0013	1.9424	0.0011			0.0019	0.0013	1.3874
9	TC001933	TTCCC TTT	TCCCT TTTC	18.68	0.49	-0.01	0.37	0.0015	0.0017	0.8854	0.0003	0.0011	0.2656	0.0009	0.0028	0.3187
9	TC001937	TTT	T	15.54	0.40	-0.14	-0.41				0.0006	0.0024	0.2737	0.0016	0.0024	0.6844
9	TC001938	TTTC	TTTTT	19.62	0.47	-0.71	-0.31	0.0014	0.0013	1.1017	0.0014	0.0050	0.2754	0.0028	0.0038	0.7345
9	TC001939	TCTT	TTTC	16.48	0.43	-1.25	-0.54	0.0007	0.0013	0.5530	0.0028					
9	TC001941	CCTTT TT	TTTTT C	20.24	0.46	-0.75	-0.57	0.0006	0.0010	0.5465	0.0006			0.0039	0.0021	1.9129
9	TC001944	T	T	32.44	0.48	0.54	-0.08							0.0009		
9	TC001946	TTC	CT	14.14	0.41	0.31	-0.42	0.0034	0.0063	0.5383				0.0051		
9	TC004315	TCTT	C	10.20	0.43	1.20	-1.21	0.0029			0.0043					

Table 3-1—*Continued*

9	TC005213	TTCTTT	CTCTCT	19.22	0.43	-0.53	-0.32							0.0011	0.0056	0.1885
9	TC005217	TTTTTTTT	TCC	18.12	0.41	-0.83	-0.61	0.0006	0.0011	0.5723				0.0024	0.0016	1.5262
9	TC005218	TT	CC	17.90	0.44	-1.13	-0.58		0.0038					0.0043		
9	TC005219	TTCTTT	CCTCCTTC	15.34	0.43	0.03	-0.45	0.0026						0.0026	0.0012	2.2107
9	TC005222	T	T	21.67	0.50	-0.30	-0.47				0.0002			0.0006		
9	TC010964	T	C	11.15	0.31	-0.97	-0.06									
9	TC010969	T	C	12.56	0.34	-1.09	-0.63								0.0062	
9	TC010973	C	T	10.95	0.33	-0.22	-0.52									
9	TC012081	TT	CC	15.59	0.38	-1.13	0.22	0.0019						0.0006		
9	TC012125	T	C	8.42	0.30	-0.27	0.26	0.0004								
9	TC012188	TT	TT	7.84	0.54	-0.80	-0.28	0.0030	0.0028	1.0627						
9	TC012205	T	C	2.97	0.38	-0.44	-0.04	0.0006							0.0010	
9	TC012221	C	T	13.41	0.29	-0.09	-0.08		0.0003			0.0003				
9	TC012222	C	T	11.04	0.30	-0.35	-0.21									
9	TC012268	T	T	28.40	0.45	-1.01	0.08	0.0011						0.0011		

Table 3-1—*Continued*

9	TC012284	C	T	7.77	0.29	0.43	-0.33		0.0004							
9	TC012313	C	T	11.96	0.32	-0.36	0.03									
9	TC012344	TT	TTT	7.68	0.30	0.45	0.23	0.0007								
10	TC011220	T	C	13.19	0.38	-1.24	-1.27				0.0003					
U	TC001985	T	TTT	19.91	0.38	-0.41	-0.54							0.0021		
U	TC002162	CC	TT	18.57	0.58	-0.36	-0.28							0.0025	0.0023	1.0788
U	TC002165	TTCTT TTC	TCCTT T	17.24	0.41	-0.64	-0.19	0.0012	0.0020	0.5731	0.0004			0.0023	0.0020	1.1462
U	TC002166	TT	TTT	12.29	0.38	-0.81	-0.46	0.0020						0.0013		
U	TC002258	T	TTTC	16.72	0.39	-0.74	-0.43	0.0020	0.0012	1.5970	0.0007	0.0037	0.1774	0.0007	0.0012	0.5323
U	TC002260	CCT	TTT	16.46	0.46	0.10	0.04	0.0038	0.0093	0.4110				0.0013	0.0070	0.1827
U	TC002262	TT	TC	21.07	0.44	-0.61	0.88							0.0031	0.0109	0.2798
U	TC002263	TTTTTT	TTTTC	21.42	0.51	0.28	0.02	0.0016			0.0065			0.0016		
U	TC002265	TT	C	22.05	0.49	0.43	-0.23		0.0010		0.0006			0.0017		
U	TC002308	CT	TT	19.12	0.45	-0.44	0.50	0.0015	0.0013	1.1248				0.0007	0.0066	0.1125
U	TC002309	TCC	TT	15.69	0.50	0.54	-0.21	0.0024	0.0020	1.1773				0.0012	0.0020	0.5887
U	TC002318	T	C	12.78	0.33	0.67	0.24		0.0062					0.0017		

Table 3-1—*Continued*

U	TC002694	T	T	21.33	0.51	-0.26	-0.26				0.0017					
U	TC003735	C	T	7.98	0.34	-0.30	-0.86							0.0019		
U	TC004170	CCCCC	TTTTTT T	13.94	0.49	0.12	-0.52	0.0026	0.0011	2.2890	0.0006	0.0011	0.5723			
U	TC004173	C	CT	7.40	0.48	0.28	-0.09		0.0049		0.0028	0.0049	0.5807			
U	TC004256	TC	TTCTC T	16.58	0.42	-0.07	-0.35	0.0055			0.0014					
U	TC006821*	CC	TT	6.77	0.31	-0.38	-0.38	0.0008								
U	TC006825*	TT	TTC	13.87	0.40	-0.22	-0.41	0.0008			0.0008	0.0043	0.1959	0.0008		
U	TC008534	TT	CC	8.33	0.30	0.93	0.03	0.0039	0.0034	1.1417						
U	TC008535	T	C	23.20	0.44	0.17	-0.23							0.0007	0.0012	0.5542
U	TC008548	T	C	19.72	0.44	-0.07	0.48							0.0026		
U	TC010622	TTT	CC	18.27	0.46	-0.17	-0.17	0.0015			0.0005					
U	TC011596	C	T	23.22	0.46	0.55	-0.16	0.0012								
U	TC006827*			7.31	0.50	-0.39	-1.17									
U	TC006828*			17.43	0.43	0.30	0.02									
U	TC006829*			17.43	0.42	-1.16	-1.34									
U	TC006835*			22.49	0.48	-0.22	-0.41									

Table 3-1—*Continued*

U	TC006836*			23.64	0.43	-1.28	-0.66									
U	TC006838*			11.16	0.31	-0.46	-0.76									

Structural Variation

The potential for structural variation leading to the still hybrid incompatibility was investigated using read depth analysis. Structural variants inherited from either parent population are revealed as increases (duplication) or decreases (deletion) relative to the average of sliding windows across a chromosome. Figure 3-2 shows the difference in the proportion of mapped reads between normal and still siblings. The read depths for normal and still siblings were remarkably similar across the whole genome. The expectation for a causative structural variant is that it would be shared by most of the individuals in either the still or normal pool, but absent in the other pool, leading to a minimum expected difference of 50% in mapped read depth between the pools. Because read depths were so similar between pools, several peaks across the genome were greater than 4 standard deviations from the average of sliding windows (i.e. we had great power to detect non-zero differenc in read depth); however, the maximum deviation between pools was at a locus on chromosome 10 where normal siblings had only 7.4% greater depth than still (normal = 24.4% and still = 16.9% greater than the chromosome maximum). Such a low differentiation between pools seems very unlikely to be causative.

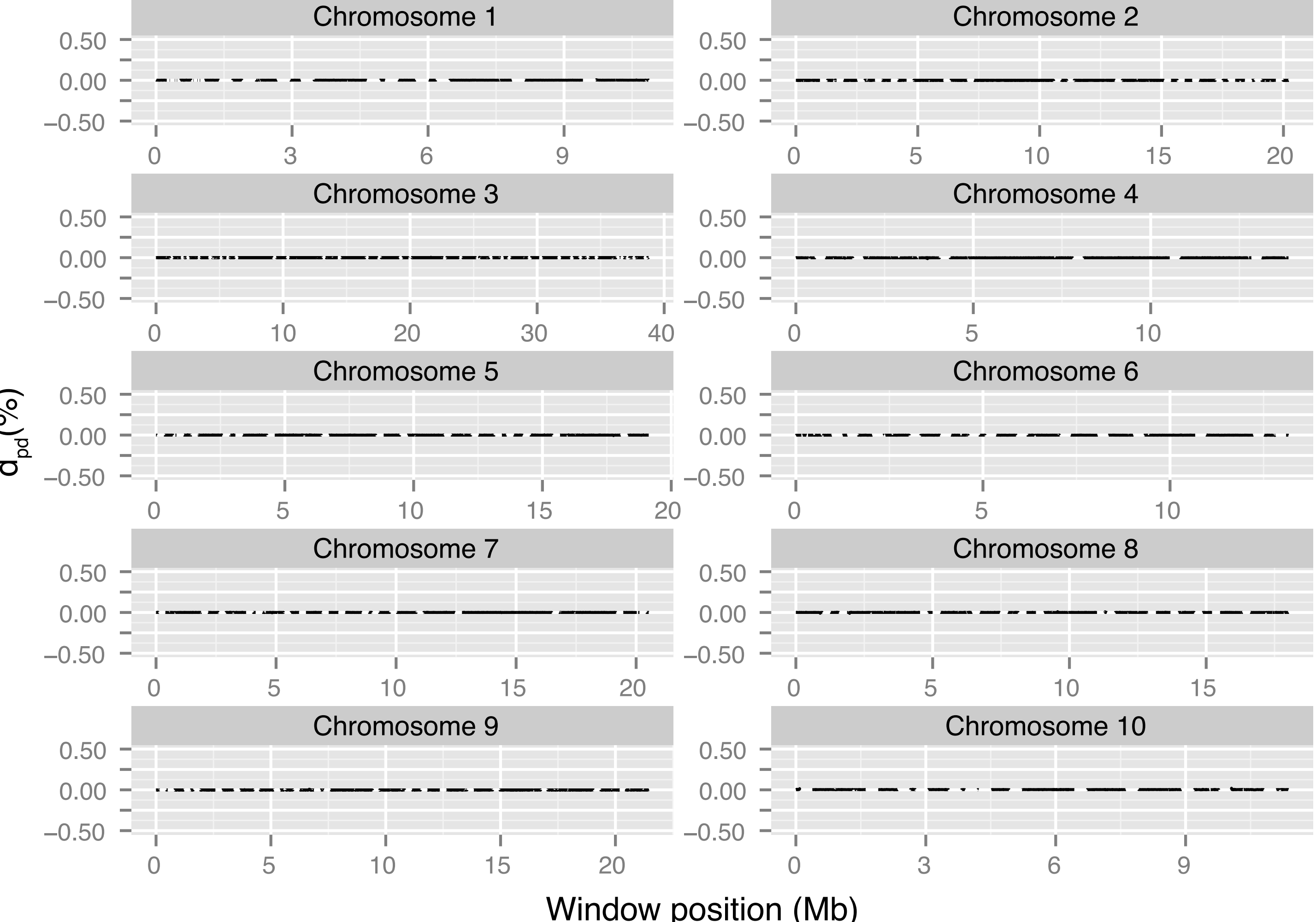

Chromosome 1
Chromosome 2
Chromosome 3
Chromosome 4
Chromosome 5
Chromosome 6
Chromosome 7
Chromosome 8
Chromosome 9
Chromosome 10
0.50
0.00
−0.50
d_{pd}(%)
Window position (Mb)

Figure 3-4 Structural variation across the genomes of hybrid siblings. Percent difference in proportional deviation (d_{pd}) was found by subtracting the measurement from still siblings from the normal siblings giving a positive value where fewer reads mapped in still, and negative values where fewer reads mapped to normal (see methods for proportional deviation calculation).

Model type	Allelic substitutions	Divergence in allopatry	Hybridization in sympatry	Direction of DMI
Classic model				
Two-locus two-allele	Two substitutions in the same lineage			Ancestral–derived
Two-locus two-allele	One substitution in each lineage			Derived–derived
Co-evolutionary model				
Two-locus two-allele	Two substitutions in each lineage			Derived–derived
Four-locus two-allele	Two substitutions in each lineage			Derived–derived; ancestral–derived

Figure 3-5 Models of hybrid incompatibility in a genomic conflict scenario. Circles represent derived alleles. In co-evolutionary models, DMIs are continually fixed either at the same loci (that is, two-locus, two-allele) or at different loci (that is, four-locus, two-allele) (see the table). In all examples with two substitutions in a lineage, the selfish locus (left) drives the evolution of the restorer locus (right). Grey arrows indicate negative epistatic interactions between complementary loci. In all models, the ancestral state is wild type except for the two-locus two-allele co-evolutionary model. In this model, the ancestral state is a co-evolving selfish element-restorer system; numbers represent the lineages in which the derived alleles originated. (from (Seehausen et al. 2014)

Discussion

Identifying the genetic changes underlying hybrid incompatibility phenotypes is a central problem of speciation research and a formidable task. One reason for this difficulty is that newly derived alleles may be involved in DMIs in taxa where speciation is already complete. Polymorphic DMIs are likely over the course of speciation, but are of particular interest when operating at early stages in the ‘speciation continuum’, where their relevance to speciation is direct and causative. In the present study, we have used the polymorphic nature of hybrid incompatibility in *Tribolium castaneum* to our advantage, by comparing the genomes of siblings with extremely discordant phenotypes. This effectively increases the percentage of the genome that is shared between the sequenced individuals without the need to construct introgression lines. Our results also demonstrate that it is possible to rapidly identify hybrid incompatibility loci with F_1 individuals, where maladaptive hybrids are unable to mate.

Our approach uncovered two genomic regions located on chromosomes 3 and 9, which harbor the majority of candidate genes underlying the still hybrid neuromuscular disorder (Figure 3-1). The results of our SNP analysis showed these regions to contain highly discordant allele frequencies between full siblings distinguished only by their expression of the still phenotype.

A Potential Role For Pesticide Resistance In Tribolium Speciation

A set of 19 candidate loci that we mapped to a 200 kb unassigned genomic scaffold (DS497671, Uns-7) were recently linked to chromosome 9 between scaffolds

NW_001093382.1 and NW_001093536.1, and identified as candidate phosphine (PH_3) resistance genes(Jagadeesan et al. 2013). Phosphine is an insecticidal fumigant which acts as a metabolic toxin by targeting redox reactions, and is used worldwide in grain storage and processing facilities. It has been common and recurrent selective force for *T.castaneum* for over 50 years, leading to the evolution of highly heritable resistance alleles (Chaudhry 1999; Bell 2000; Jagadeesan et al. 2012; 2013). Resistance in phosphine gas has evolved repeatedly in insects (*T. castaneum* and *Rhyzopertha dominica*) and in nematodes (*Caenorhabditis elegans*) and is due to polymorphisms at the dihydrolipoamide dehydrogenase (*dld*) gene, an essential metabolic enzyme involved in four multienzyme complexes in the mitochondrion: pyruvate dehydrogenase a-ketagutarate dehydrogenase, branced-chain ketoacid dehydrogenase, and glycine cleavage (Schlipalius et al. 2012). Despite the convergence of phosphine resistance at the *dld* locus across taxa, an additional locus in *T.castaneum* on chromosome 8 (*rph1)* containing 17 genes was found to interact epistatically with Uns-7, resulting in a 100-fold increase in phosphine resistance, and showing a selective fitness advantage in the absence of phosphine selection over 18 generations. In contrast, Uns-7 showed a fitness cost associated with resistance alleles in the absence of selection, with a reduction of homozygous resistance alleles and an increase in heterozygotes and homozygous sensitive alleles (Jagadeesan et al. 2013). Thus Uns-7 would seem to be under balancing selection in nature and perhaps we should see an increase in heterozygosity leading to positive Tajima's D across the region. Instead, the average Tajima's D for this region in Tanzania is -0.37. One explanation for this is that balancing selection at this locus is due to fluctuating selection, where it experiences positive selection in the presence of phosphine and purifying selection in its absence. Under long term balancing selection, allele frequencies should change at a slower rate than they would under neutral genetic

drift, yet under fluctuating selection, allele frequency change should be rapid in comparison. This should lead to negative Tajima's D values, since both purifying and positive directional selection have the same effect of reducing variation near the selected locus via hitchhiking. An alternative explanation is that recurrent bottlenecks have shaped the allele frequency spectrum, leading to an abundance of low frequency alleles and overall negative values of Tajima's D. Little is known about the maintenance of the Tanzania stock, however the Chicago stock has historically been maintained in the laboratory as large populations (n > 200), without a rigorous program of mating sibs for many generations to create inbred lines. In contrary, the Chicago stock has been historically maintained with efforts to increase or maintain it's high level of heterozygosity (Sokoloff 1974; Drury and Wade 2010).

In *C. elegans,* phosphine treatment results in aberrant mitochondrial morphology, a 70% reduction in oxidative respiration, and a severe drop in mitochondrial membrane potential ($\Delta\Psi_m$) within several hours (Zuryn et al. 2008). Additionally, phosphine-resistant mutants had more mitochondrial genome copies than wild-type animals. Strong selection acting on mitochondrial function imposed by phosphene fumigants have potentially driven changes in mitochondrial and nuclear-encoded metabolic enzyme complexes, such as the *dld* phosphine resistance mutant, resulting in mitonuclear coadaptation and thus hybrid incompatibility in crosses between Chicago mitochondrial genes and Tanzania nuclear genes. Although it is difficult to determine the potential role of each candidate, many are involved in oxidative stress pathways (Appendix A). One such gene, TC002698 on chromosome 3, is predicted to be a Thioredoxin reductase (*Trx2*), a redox-active enzyme that helps protect cells from damage due to reactive oxygen species in the mitochondria (Conrad et al. 2004; Meyer et al. 2009). Additionally, many mitochondrial genes were nearly fixed for allelic

differences between Chicago and Tanzania, with 21 alleles in 9 genes all fixed for alternate alleles, where Chicago populations all had the reference (GA-2) allele (Appendix D).

Because the Chicago population has been in laboratory culture for nearly seven decades, it was likely collected well before the widespread use of phosphine fumigants, whereas the Tanzania population was potentially subjected to 30 years, or roughly 330 generations of routine phosphine exposure before it was collected and kept in the laboratory. Our genomic survey of the *T. castaneum* populations revealed a 350% increase in genetic polymorphisms in Tanzania relative to Chicago at mitochondrial loci (0.0053 vs 0.0015 polymorphisms per site) and a 120% increase at nuclear loci (0.023 vs 0.019 polymorphisms per site), suggesting a potential for cytonuclear coadaptation driven by increased a mitochondrial mutation rate. However, it remains a curiosity that divergent mitochondrial genes from Tanzania remain functional in the reciprocal cross. Perhaps the presence of the two outlier families shown in figure 3-1 with close to 27% dead offspring hint at a small degree of negative mitonuclear epistasis in the reciprocal cross.

Other Candidate Genes

Our mapping results uncovered several additional candidate genes that potential contribute to the still incompatibility in *T.castaneum.* Most of the additional candidate genes were identified on genomic scaffolds that are not present in the assembled reference genome. The unassembled scaffolds in *T. castaneum* include approximately 22 megabases (~10% of the estimated genome size). These scaffolds remains unassembled largely due to their very high repetitive element content (simple repeats and transposons (Wang et al. 2008).

Analysis of predicted biological function and identity revealed three major types of gene in our remaining candidate list: transposable element – derived genes (n=21), membrane-associated genes (n=20), and genes involved in neural pathways (n=9, Appendix C). Of the 89 identified candidates segregating between Chicago and Tanzania populations for amino acid changes, 21 contained hits to transposable elements in Repbase with alignment scores greater than s=80. Of these, 13 were Gypsy elements; 6 were LINEs; and Polinton, Poseidon, and Rehavkus elements were each represented twice. The gene with the best hit (s=6020), TC003732 was previously identified in *T.castaneum* as Rehavkus-3, a DNA transposon encoding a 1200 amino acid transposase, a C-terminal Ulp1 cysteine protease and Cys_4-His-Cys_3 plant homeodomain (PDH) finger involved in chromatin-mediated gene regulation in *Drosophila* and *C. elegans* (Kapitonov et al. 2006; Iwase et al. 2007).

The Potential General Role Of Genetic Conflict In Driving Hybrid Incompatibility Genes

Many authors have recently pointed to the role of selection in driving the fixation of DMIs in *Drosophila* and *Xiphophorus*, and a role for genetic conflict in speciation has emerged in *Drosophila* and mouse(Presgraves 2010; Johnson 2010; Crespi and Nosil 2012). Genetic conflict arises in individuals that inherit DNA sequences with antagonistic fitness interests, and can contribute to speciation by generating genetic divergence between populations in sets of antagonistically-interacting loci that interact epistatically as DMIs in hybrid genomes (McClintock 1980; COSMIDES and TOOBY 1981; HURST and POMIANKOWSKI 1991)(Thompson & Beeman 1995??, NRG-paper). In the classic model, DMIs are envisioned as two-locus, two-allele interactions, in which incompatibilities arise either between an ancestral allele and an allele that is derived in

one lineage or between alleles that are derived in two separate lineages (figure 3-3). A special case of the model with separately derived alleles can refer to maternal-effect 'selfish' loci in which maternal 'poison' and zygotic 'antidote' are both due to divergence in developmental expression of the same locus. In a co-evolutionary framework, DMIs are continually fixed either at the same loci (that is, two-locus, two-allele) or at different loci (that is, four-locus, two-allele, figure 3-3). Insights into the role of genomic conflict in speciation reveal the potential for further development of models of hybrid incompatibility. In evolutionary models, fitness landscapes define the relationship between genotype and fitness. Individual genotypes in the adaptive landscape have to evolve to keep up with fluctuations in the adaptive landscape due to changing environments, but at a given time will lag behind potential optimal genotypes. The lag-load defines the degree in which contemporaneous genotypes fall below the potential local maximum. Models that incorporate the possibility for increased lag-load due to ongoing co-evolution predict successively more severe incompatibilities as the lag load increases. Additional theoretical work is needed to investigate such co-evolutionary models.

Many forms of genetic conflict have been putatively linked with speciation in diverse taxa where DMIs are yet to be identified (e.g. chromosomal conflicts, drive, imprinting, transposable elements, and cytonuclear conflicts; reviewed in (Crespi and Nosil 2012; Seehausen et al. 2014). A clear role in speciation for epistasis between cytoplasmic (usually mitochondrial) and nuclear genomes has emerged in *Saccharomyces* yeast, *Mimlus* monkeyflowers, and *Tigriopus* copepods (Willett and Burton 2001; Lee et al. 2008; Barr and Fishman 2010). Additional evidence for cytonuclear barriers has also been shown in cases where DMIs have not yet been physically mapped in *Passer* sparrows and in many species of Angiosperms(Tiffin et al. 2001; Turelli and Moyle 2007; Trier et al. 2014).Essentially all mitochondrial processes

(replication, transcription, translation) and functions (oxidative phosphorylation, OXPHOS) require interaction with nuclear-encoded genes. These include protein interactions, protein-DNA, and protein-RNA interactions (Gaspari et al. 2004; Ballard and Melvin 2010; Meiklejohn et al. 2013). The potential for negative epistasis exists for all of these interactions in crosses between individuals from parent populations that have diverged in cytoplasmic factors, since nuclear-encoded DMIs in one population interact with cytoplasmic factors of the other population. (Grun 1976; Burton and Barreto 2012).

Genetic conflicts also provide an explanation for the evidence of many well known 'speciation genes' that are still segregating between species (e.g. *Hybrid male rescue (Hmr), Lethal hybrid rescue (LHR), Odysseus (OdsH), Zygotic hybrid rescue (Zhr),* all of which show signatures of positive selection (WATANABE 1979; Hutter and Ashburner 1987; Sawamura et al. 1993; Ting et al. 1998; Orr and Irving 2005). In addition, segregating genetic conflicts appear to contribute to reproductive isolation between populations in *Drosophila (Overdrive, Ovd), Mus* (Meisetz, *Prdm9),* and *Tigriopus* (Cytochrome c, *cytc*) (Lee et al. 2008; Phadnis and Orr 2009; Mihola et al. 2009). These examples suggest that genomic conflicts retain their potential to contribute to reproductive isolation throughout all phases of the speciation process.

Future Prospects For Speciation Research

Although many well known 'speciation genes' remain segregating between species, little is known about the role of segregating incompatibilities at the onset of speciation. Genetic mapping experiments, such as this one, in combination with quantitative and molecular genetic analyses of hybrid crosses will reveal the prevalence

and architecture of segregating incompatibilities. In addition, meta-analysis of molecularly identified loci will help distiguish any differences in evolutionary forces behind fixed and segregating incompatibility loci. For example, how often does genetic conflict lead to hybrid incompatibility when alleles are segregating versus when they are fixed, and how many loci does each scenario involve? The study of genetic incompatibilities early in the speciation continuum is essential to understanding these and other questions that have remained elusive to evolutionary biologists for many decades.

Appendix A

List Of Published Cases Of Marsupial Hybrids

Appendix A 1 Published cases of marsupial hybrids.

Species Pair		Hybrid inviability								Hybrid sterility		Reference
A	B	AxB	Fem.	Mal.	p	BxA	Fem.	Mal.	p	AxB	BxA	
M.dorsalis	*M.eugenii*	N	0	1	0.5	–	–	–	–	M	–	Smith *et al.* 1979
M.eugenii	*M.parma*	F	–	–		–	–	–	–	M	–	Close and Lowry 1990
M.giganteus	*M.fuliginosus*	N	4	3	0.27	–	–	–	–	M	–	Poole and Catling 1974
M.parryi	*M.dorsalis*	F	–	–		–	–	–	–	–	–	Close and Lowry 1990
M.rob.erubescens	*M.rob.robustus*	N	1	2	0.38	N	–	–	–	M	–	Poole and Merchant 1987, Johnson et al. 1978
M.rob.robustus	*W.bicolor*	F	–	–		–	–	–	–	M	–	Close and Lowry 1990
M.rufo.banksianus	*M.rufo.rufogriseus*	F				F	–	–	–	N	N	Johnson *et al.* 1978, Merchant and Calaby 1981
M.rufogriseus	*M.agilis*	N	0	1	0.5	F	–	–	–	B	M	Smith et al. 1979, Johnson 1985, Lowry 1988, Newsome et al 1977
M.rufogriseus	*E.eugenii*	F				–				–	–	Johnson *et al.* 1978
M.rufogriseus	*M.giganteus*	F				–				M	–	Johnson *et al.* 1978
M.rufogriseus	*W.bicolor*	N	0	1	0.5	–	–	–	–	M	–	Smith *et al.* 1979, Johnson 1985
M.rufus	*M.agilis*	N				–				M	–	Johnson 1985
P.assimilis	*P.inornata mareeba*	M	4	0	0.061	F	0	8	0.004	M	M/B	Sharman et al. 1990
P.assimilis	*P.sharmani*	N	0	3	0.125	–	–	–	–	M	–	Close and Lowry 1990
P.assimilis	*P.penicillata*	N	9	1	0.010		–	–	–	M		Sharman et al. 1990
P.assimilis	*P.inornata*	N	0	1	0.5	N	0	1	0.5	M/?	–	Sharman et al. 1990
P.assimilis	*P.inornata mtclaro*	–				–				M/?	–	Sharman et al. 1990

P.godmani	*P.purpureicollis*	N				−				M	−	Sharman et al. 1990
P.lateralis MacD	*P.assimilis*	N	1	1	0.5	−	−	−	−	M		Close and Bell 1997
P.lateralis MacD	*P. lateralis pearsoni*	N	4	2	0.234	N	1	4	0.156	F?		Close and Bell 1997
P.purpuretcollis	*P.lateralis MacD*	N	2	0	0.25	−	−	−	−			Close and Bell 1997
P.purpuretcollis	*P.penicillata*	N	2	0	0.25	−	−	−	−			Close and Bell 1997
P.herberti	*P.penicillata*		2	2	0.375	−	−	−	−		M?	Close and Bell 1997, Sharman et al 1990
P.inornata mareeba	*P.herberti*		0	2	0.25	−	−	−	−			Close and Bell 1997
P.inornata mareeba	*P.penicillata*	N	10	7	0.150	−	−	−	−			Close and Bell 1997
P.inornata mareeba	*P.godmani*	−				N	−	−	−	M	M	Briscoe et al. 1982, Sharman *et al.* 1990
P.inornata mareeba	*P.sharmani*	N	3	2	0.313	−	−	−	−	M	−	Close and Lowry 1990
P.inornata mareeba	*P. inornata mtclaro*	−				−				?	−	Sharman et al. 1990
P.persephone	*P.xanthopus*	N	2	0	0.25	−	−	−	−	N	−	Briscoe et al. 1982
T.thetis	*T.stigmatica*	N	13	16	0.126	B	0	0		M	−	Calaby and Poole 1971
M.agilis	*M.eugenii*	B	−	−	−	−	−	−	−	−	−	Close and Lowry 1990
M.rufus	*M.giganteus*	N	1	1	0.5	N	−	−	−	B	B	Poole and Catling 1974
M.rufus	*M.fuliginosus*	N				N				B	B	Poole and Catling 1974
M.rufus	*M.robustus*	N	2	1	0.375	N	1	1	0.5	B	B	Poole and Catling 1974, Smith et al. 1979
B.giamardi	*B.pencillata*	N				−				−	−	Gray 1971
Dendrolagus_inustus	*D.ursinus*	−				−				−	−	Gray 1971
M.agilis	*W.bicolor*	−				−				−	−	Johnson 1985
M.dorsalis	*M.rufogriseus*	−				−				−	−	Johnson 1985
M.dorsalis	*M.parma*	−	−	−	−	−	−	−	−	N	N	Close and Lowry 1990

Table A-1—*Continued*

M.giganteus	*M.parryi*	–	0	1	0.5	–				–	–	Calaby and Poole 1971, Johnson 1985
M.giganteus	*A.agilis*	–				–				–	–	Johnson 1985
M.rob.robustus	*M.antilopinus*	–	–	–	–	–	–	–	–	–	–	Close and Lowry 1990
M.rufogriseus	*M.parryi*	–	–	–	–	–	–	–	–	–	–	Close and Lowry 1990
P.concinna	*P.brachyotis*	–	–	–	–	–	–	–	–	–	–	Close and Lowry 1990
P.lateralis_hacketti	*P.rothschildi*	–				–				–	–	Hayman and Martin 1969
Setonix_brachyurus	*M.eugenii*	–	–	–	–	–	–	–	–	–	–	Close and Lowry 1990
T.thetis	*M.rufogriseus*	–				–				–	–	Gray 1971

Appendix B

Potential Function Of Candidate Genes

Appendix B 2 Potential function of candidate genes. Amino acid sequences of candidates were searched against multiple annotated reference sequences as implemented in Blast2GO (Conesa et al. 2005), using an Expect score cutoff of 0.001.

Chr	Gene	sequence description
1	TC001728	gag-pol protein
3	TC002697	protein melted-like
3	TC002698	thioredoxin reductase
3	TC002699	hypothetical protein
3	TC003729	cg13532 cg13532-pa
3	TC003732	short-chain dehydrogenase
3	TC003733	solute carrier family 25 member 35-like
3	TC010489	laccase-like multicopper oxidase 1
3	TC010495	map kinase-interacting serine threonine-protein kinase 1-like
3	TC010509	tyrosine partial
3	TC010557	transport protein sec61 subunit alpha 2
3	TC010559	superkiller viralicidic activity 2-like
8	TC002178	bel12_ag transposon polyprotein
9	TC001643	ER resident protein 44-like isoform
9	TC001651	gag-like protein
9	TC001654	hypothetical protein
9	TC001660	p-element transposase
9	TC001664	RNA-directed DNA polymerase from jockey-like ME
9	TC001667	integrin alpha-ps2
9	TC001668	sorting nexin-19
9	TC001669	hypothetical protein
9	TC001670	hypothetical protein
9	TC001672	pancreatic triacylglycerol lipase-like
9	TC001674	sh3 domain-binding protein 5 homolog isoform x1
9	TC001677	tripartite motif-containing protein 2-like isoform x3
9	TC001914	hypothetical protein
9	TC001915	transposable element tc3 transposase
9	TC001916	cytochrome p450 306a1-like
9	TC001918	zinc metalloproteinase yil108w-like
9	TC001920	phospholipase a-2-activating
9	TC001922	lethal malignant brain tumor-like protein 3-like
9	TC001927	hypothetical protein
9	TC001933	hypothetical protein
9	TC001937	polycystic kidney disease 2-like 1 isoform X2
9	TC001938	low quality protein: polycystin-1
9	TC001939	rna-directed dna polymerase from mobile element jockey-like
9	TC001941	low affinity cationic amino acid transporter 2-like
9	TC001944	protein-l-isoaspartate o-methyltransferase
9	TC001946	serine protease h31
9	TC004315	hypothetical protein
9	TC005213	hypothetical protein

Table B-1—*Continued*

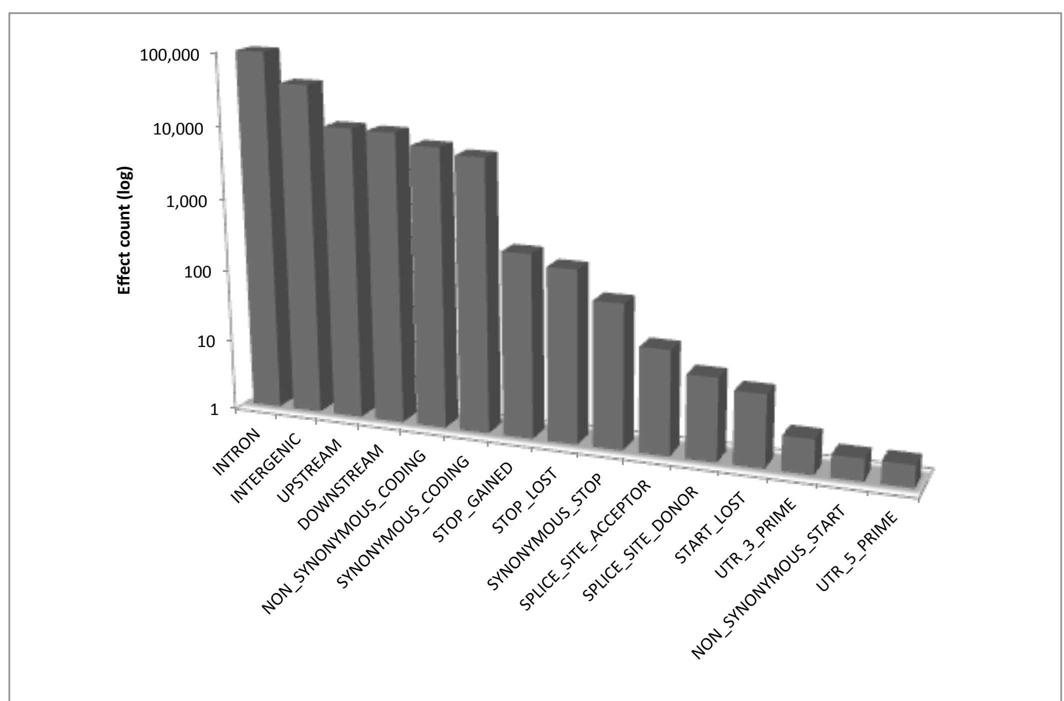
9	TC005217	rough deal
9	TC005218	hypothetical protein
9	TC005219	dopamine transporter
9	TC005222	ankyrin repeat
9	TC010964	retinol dehydrogenase 11
9	TC010969	hypothetical protein
9	TC010973	bel12_ag transposon polyprotein
9	TC012081	type ii inositol- -trisphosphate 5-phosphatase
9	TC012125	4-aminobutyrate aminotransferase
9	TC012188	anterior fat body protein
9	TC012205	mical-like protein 2
9	TC012221	huntingtin
9	TC012222	protein kinase protein
9	TC012268	slit protein
9	TC012284	glucosyl glucuronosyl transferases
9	TC012313	serine threonine-protein phosphatase
9	TC012344	protein sidekick-like
10	TC011220	pol protein
U	TC001985	piggybac transposable element-derived protein 3-like
U	TC002162	sec14-like protein 2-like
U	TC002165	ankyrin unc44
U	TC002166	ankyrin unc44
U	TC002258	hypothetical protein
U	TC002260	sodium channel protein nach-like
U	TC002262	cg7675 cg7675-pb
U	TC002263	isoform c
U	TC002265	tbc1 domain family member 30
U	TC002308	bride of sevenless
U	TC002309	synaptic vesicular amine transporter-like
U	TC002318	hypothetical protein
U	TC002694	calcitonin receptor
U	TC003735	reverse partial
U	TC004170	reverse transcriptase
U	TC004173	hypothetical protein
U	TC004256	hypothetical protein
U	TC006843	zinc finger protein 271-like
U	TC006964	endonuclease and reverse transcriptase-like protein
U	TC008534	creb 7 protein
U	TC008535	hypothetical protein
U	TC008548	hypothetical protein
U	TC010622	endonuclease and reverse transcriptase-like protein
U	TC011596	metabotropic glutamate receptor 2
U	TC006827*	serine/threonine protein kinase
U	TC006828*	fatty acetyl coA synthetase activity
U	TC006829*	cytochrome p450; Cypg15
U	TC006835*	dopamine transporter; ankyrin
U	TC006836*	hypothetical protein

Table B-1—*Continued*

U	*	nucleic acid/zinc ion binding

Appendix C

Distribution Of Discordant Snps Across Genomic Feature



Appendix C 1 Distribution of discordant SNPs across genomic features. All 14,669 discordant SNPs were binned according to their location in genomic features within 1 kb upstream and downstream of candidate genes as well as intergenic regions between genes, given as the effect count per feature class.

Appendix D

Mitochondrial Divergence Between Chicago And Tanzania Populations

Appendix D 1 Distribution of coding differences between Chicago and Tanzania mitochondria

Gene	Position	Ref	Chicago	Tanzania	frequency diff.
COX1	1484	G	G	A	0.983
COX1	1997	C	C	T	0.991
COX1	2126	A	A	G	0.992
COX1	2633	A	A	G	0.979
COX2	3293	C	C	T	0.988
ATP6	4288	T	T	C	0.988
COX3	5112	G	G	A	0.984
COX3	5346	G	G	A	0.995
ND5	6515	C	C	T	0.975
ND5	6705	T	T	A	0.979
ND4	8439	T	T	C	0.979
ND4	8651	C	C	T	0.982
ND4	8763	A	A	T	0.989
ND6	9962	A	A	G	0.973
CYTB	11206	A	A	G	0.992
ND1	12277	T	T	C	0.952
ND1	12442	C	C	T	0.982
REP_ORI	14732	T	T	A	0.996
REP_ORI	14777	A	A	G	0.991
REP_ORI	14844	G	G	A	0.984
REP_ORI	15155	C	C	T	0.973

References

Abe, H., F. Ohbayashi, T. Shimada, T. Sugasaki, S. Kawai, K. Mita and T. Oshiki. 2000. Molecular structure of a novel gypsy -Ty3-like retrotrans- poson (Kabuki) and nested retrotransposable elements on the W chromosome of the silkworm *Bombyx mori*. *Mol Gen Genet* 263:916–924.

Al Nadaf S, Waters PD, Koina E, Deakin JE, Jordan KS, Graves JAM. 2010. Activity map of the tammar X chromosome shows that marsupial X inactivation is incomplete and escape is stochastic. Genome Biol. 11(12):1–18.

Aravin, A. A., Hannon G. J. and Brennecke J. 2007. The Piwi-piRNA Pathway provides an adaptive defense in the transposon arms race. *Science* 318: 761–764.

Bachtrog, D. 2005. Sex chromosome evolution: Molecular aspects of Y-chromosome degeneration in *Drosophila*. *Genome Research* 15:1393–1401.

Ballard, J. W. O., and R. G. Melvin. 2010. Linking the mitochondrial genotype to the organismal phenotype. Mol Ecol 19:1523–1539.

Barr, C. M., and L. Fishman. 2010. The nuclear component of a cytonuclear hybrid incompatibility in Mimulus maps to a cluster of pentatricopeptide repeat genes. Genetics 184:455–465.

Basset, P., Yannic G. and Hausser J. 2008. Chromosomal rearrangements and genetic structure at different evolutionary levels of the *Sorex araneus* group. *J Evol Biol* 21:842–852.

Bateson, W. 1909. Heredity and variation in modern lights. Pp. 85–101 *in* A. C. Seward, ed. Darwin and Modern Science. Cambridge University Press, Cambridge.

Bayes, J. J., and Malik H. S., 2009. Altered heterochromatin binding by a hybrid sterility protein in *Drosophila* Sibling Species. *Science* 326:1538–1541.

Beeman, R. W. and Friesen K. S. 1999. Properties and natural occurrence of maternal-effect selfish genes ("Medea" factors) in the red flour beetle, *Tribolium castaneum*. *Heredity* 82:529–534.

Beeman, R. W., Friesen K. S. and Denell R. E. 1992. Maternal-effect selfish genes in flour beetles. *Science* 256:89–92.

Belancio, V. P., Hedges D. J. and Deininger P. 2008. Mammalian non-LTR retrotransposons: For better or worse, in sickness and in health. *Genome Research* 18:343–358.

Bell, C. H. 2000. Fumigation in the 21st century. Crop Protection 19:563–569.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B

Biemont, C. and Vieira C. 2005. What transposable elements tell us about genome organization and evolution: the case of *Drosophila*. *Cytogenet Genome Res* 110:25–34.

Biémont, C. and Vieira C. 2006. Genetics: Junk DNA as an evolutionary force. *Nature* 443:521–524.

Bingham, P. M., Kidwell M. G. and Rubin G. M. 1982. The molecular basis of *P-M* hybrid dysgenesis: the role of the *P* element, a *P*-strain-specific transposon family. *Cell* 29:995–1004.

Blumenstiel, J. P. and Hartl D. L. 2005. Evidence for maternally transmitted small interfering RNA in the repression of transposition in *Drosophila virilis*. *Proc Natl Acad Sci USA* 102:15965–15970.

Bourc'his, D. and Voinnet O. 2010. A small-RNA perspective on gametogenesis, fertilization, and early zygotic development. *Science* 330:617–622.

Bourque, G., Leong B., Vega V. B., Chen X., Lee Y. L., Srinivasan K. G., Chew J. L., Ruan Y., Wei C.-L., Ng H. H., and Liu E. T. 2008. Evolution of the mammalian transcription

factor binding repertoire via transposable elements. *Genome Research* 18:1752–1762.

Brandt, J., Schrauth S., Veith A. M., Froschauer A., Haneke T., Schultheis C., Gessler M., Leimeister C. and Volff J. N. 2005. Transposable elements as a source of genetic innovation: expression and evolution of a family of retrotransposon-derived neogenes in mammals. *Gene* 345:101–111.

Bregliano, J., Picard G., Bucheton A., Pelisson A., Lavige J. and L'Heritier P. 1980. Hybrid dysgenesis in *Drosophila melanogaster*. *Science* 207:606–611.

Brennecke, J., Malone C. D., Aravin A. A., Sachidanandam R., Stark A. and Hannon G. J. 2008. An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* 322:1387–1392.

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

Britten, R. J. 1997. Mobile elements inserted in the distant past have taken on important functions. *Gene* 205:177–182.

Britten, R. J. and Davidson E. H. 1971. Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. *Q Rev Biol* 46:111–138.

Brookfield, J. F. Y. and Badge R. M. 1997. Population genetics models of transposable elements. *Genetica* 100:281–294.

Brothers AN, Delph LF. 2010. Haldane’s rule is extended to plants with sex chromosomes. Evolution. 64:3643–3648.

Brown, J. D. and O'Neill R. J. 2010. Chromosomes, conflict, and epigenetics: chromosomal speciation revisited. *Annu. Rev. Genom. Human Genet*. 11: 291–316.

Bucheton, A., Lavige J. M., Picard G.and L'Heritier P. 1976. Non-Mendelian female sterility in *Drosophila melanogaster*: Quantitative variations in the efficiency of inducer and

reactive strains. *Heredity* 36:305–314.

Bucheton, A., Paro R., Sang H. M., Pelisson A. and Finnegan D. J. 1984. The molecular basis of *I-R* hybrid Dysgenesis in *Drosophila melanogaster*: Identification, cloning, and properties of the I factor. *Cell* 38:153–163.

Burgoyne PS, Mahadevaiah SK, Sutcliffe MJ, Palmer SJ. 1992. Fertility in mice requires X–Y pairing and a Y-chromosomal ‘‘spermiogenesis’’ gene mapping to the long arm. Cell. 71(3):391–398.

Burton, R. S., and F. S. Barreto. 2012. A disproportionate role for mtDNA in Dobzhansky–Muller incompatibilities? Mol Ecol n/a–n/a.

Butter, F., Kappei D., Buchholz F., Vermeulen M. and Mann M. 2010. A domesticated transposon mediates the effects of a single-nucleotide polymorphism responsible for en- hanced muscle growth. *Nat. Rev. Genet*. 11:305–311.

Cáceres, M., Ranz, J. M., Barbadilla, A., Long, M., & Ruiz, A. 1999. Generation of a widespread *Drosophila* inversion by a transposable element. *Science*, 285:415-418.

Calarco, J. P. and Martienssen R. A. 2011. Genome reprogramming and small interfering RNA in the *Arabidopsis* germline. *Curr. Opin. Genet. Dev*. 21: 134–139.

Callinan, P. A. and Batzer M. A. 2006. Retrotransposable elements and human disease. *Genome dynamics* 1:104–115.

Casals, F., M. Cáceres and A. Ruiz. 2003. The foldback-like transposon *Galileo* is involved in the generation of two different natural chromosomal inversions of *Drosophila buzzatii*. Mol Biol Evol 20:674–685.

Castillo, D. M. and Moyle L. C. 2012. Evolutionary Implications of Mechanistic Models of TE-Mediated Hybrid Incompatibility. *Evolution* 2012:1–12.

Cattani, M. V., and 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.

Chambeyron, S., Popkova A., Payen-Groschene G., Brun C., Laouini D., Pelisson A., and Bucheton A. 2008. piRNA-mediated nuclear accumula- tion of retrotransposon trans- cripts in the *Drosophila* female germline. *Proc Natl Acad Sci USA* 105:14964–14969.

Charlesworth, B. and Charlesworth D. 1983. The population dynamics of transposable elements. *Genet Res* 42:1–27.

Charlesworth, B. and Langley C. H. 1989. The population genetics of *Drosophila* transposable elements. *Annu Rev Genet* 23:251–287.

Charlesworth, B., Sniegowski P. and Stephan W. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nat Rev Genet* 371:215–220.

Charlesworth, D., Charlesworth B. and Marais G. 2005. Steps in the evolution of heteromorphic sex chromosomes. *Nat Rev Genet* 95:118–128.

Chaudhry, M. Q. 1999. Review A Review of the Mechanisms Involved in the Action of Phosphine as an Insecticide and Phosphine Resistance in Stored-Product Insects. Pestic. Sci. 49:213–228.

Chénais, B., Caruso A., Hiard S. and Casse N. 2012. The impact of transpo- sable elements on eukaryotic genomes: From genome size increase to genetic adaptation to stressful environments. *Cell* 509:7–15.

Cingolani, P., A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, and D. M. Ruden. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly 6:80–92.

Clark NL, Swanson WJ. 2005. Pervasive adaptive evolution in primate seminal proteins. PLoS Genet. 1(3):e35.

Close R, Lowry P. 1990. Hybrids in marsupial research. Aust J Zool. 37(2–4): 259–267.

Collins, M. and Rubin G. M. 1984. Structure of chromosomal rearrangements induced by the FB transposable element in *Drosophila*. *Nature* 308:323–327.

Comai, L., Tyagi A. P., Winter K., Holmes-Davis R., Reynolds S. H., Stevens Y. and Byers B. 2000. Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. *Plant Cell* 12:1551–1568.

Conesa, A., S. Gotz, J. M. Garcia-Gomez, J. Terol, M. Talon, and M. Robles. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676.

Conrad, M., C. Jakupoglu, S. G. Moreno, S. Lippl, A. Banjac, M. Schneider, H. Beck, A. K. Hatzopoulos, U. Just, F. Sinowatz, W. Schmahl, K. R. Chien, W. Wurst, G. W. Bornkamm, and M. Brielmeier. 2004. Essential Role for Mitochondrial Thioredoxin Reductase in Hematopoiesis, Heart Development, and Heart Function. Molecular and Cellular Biology 24:9414–9423.

Cooper D, Vandeberg J, Sharman G, Poole W. 1971. Phosphoglycerate kinase polymorphism in kangaroos provides further evidence for paternal X-inactivation. Nature. 230(13):155–157.

Cordaux, R., Udit S., Batzer M. A., and Feschotte C. 2006. Birth of a chimeric primate gene by capture of the transposase gene from a mobile element. *Proc Natl Acad Sci USA* 103:8101–8106.

COSMIDES, L. M., and J. TOOBY. 1981. Cytoplasmic Inheritance and Intragenomic Conflict. J Theor Biol 89:83–129.

Coyne JA, Orr HA. 1989a. Patterns of speciation in Drosophila. Evolution. 43(2):362–381.

Coyne JA, Orr HA. 1989b. Two rules of speciation. In: Otte D, Endler JA, editors. Speciation and its consequences. Sunderland (MA): Sinauere Associates, Inc.

Coyne JA, Orr HA. 2004. Speciation. Sunderland (MA): Sinauer Associates, Inc.

Coyne JA. 1985. The genetic basis of Haldane's rule. Nature. 314(6013):736–738.

Coyne JA. 1992. Genetics and speciation. Nature. 355(6360):511–515.

Coyne, J. A. 1985. Genetic studies of three sibling species of *Drosophila* with relationship to theories of speciation. *Genet Res* 46:169–192.

Coyne, J. A. 1986. Meiotic segregation and male recombination in interspeci- fic hybrids of *Drosophila*. *Genetics* 114:485–494.

Coyne, J. A. 1989. Mutation rates in hybrids between sibling species of *Droso- phila*. *Heredity* 63 (Pt 2):155–162.

Coyne, J. A., and H. A. Orr. 1997. "Patterns of speciation in Drosophila"" revisited. Evolution 51:295–303.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Inc., Sunderland, MA.

Crespi, B., and P. Nosil. 2012. Conflictual speciation: species formation via genomic conflict. Trends Ecol Evol, doi: doi: 10.1016/j.tree.2012.08.015.

Cutter, A. D. 2012. The polymorphic prelude to Bateson–Dobzhansky–Muller incompatibilities. Trends Ecol Evol 27:209–218.

de Boer, J. G., Yazawa R., Davidson W. S. and Koop B. F. 2007. Bursts and horizontal evolution of DNA transposons in the speciation of pseudotetra- ploid salmonids. *BMC Genomics* 8:422.

de Koning, A. P. J., Gu, W., Castoe, T. A., Batzer, M. A., & Pollock, D. D. 2011. Repetitive Elements May Comprise Over Two-Thirds of the Human Genome. PLoS Genetics7: e1002384.

Deakin J, Chaumeil J, Hore T, Marshall J. 2009. Unravelling the evolutionary origins of X chromosome inactivation in mammals: insights from marsupials and monotremes. Chromosome Res. 17:671–685.

Delprat, A., Negre B., Puig M. and Ruiz A. 2009. The transposon galileo generates natural

chromosomal inversions in *Drosophila* by ectopic recombination. *PLoS ONE* 4:e7883.

Demuth, J. P. and Wade M. J. 2005. On the theoretical and empirical framework for studying genetic interactions within and among species. *Am. Nat.* 165:524–536.

Demuth, J. P., and M. J. Wade. 2007a. Population differentiation in the beetle *Tribolium castaneum*. I. Genetic architecture. Evolution 61:494–509.

Demuth, J. P., and M. J. Wade. 2007b. Population differentiation in the beetle *Tribolium castaneum*. II. Haldane's rule and incipient speciation. Evolution 61:694–699.

Dobigny, G., Ozouf-Costaz C., Waters P. D., Bonillo C., Coutanceau J.-P. and Volobouev V. 2004. LINE-1 amplification accompanies explosive geno- me repatterning in rodents. *Chromosome Res* 12:787–793.

Dobzhansky, T. 1937. *Genetics and the Origin of the Species.* Columbia University Press, New York, NY.

Drury, D. W., A. L. Siniard, and M. J. Wade. 2009. Genetic differentiation among wild populations of Tribolium castaneum estimated using microsatellite markers. J Hered 100:732–741.

Drury, D. W., and M. J. Wade. 2010. Genetic variation and co-variation for fitness between intra-population and inter-population backgrounds in the red flour beetle, Tribolium castaneum. J Evol Biol 24:168–176.

Drury, D. W., Jideonwo V. N., Ehmke R. C. and Wade M. J. 2011. An unusual barrier to gene flow: perpetually immature larvae from inter-population crosses in the flour beetle, *Tribolium castaneum*. *J Evol Biol* 24:2678–2686.

Eanes, W. F., Wesley C., Hey J., Houle D. and Ajioka J. W. 1988. The Fitness Consequences of *P*-Element Insertion in *Drosophila melanogaster*. *Genet Res* 52:17–26.

Eickbush, T. H. and Furano A. V. 2002. Fruit flies and humans respond differently to retrotransposons. *Curr. Opin. Genet. Dev.* 12:669–674.

Eldridge MD, Close RL. 1993. Radiation of chromosome shuffles. Curr Opin Genet Dev. 3(6):915–922.

Engels, W. R. and C. R. Preston. 1984. Formation of chromosome rearrange- ments by *P* factors in *Drosophila*. *Genetics* 107:657–678.

Engels, W. R. and Preston C. R. 1979. Hybrid dysgenesis in *Drosophila me- lanogaster*: the biology of female and male sterility. *Genetics* 92:161–174.

Feder, J. L. and Nosil P. 2009. Chromosomal inversions and species differences: when are genes affecting adaptive divergence and reproduce- tive isolation expected to reside within inversions? *Evolution* 63: 3061–3075.

Feder, J. L., Roethele J. B., Filchak K., Niedbalski J. and Romero-Severson J. 2003. Evidence for inversion polymorphism related to sympatric host race formation in the apple maggot fly, *Rhagoletispomonella*. *Genetics* 163: 939–953.

Fedoroff, N. V. 1999. Transposable Elements As a Molecular Evolutionary Force. *Annals NY Acad Sci* 870:251–264.

Fedoroff, N. V. 2012. Transposable Elements, Epigenetics, and Genome Evolution. *Science.* 338:758-767.

Ferree, P. M. and Barbash D. A. 2009. Species-specific heterochromatin prevents mitotic chromosome segregation to cause hybrid lethality in *Drosophila*. *PLoS Biology* 7:e1000234.

Feschotte, C. 2008. Transposable elements and the evolution of regulatory net- works. *Nat. Rev. Genet.* 9:397–405.

Feschotte, C. and Pritham E. J. 2007a. *Computational Analysis and Paleogenomics of Interspersed Repeats in Eukaryotes.* Pp. 31–54 *in* N. Stojanovic, ed. Computational Genomics: Current Methods. Horizon Bioscience, Norfolk, UK.

Feschotte, C. and Pritham E. J. 2007b. DNA transposons and the evolution of eukaryotic

genomes. *Annu Rev Genet* 41:331–368.

Finnegan, D. J. 1992. *Transposable elements.* Pp. 1096–1107 *in* D. L. Lindsley and G. G. Zimm, eds. The genome of *Drosophila melanogaster*. San Diego Academic Press, San Diego, CA.

Frank SA. 1991. Divergence of meiotic drive-suppression systems as an explanation for sex-biased hybrid sterility and inviability. Evolution. 45:262–267.

Futschik, A., and C. Schlötterer. 2010. The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. Genetics 186:207–218.

Gaspari, M., M. Falkenberg, N.-G. Larsson, and C. M. Gustafsson. 2004. The mitochondrial RNA polymerase contributes critically to promoter specificity in mammalian cells. EMBO J 23:4606–4614.

Gavrilets, S. 2004. *Fitness landscapes and the origin of species.* Princeton Univ Pr.

Ghildiyal, M., Seitz H., Horwich M. D., Li C., Du T., Lee S., Xu J., W. Kittler E. L., Zapp M. L., Weng Z. and Zamore P. D. 2008. Endogenous siRNAs Derived from Transposons and mRNAs in *Drosophila* Somatic Cells. *Science* 320:1077–1081.

Gifford R, Kabat P, Martin J, Lynch C, Tristem M. 2005. Evolution and distribution of class II-related endogenous retroviruses. J Virol. 79(10):6478–6486.

Ginzburg, L. R., Bingham P. M. and Yoo S. 1984. On the theory of speciation induced by transposable elements. *Genetics* 107:331–341.

Girard, A. and Hannon G. J. 2008. Conserved themes in small-RNA-mediated transposon control. *Trends Cell Biol.* 18:136–148.

Good JM, Nachman MW. 2005. Rates of protein evolution are positively correlated with developmental timing of expression during mouse spermatogenesis. Mol Biol Evol. 22(4):1044–1052.

Goodier, J. L. and Kazazian H. H. Jr. 2008. Retrotransposons Revisited: The Restraint and

Rehabilitation of Parasites. *Cell* 135:23–35.

Graves J, ONeill R. 1997. Sex chromosome evolution and Haldane's rule. J Hered. 88(5):358–360.

Graves J. 1996. Mammals that break the rules: genetics of marsupials and monotremes. Annu Rev Genet. 30:233–260.

Gray AP. 1954. Mammal hybrids. Farnham Royal (UK): Commonwealth Agricultural Bureaux.

Gray AP. 1958. Bird hybrids. Farnham Royal (UK): Commonwealth Agricultural Bureaux.

Gray, Y. H. M. 2000. It takes two transposons to tango:transposable-element-mediated chromosomal rearrangements. Cell 16:461–468.

Gross, B. L., Schwarzbach A. E. and Rieseberg L. H. 2003. Origin (s) of the diploid hybrid species *Helianthus deserticola* (Asteraceae). *Am. J. Bot.* 90: 1708–1719.

Grun, P. 1976. Cytoplasmic Genetics and Evolution. Columbia Univ. Press, New York.

Guillén, Y. and Ruiz A. 2012. Gene alterations at *Drosophila* inversion breakpoints provide *prima facie* evidence for natural selection as an explanation for rapid chromosomal evolution. *BMC Genomics* 13:53.

Gupta V, Parisi M, Sturgill D, Nuttall R, Doctolero M, Dudko OK, Malley JD, Eastman PS, Oliver B. 2006. Global analysis of X-chromosome dosage compensation. J Biol. 5(1):1–22.

Haas, B. J., Wortman J. R., Ronning C. M., Hannick L. I., Smith R. K. J., Maiti R., Chan A. P., Yu C., Farzad M., Wu D., White O. and Town C. D. 2005. Complete reannotation of the *Arabidopsis* genome: methods, tools, protocols and the final release. *BMC Biol* 3:7.

Hahn MW, Lanzaro GC. 2005. Female-biased gene expression in the malaria mosquito Anopheles gambiae. Curr Biol. 15(6):R192–R193.

Haldane JBS. 1932. The causes of evolution. London: Longmans, Green & Co.

Haldane, J. 1922. Sex ratio and unisexual sterility in hybrid animals. *J Genet* 12:101–109.

Harr, B. 2006. Genomic islands of differentiation between house mouse subspecies. *Genome*

Res 16:730–737.

Henikoff S, Ahmad K, Malik HS. 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. Science. 293(5532):1098–1102.

Hey, J. 1988. Speciation via hybrid dysgenesis: negative evidence from the *Drosophila affinis* subgroup. *Genetica* 78:97–103.

Hillis DM, Green DM. 1990. Evolutionary changes of heterogametic sex in the phylogenetic history of amphibians. J Evol Biol. 3(1–2):49–64.

Hiraizumi, Y., Slatko B., Langley C. and Nill A. 1973. Recombination in *Drosophila melanogaster* male. *Genetics* 73:439–444.

Hoffmann, A. A. and Rieseberg L. H. 2008. Revisiting the Impact of Inversions in Evolution: From Population Genetic Markers to Drivers of Adaptive Shifts and Speciation? *Annu. Rev. Ecol. Evol. Syst.* 39:21–42.

Hua-Van, A., Langin T. and Daboussi M. J. 2002. Aberrant transposition of a Tc1-mariner element, impala, in the fungus *Fusarium oxysporum*. *Mol Gen and Genom* 267:79–87.

Hurst LD, Pomiankowski A. 1991. Causes of sex ratio bias may account for unisexual sterility in hybrids: a new explanation of Haldane's rule and related phenomena. Genetics. 128:841–858.

Hutter, P., and M. Ashburner. 1987. Genetic rescue of inviable hybrids between Drosophila melanogaster and its sibling species. Nature 327:331–333.

International Chicken Genome Sequencing Consortium. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432:695–716.

International Human Genome Sequencing Consortium.2001. Initial sequenc- ing and analysis of the human genome. *Nature* 409:860–921.

Itoh Y, Replogle K, Kim Y-H, Wade J, Clayton DF, Arnold AP. 2010. Sex bias and dosage

compensation in the zebra finch versus chicken genomes: general and specialized patterns among birds. Genome Res. 20(4):512–518.

Iwase, S., F. Lan, P. Bayliss, L. de la Torre-Ubieta, M. Huarte, H. H. Qi, J. R. Whetstine, A. Bonni, T. M. Roberts, and Y. Shi. 2007. The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. Cell 128:1077–1088.

Jagadeesan, R., A. Fotheringham, P. R. Ebert, and D. I. Schlipalius. 2013. Rapid genome wide mapping of phosphine resistance loci by a simple regional averaging analysis in the red flour beetle, Tribolium castaneum. BMC Genomics 14:650.

Jagadeesan, R., P. J. Collins, G. J. Daglish, P. R. Ebert, and D. I. Schlipalius. 2012. Phosphine Resistance in the Rust Red Flour Beetle, Tribolium castaneum (Coleoptera: Tenebrionidae): Inheritance, Gene Interactions and Fitness Costs. PLoS ONE 7.

Jensen, S., Gassama M.-P. and Heidmann T. 1999. Taming of transposable elements by homology-dependent gene silencing. *Nat. Rev. Genet.* 21: 209–212.

Jiang, N., Bao Z., Zhang X., Hirochika H., Eddy S. R., McCouch S. R. and Wessler S. R. 2003. An active DNA transposon family in rice. *Nat. Rev. Genet.* 421:163–167.

Johnson NA, Perez DE, Cabot EL, Hollocher H, Wu CI. 1992. A test of reciprocal X–Y interactions as a cause of hybrid sterility in Drosophila. Nature. 358(6389):751–753.

Johnson, N. A. 2010. Hybrid incompatibility genes: remnants of a genomic battlefield? Trends Genet 26:317–325.

Josefsson, C., Dilkes B. and Comai L. 2006. Parent-dependent loss of gene silencing during interspecies hybridization. *Curr. Biol.* 16:1322–1328.

Jurka, J., Bao W. and Kojima K. K. 2011. Families of transposable elements, population structure and the origin of species. *Biol. Direct* 6:44.

Kapitonov, V. V., A. J. Gentles, and J. Jurka. 2006. Rehavkus-3_TC, a family of Rehavkus DNA

transposons from the red flour beetle genome. Repbase Reports 6:196–196.

Kapitonov, V. V., and Jurka J. 2003. Molecular paleontology of transposable elements in the Drosophila melanogaster genome. *Proc Natl Acad Sci USA*100: 6569–6574.

Kapitonov, V. V., and Jurka J. 2005. RAG1 core and V (D) J recombination signal sequences were derived from Transib transposons. *PLoS Biology* 3: e181.

Karlsson, E. K., I. Baranowska, C. M. Wade, N. H. C. Salmon Hillbertz, M. C. Zody, N. Anderson, T. M. Biagi, N. Patterson, G. R. Pielberg, E. J. Kulbokas, K. E. Comstock, E. T. Keller, J. P. Mesirov, H. von Euler, O. Kämpe, Å. Hedhammar, E. S. Lander, G. Andersson, L. Andersson, and K. Lindblad-Toh. 2007. Efficient mapping of mendelian traits in dogs through genome-wide association. Nature genetics 39:1321–1328.

Kawakami, T., Dhakal P., Katterhenry A. N., Heatherington C. A. and Ungerer M. C. 2011. Transposable Element Proliferation and Genome Expansion Are Rare in Contemporary Sunflower Hybrid Populations Despite Widespread Transcriptional Activity of LTR Retrotransposons. *Genome Biol Evol* 3:156–167.

Kazazian, H. H. 1998. Mobile elements and disease. *Curr. Opin. Genet. Dev.* 8:343–350.

Kazazian, H. H. 2004. Mobile Elements: Drivers of Genome Evolution. *Science* 303:1626–1632.

Kelleher ES, Edelman NB, Barbash DA (2012) *Drosophila* Interspecific Hybrids Phenocopy piRNA-Pathway Mutants. PLoS Biol 10(11): e1001428.

Khaitovich P, Hellmann I, Enard W, Nowick K, Leinweber M, Franz H, WeissG,LachmannM,Pa¨a¨boS.2005.Parallelpatternsofevolutionin the genomes and transcriptomes of humans and chimpanzees. Science. 309(5742):1850–1854.

Kidwell, M. G. 1983. Evolution of hybrid dysgenesis determinants in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 80:1655–1659.

Kidwell, M. G. and Kidwell J. F. 1975. Cytoplasm–chromosome interactions in *Drosophila*

melanogaster. *Nature* 253:755–756.

Kidwell, M. G. and Lisch D. R. 2001. Perspective: Transposable elements, parasitic DNA, and genome evolution. *Evolution* 55:1–24.

Kidwell, M. G., Kidwell J. F. and Sved J. A. 1977. Hybrid dysgenesis in *Dros-ophilamelanogaster*: A syndrome of aberrant traits including mutation, sterili ty and male recombination. *Genetics* 86:813–833.

Kim, H. S., T. Murphy, J. Xia, D. Caragea, Y. Park, R. W. Beeman, M. D. Lorenzen, S. Butcher, J. R. Manak, and S. J. Brown. 2010. BeetleBase in 2010: revisions to provide comprehensive genomic information for Tribolium castaneum. Nucleic Acids Res. 38:D437–42.

Kim, T.-M., Hong S.-J. and Rhyu M.-G. 2004. Periodic explosive expansion of human retroelements associated with the evolution of the hominoid pri- mate. *J. Korean Med. Sci.* 19:177–185.

Kofler, R., R. V. Pandey, and C. Schlötterer. 2011. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27:3435–3436.

Kunarso, G., Chia N. Y., Jeyakani J., Hwang C. and Lu X. 2010. Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nature* 42:631-634.

Labrador, M., Farre M., Utzet F. and Fontdevila A. 1999. Interspecific hybridization increases transposition rates of *Osvaldo*. *Mol Biol Evol* 16: 931–937.

Lande, R. 1985. The fixation of chromosomal rearrangements in a subdivided population with local extinction and colonization. *Heredity* 54:323–332.

Laurie CC. 1997. The weaker sex is heterogametic: 75 years of Haldane's rule. Genetics. 147(3):937–951.

Leder EH, Cano JM, Leinonen T, O'Hara RB, Nikinmaa M, Primmer CR, Merila J. 2010. Female-biased expression on the X chromosome as a key step in sex chromosome evolution in threespine sticklebacks. Mol Biol Evol. 27(7):1495–1503.

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

Li, H., and R. Durbin. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079.

Li, W. L., Zhang P., Fellers J. P., Friebe B. and Gill B. S. 2004. Sequence composition, organization, and evolution of the core *Triticeae* genome. *Plant J.* 40:500–511.

Lim, J. K. and Simmons M. J. 1994. *Gross chromosome rearrangements mediated by transposable elements in Drosophila melanogaster*. 16:269–275.

Linsley, E. G. 1942. Insect food caches as reservoirs and original sources of some stored products pests. J. econ. Ent. 35:434–439.

Linsley, E. G. 1944. Natural sources, habitats and reservoirs of insects associated with stored food products. Hilgardia 16:187–224.

Lisch, D. 2009. Epigenetic regulation of transposable elements in plants. *Annu Rev Plant Biol* 60:43–66.

Lister, C., Jackson D. and Martin C. 1993. Transposon-induced inversion in Antirrhinum modifies nivea gene expression to give a novel flower color pattern under the control of cycloidearadialis. *Plant Cell* 5:1541–1553.

Litman, G. W., Rast J. P., and Fugmann S. D. 2010. The origins of vertebrate adaptive

immunity. *Nat Rev Immunology*, 10:543–553.

Lorenzen, M. D., Gnirke A., Margolis J., Garnes J., Campbell M., Stuart J. J., Aggarwal R., Richards S., Park Y. and Beeman R. W. 2008. The maternal-effect, selfish genetic element *Medea* is associated with a composite Tc1 transposon. *Proc Natl Acad Sci USA* 105:10085–10089.

Lorenzi, H., Thiagarajan M., Haas B., Wortman J., Hall N. and Caler E. 2008. Genome wide survey, discovery and evolution of repetitive elements in three *Entamoeba* species. *BMC Genomics* 9:595.

Lowry, D. B. and Willis J. H. 2010. A Widespread Chromosomal Inversion Polymorphism Contributes to a Major Life-History Transition, Local Adaptation, and Reproductive Isolation. *PLoS Biol* 8:e1000500.

Lu, C., Chen J., Zhang Y., Hu Q., Su W. and Kuang H. 2012. Miniature Inver- ted-Repeat Transposable Elements (MITEs) Have Been Accumulated through Amplification Bursts and Play Important Roles in Gene Expression and Species Diversity in *Oryza sativa*. *Mol Biol Evol* 29: 1005–1017.

Luan L, Liu JG, Hoja MR, Lightfoot DA, Hoog C. 2001. The checkpoint monitoring chromosomal pairing in male meiotic cells is p53-independent. Cell Death Differ. 8(3):316–317.

Lucchesi JC. 1973. Dosage compensation in Drosophila. Annu Rev Genet. 7(1):225–237.

Lynch, V. J., Leclerc R. D., May G. and Wagner G. P. 2011. Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. *Nature Reviews Genetics* 43:1154–1159.

Lyon MF. 1961. Gene action in the X chromosome of the mouse (Mus musculus L.). Nature. 190:372–373.

Machado, C. A., Haselkorn T. S. and Noor M. A. F. 2007b. Evaluation of the genomic extent of

effects of fixed inversion differences on intraspecific variation and interspecific gene flow in *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 175:1289–1306.

Machado, C. A., Matzkin L. M., Reed L. K. and Markow T. A. 2007a. Multilocus nuclear sequences reveal intra- and interspecific relationships among chromosomally polymorphic species of cactophilic *Drosophila*. *Mol Ecol* 16:3009–3024.

Maheshwari, S., and D. A. Barbash. 2011. The Genetics of Hybrid Incompatibilities. Annu Rev Genet 45:331–355.

Maksakova, I. A., Romanish M. T., Gagnier L., Dunn C. A., van de Lagemaat L. N. and Mager D. L. 2006. Retroviral Elements and Their Hosts: Insertional Mutagenesis in the Mouse Germ Line. *PLoS Genet* 2(1):e2.

Mallet, J. 2006. What does Drosophila genetics tell us about speciation? Trends Ecol Evol 21:386–393.

Malone JH, Fontenot BE. 2008. Patterns of reproductive isolation in toads. PLoS One. 3(12):1–11.

Malone, C. D. and Hannon G. J. 2009. Small RNAs as guardians of the genome. *Cell* 136:656–668.

Mandal, P. K. and Kazazian H. H. J. 2008. SnapShot: Vertebrate transposons. *Cell* 135:192–192.e1.

Mank JE, Hosken DJ, Wedell N. 2011. Some inconvenient truths about sex chromosome dosage compensation and the potential role of sexual conflict. Evolution. 65(8):2133–2144.

Margraf, R. L., J. D. Durtschi, S. Dames, D. C. Pattison, J. E. Stephens, R. Mao, and K. V. Voelkerding. 2010. Multi-sample pooling and illumina genome analyzer sequencing methods to determine gene sequence variation for database development. Journal of Biomolecular Techniques : JBT 21:126–140.

Market, E. and Papavasiliou F. N. 2003. V (D) J Recombination and the Evolution of the Adaptive Immune System. *PLoS Biol* 1:e16.

Martienssen, R. A. 2010. Heterochromatin, small RNA and post-fertilization dysgenesis in allopolyploid and interploid hybrids of *Arabidopsis*. *New Phytol* 186:46–53.

Mayr, E. 1942. *Systematics and the origin of species, from the viewpoint of a zoologist.* Columbia Univ. Press, New York.

McClintock, B. 1950. The Origin and Behavior of Mutable Loci in Maize. *Proc Natl Acad Sci USA* 36(6):344.

McClintock, B. 1956. Controlling Elements and the Gene. *Cold Spring Harb Symp Quant Biol* 21:197–216.

McClintock, B. 1980. Modified gene expressions induced by transposable elements. Pp. 11–19 *in* W. A. Scott, R. Werner, D. R. Joseph, and J. Schulz, eds. Mobilization and Reassembly of Genetic Information. Academic Press, London.

McClintock, B. 1984. The Significance of Responses of the Genome to Chal- lenge. *Science* 226:792–801.

McCormick, S. 1993. Male gametophyte development. Plant Cell 5:1265–1275.

McDermott SR, Noor MAF. 2010. The role of meiotic drive in hybrid male sterility. Philos Trans R Soc Lond B Biol Sci. 365:1265–1272.

McGaugh, S. E. and Noor M. A. F. 2012. Genomic impacts of chromosomal inversions in parapatric *Drosophila* species. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 367:422–429.

McKee B. 1997. Pairing sites and the role of chromosome pairing in meiosis and spermatogenesis in male Drosophila. Curr Top Dev Biol. 37:77–115.

Meiklejohn, C. D., M. A. Holmbeck, M. A. Siddiq, D. N. Abt, D. M. Rand, and K. L. Montooth. 2013. An Incompatibility between a mitochondrial tRNA and its nuclear-encoded tRNA synthetase compromises development and fitness in Drosophila. PLoS Genet

9:e1003238.

Metcalfe CJ, Bulazel KV, Ferreri GC, Schroeder-Reiter E, Wanner G, Rem W, Obergfell C, Eldridge MDB, O'Neill RJ. 2007. Genomic instability within centromeres of interspecific marsupial hybrids. Genetics. 177(4):2507–2517.

Metcalfe, C. J., Bulazel K. V., Ferreri G. C., Schroeder-Reiter E., Wanner G., Rem W., Obergfell C., Eldridge M. D. B. and O'Neill R. J. 2007. Genomic instability within centromeres of interspecific marsupial hybrids. *Genetics* 177:2507–2517.

Meyer, Y., B. B. Buchanan, F. Vignols, and J.-P. Reichheld. 2009. Thioredoxins and Glutaredoxins: Unifying Elements in Redox Biology. Annu Rev Genet 43:335–367.

Michalak, P. 2009. Epigenetic, transposon and small RNA determinants of hybrid dysfunctions. *Heredity* 102:45–50.

Michel, A. P. 2006. Divergence With Gene Flow in *Anopheles funestus* From the Sudan Savanna of Burkina Faso, West Africa. *Genetics* 173:1389–1395.

Mihola, O., Z. Trachtulec, C. Vlcek, J. C. Schimenti, and J. Forejt. 2009. A mouse speciation gene encodes a meiotic histone H3 methyltransferase. Science 323:373–375.

Moran, T., and A. Fontdevila. 2014. Genome-Wide Dissection of Hybrid Sterility in Drosophila Confirms a Polygenic Threshold Architecture. J Hered, doi: 10.1093/jhered/esu003.

Muller HJ, Pontecarvo G. 1942. Recessive genes causing interspecific sterility and other disharmonies between Drosophila melanogaster and simulans. Genetics. 27:157.

Muller HJ. 1940. Bearing of the Drosophila work on systematics. In: Huxley JS, editor. The new systematics. Oxford: Clarendon Press. p. 185–268.

Muller HJ. 1942. Isolating mechanisms, evolution, and temperature. Biol Symp. 6:71–125.

Nadeau, N. J., Whibley A., Jones R. T., Davey J. W., Dasmahapatra K. K., Baxter S. W., Quail M. A., Joron M., ffrench-Constant R. H., Blaxter M. L., Mallet J. and Jiggins C. D. 2012. Genomic islands of divergence in hybridizing *Heliconius* butterflies identified by large-

scale targeted sequencing. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 367:343–353.

Naito, K., Zhang F., Tsukiyama T., Saito H., Hancock C. N., Richardson A. O., Okumoto Y., Richardson A. O., Okumoto Y., Tanisaka T. and Wessler S. R. 2009. Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. *Nature* 461:1130–1134.

Namekawa SH, VandeBerg JL, McCarrey JR, Lee JT. 2007. Sex chromosome silencing in the marsupial male germ line. Proc Natl Acad Sci U S A. 104(23):9730–9735.

Navarro, A. and Barton N. H. 2003. Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. *Evolution* 57:447–459.

Nguyen DK, Disteche CM. 2006. Dosage compensation of the active X chromosome in mammals. Nat Genet. 38(1):47–53.

Noor, M. A. F. and Bennett S. M. 2009. Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity* 103:439–444.

Noor, M. A. F., Grams K., Bertucci L. and Reiland J. 2001. Chromosomal inversions and the reproductive isolation of species. *Proc Natl Acad Sci USA* 98:12084–12088.

Nosil, P. 2012. Ecological Speciation. Oxford University Press, New York.

Nosil, P. and Schluter D. 2011. The genes underlying the process of specia- tion. *Trends Ecol. Evol.* 26:160–167.

Nosil, P., and D. Schluter. 2011. The genes underlying the process of speciation. Trends Ecol Evol 26:160–167.

O'Donnell, K. A. and Boeke J. D. 2007. Mighty Piwis Defend the Germline against Genome Intruders. *Cell* 129:37–44.

O'Neill, R. J., O'Neill M. J. and Graves J. A. M. 1998. Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian

hybrid. *Nature* 393:68–72.

O'Neill RJ, O'Neill MJ, Graves JA. 1998. Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. Nature. 393(6680):68–72.

Oliver, K. R. and Greene W. K. 2009. Transposable elements: powerful facilitators of evolution. *Bioessays* 31:703–714.

Oliver, K. R. and Greene W. K. 2011. Mobile DNA and the TE-Thrust hypothesis: supporting evidence from the primates. *Mob DNA* 2:8.

Olmedo-Monfil, V., Durán-Figueroa N., Arteaga-Vázquez M., Demesa-Arévalo E., Autran D., Grimanelli D., Slotkin R. K., Martienssen R. A. and Vielle-Calzada J. P. 2010. Control of female gamete formation by a small RNA pathway in *Arabidopsis*. *Nature* 464:628–632.

Orr HA, Turelli M. 1996. Dominance and Haldane's rule. Genetics. 143(1):613–616.

Orr HA. 1993a. A mathematical model of Haldane's rule. Evolution. 47(5):1606–1611.

Orr HA. 1993b. Haldane's rule has multiple genetic causes. Nature. 361(6412):532–533.

Orr, H. A. 1995. THE POPULATION-GENETICS OF SPECIATION - THE EVOLUTION OF HYBRID INCOMPATIBILITIES. Genetics 139:1805–1813.

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

Orr, H. A. and Turelli M. 2001. The evolution of postzygotic isolation: Accumulating Dobzhansky-Muller incompatibilities. *Evolution* 55:1085–1094.

Orr, H. A., and S. Irving. 2005. Segregation distortion in hybrids between the Bogota and USA subspecies of Drosophila pseudoobscura. Genetics 169:671–682.

Ortiz-Barrientos, D. and Noor M. A. F. 2005. Evidence for a one-allele assortative mating locus. *Science* 310:1467–1467.

Ortiz-Barrientos, D., Counterman B. and Noor M. A. F. 2004. The genetics of speciation by reinforcement. *PLoS Biol* 2:2256–2263.

Pardo-Manuel de Villena F, Sapienza C. 2001. Female meiosis drives karyotypic evolution in mammals. Genetics. 159(3):1179–1189.

Park, T., D. B. Mertz, and K. Petrusewicz. 1961. Genetic Strains of *Tribolium*: Their Primary Characteristics. Physiological Zoology 34:62–80. The University of Chicago Press.

Pascale, E., Valle E. and Furano A. V. 1990. Amplification of an ancestral mammalian L1 family of long interspersed repeated DNA occurred just before the murine radiation. *Proc Natl Acad Sci USA* 87:9481–9485.

Pelisson, A. 1981. The *I-R* system of hybrid dysgenesis in *Drosophila melanogaster*: Are *I* factor insertions responsible for the mutator effect of the *I-R* interaction? *Mol Gen Genet* 183:123–129.

Pelisson, A. and Bregliano J. C. 1987. Evidence for rapid limitation of the *I* element copy number in a genome submitted to several generations of *I-R* hybrid dysgenesis in *Drosophila melanogaster*. *Mol Gen Genet* 207:306–313.

Pelisson, A., Song S. U., Prud'homme N., Smith P. A., Bucheton A. and Corces V. G. 1994. *Gypsy transposition correlates with the production of a retroviral envelope-like protein under the tissue-specific control of the Drosophila flamenco gene.* 13:4401.

Phadnis, N., and H. A. Orr. 2009. A Single Gene Causes Both Male Sterility and Segregation Distortion in Drosophila Hybrids. Science 323:376–379.

Picard, G. 1976. Non-Mendelian female sterility in *Drosophila melanogaster*: hereditary transmission of *I* factor. *Genetics* 83:107–123.

Picard, G. and L'Heritier P. 1971. A maternally inerited factor inducing sterility in *D. melanogaster. Drosophila Information Service* 46:54.

Presgraves DC, Orr HA. 1998. Haldane’s rule in taxa lacking a hemizygous X. Science.

282(5390):952–954.

Presgraves DC. 2002. Patterns of postzygotic isolation in Lepidoptera. Evolution. 56(6):1168–1183.

Presgraves, D. C. 2010. Darwin and the origin of interspecific genetic incompatibilities. Am Nat 176 Suppl 1:S45–S60.

Presgraves, D. C., L. Balagopalan, S. M. Abmayr, and H. A. Orr. 2003. Adaptive evolution drives divergence of a hybrid inviability gene between two species of Drosophila. Nature 423:715–719.

Prince EG, Kirkland D, Demuth JP. 2010. Hyperexpression of the X chromosome in both sexes results in extensive female bias of X-linked genes in the flour beetle. Genome Biol Evol. 2:336–346.

Prud'homme, N., Gans M., Masson M., Terzian C. and Bucheton A. 1995. Flamenco, a gene controlling the gypsy retrovirus of *Drosophila melanogaster*. *Genetics* 139:697–711.

Quinlan, A. R., Clark R. A., Sokolova S., Leibowitz M. L., Zhang Y., Hurles M. E., Mell J. C. and Hall I. M. 2010. Genome-wide mapping and assembly of structural variant breakpoints in the mouse genome. *Genome Research* 20:623–635.

Rabosky, D. L., and D. R. Matute. 2013. Macroevolutionary speciation rates are decoupled from the evolution of intrinsic reproductive isolation in Drosophila and birds. Proc Natl Acad Sci USA 110:15354–15359.

Ray, D. A., Feschotte C., Pagan H. J. T., Smith J. D., Pritham E. J., Arensburger P., Atkinson P. W. and Craig N. L. 2008. Multiple waves of recent DNA transposon activity in the bat, *Myotis lucifugus. Genome Res* 18:717–728.

Reiss, D., Josse T., Anxolabehere D. and Ronsseray S. 2004. *aubergine* mutations in *Drosophila melanogaster* impair *P* cytotype determination by telomeric *P* elements inserted in heterochromatin. *Molecular Genetics and Genomics* 272:336–343.

Renaut, S., Nolte A. W. and Bernatchez L. 2010. Mining transcriptome sequences towards identifying adaptive single nucleotide polymorphisms in lake whitefish species pairs (*Coregonusspp*. Salmonidae). *Mol Ecol* 19: 115–131.

Ribeiro, J. M. and Kidwell M. G. 1994. Transposable elements as population drive mechanisms: specification of critical parameter values. *J Med Ento- mol* 31:10–16.

Richards, S., R. A. Gibbs, G. M. Weinstock, S. J. Brown, R. E. Denell, R. W. Beeman, R. A. Gibbs, R. W. Beeman, S. J. Brown, G. Bucher, M. Friedrich, C. J. P. Grimmelikhuijzen, M. Klingler, M. D. Lorenzen, S. Richards, S. Roth, R. Schröder, D. Tautz, E. M. Zdobnov, D. M. Muzny, R. A. Gibbs, G. M. Weinstock, T. Attaway, S. Bell, C. J. Buhay, M. N. Chandrabose, D. Chavez, K. P. Clerk-Blankenburg, A. Cree, M. Dao, C. Davis, J. Chacko, H. Dinh, S. Dugan-Rocha, G. Fowler, T. T. Garner, J. Garnes, A. Gnirke, A. Hawes, J. Hernandez, S. Hines, M. Holder, J. Hume, S. N. Jhangiani, V. Joshi, Z. M. Khan, L. Jackson, C. Kovar, A. Kowis, S. Lee, L. R. Lewis, J. Margolis, M. Morgan, L. V. Nazareth, N. Nguyen, G. Okwuonu, D. Parker, S. Richards, S.-J. Ruiz, J. Santibanez, J. Savard, S. E. Scherer, B. Schneider, E. Sodergren, D. Tautz, S. Vattahil, D. Villasana, C. S. White, R. Wright, Y. Park, R. W. Beeman, J. Lord, B. Oppert, M. D. Lorenzen, S. J. Brown, L. Wang, J. Savard, D. Tautz, S. Richards, G. M. Weinstock, R. A. Gibbs, Y. Liu, K. Worley, G. M. Weinstock, C. G. Elsik, J. T. Reese, E. Elhaik, G. Landan, D. Graur, P. Arensburger, P. Atkinson, R. W. Beeman, J. Beidler, S. J. Brown, J. P. Demuth, D. W. Drury, Y.-Z. Du, H. Fujiwara, M. D. Lorenzen, V. Maselli, M. Osanai, Y. Park, H. M. Robertson, Z. Tu, J.-J. Wang, S. Wang, S. Richards, H. Song, L. Zhang, E. Sodergren, D. Werner, M. Stanke, B. Morgenstern, V. Solovyev, P. Kosarev, G. Brown, H.-C. Chen, O. Ermolaeva, W. Hlavina, Y. Kapustin, B. Kiryutin, P. Kitts, D. Maglott, K. Pruitt, V. Sapojnikov, A. Souvorov, A. J. Mackey, R. M. Waterhouse, S. Wyder, E. M. Zdobnov, E. M. Zdobnov, S. Wyder, E. V. Kriventseva, T.

Kadowaki, P. Bork, M. Aranda, R. Bao, A. Beermann, N. Berns, R. Bolognesi, F. Bonneton, D. Bopp, S. J. Brown, G. Bucher, T. Butts, A. Chaumot, R. E. Denell, D. E. K. Ferrier, M. Friedrich, C. M. Gordon, M. Jindra, M. Klingler, Q. Lan, H. M. G. Lattorff, V. Laudet, C. von Levetsow, Z. Liu, R. Lutz, J. A. Lynch, R. N. da Fonseca, N. Posnien, R. Reuter, S. Roth, J. Savard, J. B. Schinko, C. Schmitt, M. Schoppmeier, R. Schröder, T. D. Shippy, F. Simonnet, H. Marques-Souza, D. Tautz, Y. Tomoyasu, J. Trauner, M. Van der Zee, M. Vervoort, N. Wittkopp, E. A. Wimmer, X. Yang, A. K. Jones, D. B. Sattelle, P. R. Ebert, D. Nelson, J. G. Scott, R. W. Beeman, S. Muthukrishnan, K. J. Kramer, Y. Arakane, R. W. Beeman, Q. Zhu, D. Hogenkamp, R. Dixit, B. Oppert, H. Jiang, Z. Zou, J. Marshall, E. Elpidina, K. Vinokurov, C. Oppert, Z. Zou, J. Evans, Z. Lu, P. Zhao, N. Sumathipala, B. Altincicek, A. Vilcinskas, M. Williams, D. Hultmark, C. Hetru, H. Jiang, C. J. P. Grimmelikhuijzen, F. Hauser, G. Cazzamali, M. Williamson, Y. Park, B. Li, Y. Tanaka, R. Predel, S. Neupert, J. Schachtner, P. Verleyen, F. Raible, P. Bork, M. Friedrich, K. K. O. Walden, H. M. Robertson, S. Angeli, S. Forêt, G. Bucher, S. Schuetz, R. Maleszka, E. A. Wimmer, R. W. Beeman, M. D. Lorenzen, Y. Tomoyasu, S. C. Miller, D. Grossmann, and G. Bucher. 2008. The genome of the model beetle and pest Tribolium castaneum. Nature 452:949–955. Nature Publishing Group.

Richardson BJ, Czuppon AB, Sharman GB. 1971. Inheritance of glucose-6- phosphate dehydrogenase variation in Kangaroos. Nature. 230(13):154–155.

Rieseberg, L. H. 2001. Chromosomal rearrangements and speciation. *Trends Ecol. Evol.* 16:351–358.

Rose, M. R. and Doolittle W. F. 1983. Molecular biological mechanisms of speciation. *Science* 220:157–162.

Royo H, Polikiewicz G, Mahadevaiah SK, Prosser H, Mitchell M, Bradley A, de Rooij DG,

Burgoyne PS, Turner JMA. 2010. Evidence that meiotic sex chromosome inactivation is essential for male fertility. Curr Biol. 20(23):2117–2123.

Rubin, G. M., Kidwell M. G. and Bingham P. M. 1982. The molecular basis of *P-M* hybrid dysgenesis: The nature of induced mutations. *Cell* 29:987–994.

Sarot, E., Payen-Groschene G., Bucheton A. and Pelisson A. 2004. Evidence for a piwi-dependent RNA silencing of the gypsy endogenous retrovirus by the *Drosophila melanogaster* flamenco gene. *Genetics* 166:1313–1321.

Sasa M, Chippindale P, Johnson N. 1998. Patterns of postzygotic isolation in frogs. Evolution. 52(6):1811–1820.

Sawamura, K., M. T. YAMAMOTO, and T. K. WATANABE. 1993. Hybrid Lethal Systems in the Drosophila-Melanogaster Species Complex .2. the Zygotic-Hybrid-Rescue (Zhr) Gene of Drosophila-Melanogaster. Genetics 133:307–313.

Schaefer, R. E., Kidwell M. G. and Fausto-Sterling A. 1979. Hybrid Dysgenesis in *Drosophila melanogster*: Morphological and Cytological Studies of Ovarian Dysgenesis. *Genetics* 92:1141–1152.

Schilthuizen M, Giesbers MC, Beukeboom LW. 2011. Haldane's rule in the 21st century. Heredity. 107:95–102.

Schlipalius, D. I., N. Valmas, A. G. Tuck, R. Jagadeesan, L. Ma, R. Kaur, A. Goldinger, C. Anderson, J. Kuang, S. Zuryn, Y. S. Mau, Q. Cheng, P. J. Collins, M. K. Nayak, H. J. Schirra, M. A. Hilliard, and P. R. Ebert. 2012. A Core Metabolic Enzyme Mediates Resistance to Phosphine Gas. Science 338:807–810.

Schmidt, D., Schwalie P. C., Wilson M. D., Ballester B., Gonçalves Â., Kutter C., Brown G. D., Marshall A., Flicek P. and Odom D. T. 2012. Waves of Retrotransposon Expansion Remodel Genome Organization and CTCF Binding in Multiple Mammalian Lineages. *Cell* 148:335–348.

Schwarzbach, A. E. and Rieseberg L. H. 2002. Likely multiple origins of a diploid hybrid sunflower species. *Mol Ecol* 11:1703–1715.

Seehausen, O., R. K. Butlin, I. Keller, C. E. Wagner, J. W. BOUGHMAN, P. A. Hohenlohe, C. L. Peichel, G.-P. Saetre, C. Bank, Å. Brännström, A. Brelsford, C. S. Clarkson, F. Eroukhmanoff, J. L. Feder, M. C. Fischer, A. D. Foote, P. Franchini, C. D. Jiggins, F. C. Jones, A. K. Lindholm, K. Lucek, M. E. Maan, D. A. Marques, S. H. Martin, B. Matthews, J. I. Meier, M. Möst, M. W. Nachman, E. Nonaka, D. J. Rennison, J. Schwarzer, Eric T Watson, A. M. Westram, and A. Widmer. 2014. Genomics and the origin of species. Nat Rev Genet 15:176–192.

Shan, X., Liu Z., Dong Z., Wang Y., Chen Y., Lin X., Long L., Han F., Dong Y. and Liu B. 2005. Mobilization of the active MITE transposons mPing and Pong in rice by introgression from wild rice (*Zizania latifolia Griseb*.). *Mol Biol Evol* 22:976–990.

Sharman G, Close R, Maynes G. 1990. Chromosome evolution, phylogeny, and speciation of rock wallabies (Petrogale, Macropodidae). Aust J Zool. 37(2–4):351–363.

Sharp P. 1982. Sex chromosome pairing during male meiosis in marsupials. Chromosoma. 86(1):27–47.

Shedlock, A. Takahashi M., K. and Okada N. 2004. SINEs of speciation: tracking lineages with retroposons. *Trends Ecol. Evol.* 19:545–553.

Slotkin, R. K., Freeling M. and Lisch D. 2003. *Mu* killer causes the heritable inactivation of the Mutator family of transposable elements in *Zea mays*. *Genetics* 165:781–797.

Slotkin, R. K., Freeling M.and Lisch D. 2005. Heritable transposon silencing initiated by a naturally occurring transposon inverted duplication. *Nat. Rev. Genet.* 37:641–644.

Slotkin, R. K., Vaughn M., Borges F., Tanurdžić M., Becker J. D., Feijó J. A. and Martienssen R. A. 2009. Epigenetic Reprogramming and Small RNA Silencing of Transposable Elements in Pollen. *Cell* 136:461–472.

Sokoloff, A. 1974. The biology of Tribolium, with special emphasis on genetic aspects. Clarendon Press, Oxford.

Swanson WJ, Nielsen R, Yang Q. 2003. Pervasive adaptive evolution in mammalian fertilization proteins. Mol Biol Evol. 20(1):18–20.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Tao Y, Hartl DL. 2003. Genetic dissection of hybrid incompatibilities between Drosophila simulans and D. mauritiana. III. Heterogeneous accumulation of hybrid incompatibilities, degree of dominance, and implications for Haldane's Rule. Evolution. 57:2580–2598.

The Marie Curie SPECIATION Network. 2011. What do we need to know about speciation? Trends Ecol Evol 27:27–39.

Thompson, J. N., Woodruff R. C. and Schaefer G. B. 1978. An assay of somatic recombination in male recombination lines of *Drosophila melanogaster*. *Genetica* 49:77–80.

Thomson, M. and Beeman R. W. 1999. Assisted suicide of a selfish gene. *J Hered* 90:191–194.

Thomson, M. S., Friesen K. S., Denell R. E. and Beeman R. W. 1995. A hybrid income-patibility factor in *Tribolium castaneum*. *J Hered* 86:6.

Thomson, M. S., K. S. Friesen, R. E. Denell, and R. W. Beeman. 1995. A hybrid incompatibility factor in *Tribolium castaneum*. J Hered 86:6. Am Genetic Assoc.

Tiffin, P., S. Olson, and L. C. Moyle. 2001. Asymmetrical crossing barriers in angiosperms. … of the Royal ….

Ting, C.-T., S. C. Tsaur, M.-L. Wu, and C.-I. Wu. 1998. A rapidly evolving homeobox at the site of a hybrid sterility gene. Science 282:1501–1504.

Torgerson DG, Kulathinal RJ, Singh RS. 2002. Mammalian sperm proteins are rapidly evolving: evidence of positive selection in functionally diverse genes. Mol Biol Evol.

19(11):1973–1980.

Trier, C. N., J. S. Hermansen, G.-P. Saetre, and R. I. Bailey. 2014. Evidence for Mito-Nuclear and Sex-Linked Reproductive Barriers between the Hybrid Italian Sparrow and Its Parent Species. PLoS Genet 10:e1004075. Public Library of Science.

Turelli M, Moyle LC. 2007. Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. Genetics. 176(2):1059–1088.

Turelli M, Orr HA. 1995. The dominance theory of Haldane's rule. Genetics. 140(1):389–402.

Turelli M, Orr HA. 2000. Dominance, epistasis and the genetics of postzygotic isolation. Genetics. 154(4):1663–1679.

Turelli, M. and Moyle L. C. 2007. Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. *Genetics* 176:1059–1088.

Turner, T. L. and Hahn M. W. 2010. Genomic islands of speciation or genomic islands and speciation? *Mol Ecol* 19:848–850.

Turner, T. L., Hahn M. W. and Nuzhdin S. V. 2005. Genomic Islands of Speciation in *Anopheles gambiae. PLoS Biol* 3:e285.

Ungerer, M. C., Strakosh S. C. and Stimpson K. M. 2009. Proliferation of Ty3/gypsy-like retrotransposons in hybrid sunflower taxa inferred from phylogenetic data. *Bmc Biol* 7:40.

Ungerer, M. C., Strakosh S. C. and Zhen Y. 2006. Genome expansion in three hybrid sunflower species is associated with retrotransposon proliferation. *Curr. Biol.* 16:R872–R873.

Venner, S., Feschotte C. and Biémont C. 2009. Dynamics of transposable elements: towards a community ecology of the genome. *Trends Genet* 25: 317–323.

Verneau, O., Catzeflis F. and Furano A. V. 1998. Determining and dating recent rodent speciation events by using L1 (LINE-1) retrotransposons. *Proc Natl Acad Sci USA* 95:11284–11289.

Wade, M. J. 1977. An experimental study of group selection. Evolution 134–153. JSTOR.

Wade, M. J. and Beeman R. W. 1994. The population dynamics of maternal-effect selfish genes. *Genetics* 138:1309–1314.

Wade, M. J., and N. A. Johnson. 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 (Pt 2):155–162.

Wade, M. J., N. A. Johnson, R. Jones, V. Siguel, and M. McNaughton. 1997. Genetic variation segregating in natural populations of Tribolium castaneum affecting traits observed in hybrids with T. freemani. Genetics 147:1235–1247.

Walker, E. L., Robbins T. P., Bureau T. E., Kermicle J. and Dellaporta S. L. 1995. Transposon-mediated chromosomal rearrangements and gene dupli- cations in the formation of the maize *R-r* complex. *EMBO J* 14:2350–2363.

Walsh, J. B. 1982. Rate of Accumulation of Reproductive Isolation by Chromosome Rearrangements. *Am. Nat.* 120:510–532.

Wang, L., S. Wang, Y. Li, M. S. R. Paradesi, and S. J. Brown. 2007. BeetleBase: the model organism database for Tribolium castaneum. Nucleic Acids Res. 35:D476–9.

Wang, N., Wang H., Zhang D., Wu Y., Ou X., Liu S., Dong Z. and Liu B. 2010. Transpositional reactivation of the Dart transposon family in rice lines derived from introgressive hybridization with *Zizania latifolia. BMC Plant Biol* 10:190.

Wang, S., Lorenzen M. D., Beeman R. W. and Brown S. J. 2008. Analysis of repetitive DNA distribution patterns in the *Tribolium castaneum* genome. *Genome Biol* 9:R61.

Wang, S., M. D. Lorenzen, R. W. Beeman, and S. J. Brown. 2008. Analysis of repetitive DNA distribution patterns in the *Tribolium castaneum* genome. Genome Biol 9:R61.

Watanabe, T. K. 1979. A gene that rescues the lethal hybrids between. Drosophila melanogaster and D. simulans. Japanese Journal of Genetics 325–331.

Welch, M. E. and Rieseberg L. H. 2002. Patterns of Genetic Variation Suggest a Single, Ancient Origin for the Diploid Hybrid Species *Helianthus paradoxus*. *Evolution* 56:2126–2137.

Wessler, S. R. 1996. Plant retrotransposons: Turned on by stress. *Cell* 6:959–961.

Wicker, T., Sabot F., Hua-Van A., Bennetzen J. L., Capy P., Chalhoub B., Flavell A., Leroy P., Morgante M., Panaud O., Paux E., SanMiguel P. and Schulman A. H. 2007. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* 8:973–982.

Willett, C. S., and R. S. Burton. 2001. Viability of cytochrome c genotypes depends on cytoplasmic backgrounds in Tigriopus californicus. Evolution 55:1592–1599.

Woodruff, R. C. and Thompson J. N. 1977. An analysis of spontaneous recombination in I males. *Heredity* 38:291–307.

Wright, S. 1931. Evolution in Mendelian Populations. *Genetics* 16:97–159.

Wright, S. 1932. *The Roles of Mutation, Inbreeding, Crossbreeding, and Selection in Evolution. in* D. F. Jones, ed. Brooklyn Botanic Garden, Ithaca, New York.

Wu C-I, Davis A. 1993. Evolution of postmating reproductive isolation: the composite nature of Haldane's rule and its genetic basis. Am Nat. 142(2):187–212.

Wu C, Johnson N, Palopoli M. 1996. Haldane's rule and its legacy: why are there so many sterile males? Trends Ecol Evol. 11:281–284.

Xing, J., Zhang Y., Han K., Salem A. H., Sen S. K., Huff C. D., Zhou Q., Kirkness E. F., Levy S. and Batzer M. A. 2009. Mobile elements create structural variation: analysis of a complete human genome. *Genome Research* 19:1516–1526.

Xiong Y, Chen X, Chen Z, Wang X, Shi S, Wang X, et al. 2010. RNA sequencing shows no dosage compensation of the active X-chromosome. Nat Genet. 42(12):1043–1049.

Yamaguchi, O., Cardellino R. A. and Mukai T. 1976. High Rates of Occurrence of Spontaneous

Chromosome Aberrations in Drosophila mela- nogaster. *Genetics* 83:409.

Zeh, D. W., Zeh J. A. and Ishida Y. 2009. Transposable elements and an epigenetic basis for punctuated equilibria. *Bioessays* 31:715–726.

Zha X, Xia Q, Duan J, Wang C, He N, Xiang Z. 2009. Dosage analysis of Z chromosome genes using microarray in silkworm, Bombyx mori. Insect Biochem Mol Biol. 39:315–321.

Zhang, J., Yu C., Pulletikurti V., Lamb J., Danilova T., Weber D. F., Birchler J. and Peterson T. 2009. Alternative *Ac/Ds* transposition induces major chromosomal rearrangements in maize. *Gene Dev* 23:755–765.

Zhou, L., Mitra R., Atkinson P. W., Burgess Hickman A., Dyda F. and Craig N. L. 2004. Transposition of hAT elements links transposable elements and V (D) *J recombination.* 432:995–1001.

Zuryn, S., J. Kuang, and P. Ebert. 2008. Mitochondrial modulation of phosphine toxicity and resistance in Caenorhabditis elegans. Toxicol. Sci. 102:179–186.

Zwick ME, Salstrom JL, Langley CH. 1999. Genetic variation in rates of nondisjunction: association of two naturally occurring polymorphisms in the chromokinesin nod with increased rates of nondisjunction in Drosophila melanogaster. Genetics. 152(4):1605–1614.

Biographical Information

Eric received his Bachelor's Degree in Biological Sciences from Louisiana State University in 2006, where he began researching speciation genetics in *Drosophila* in the laboratory of Mohamed Noor, now at Duke University. After a couple years of work on the genetics and geography of ethanol preference and performance in *Drosophila pseudoobscura* and *D. persimilis*, he spent a summer performing research in the laboratory of Jerry Coyne at the University of Chicago. There, he performed an analysis of the genetics of courtship song divergence between the sister species *D. yakuba* and *D. santomea*. Upon returning to Louisiana State University, Dr. Watson spent time performing independent research in the laboratory of David Donze, working with chromatin remodelling protein domains in the yeast *Saccharomyces cerevisiae*, followed by work in the laboratory of Andrew Whitehead, now at the University of California at Davis on population genetics in the Mongolian salmonid *Hucho taimen.*

Upon graduation with his BS, Dr. Watson moved to the University of Arizona, where he worked on the National Geographic's Human Genographic project in the Human Origins Genotyping Laboratory as a Research Technician. The following year, Dr. Watson became a Research Specialist in the University of Arizona Genomics Core facility, where he ran fragment analysis on an ABI 3730, and characterized novel nuclear markers for conservation genetics in species ranging from the invasive whitefly, *Bemesia tabaccii,* various species of *Fundulus* killifish, and in the viper genera *Cerastes* and *Crotalus.*

Dr. Watson will continue his research on the genetics of hybrid incompatibility as a postdoctoral researcher at the University of Southern California in the laboratory of Suzanne Edmands. Here, he will work with the intertidal copepod *Tigriopus californicus* to understand the forces driving the evolution of segregating cytonuclear incompatibilities.