Stony Brook University

Life after beanbag genetics:

theoretical and empirical studies on epistasis and penetrance

A Dissertation Presented

by

Joseph Lachance

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Genetics

Stony Brook University

December 2010

Stony Brook University

The Graduate School

Joseph Lachance

We, the dissertation committee for the above candidate for the Doctor of Philosophy Degree, hereby recommend acceptance of the dissertation.

John R. True, Dissertation Advisor
Associate Professor, Department of Ecology and Evolution, Stony Brook University

Walter F. Eanes, Chairperson of the Defense
Professor, Department of Ecology and Evolution, Stony Brook University

Daniel E. Dykhuizen
Distinguished Professor, Department of Ecology and Evolution, Stony Brook University

Joshua Rest
Assistant Professor, Department of Ecology and Evolution, Stony Brook University

Stephan B. Munch, Outside Committee Member
Associate Professor, School of Marine and Atmospheric Sciences, Stony Brook University

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

Life after beanbag genetics:

theoretical and empirical studies on epistasis and penetrance

by

Joseph Lachance

Doctor of Philosophy

in

Genetics

Stony Brook University

2010

Population genetics reduces the description of a population to a set of genotype frequencies. Because of this, it has been called "beanbag genetics." However, greater complexity and biological realism can be incorporated into this framework. Two examples of this include interactions between different genes (epistasis) and individuals with the same genotype that have different phenotypes (incomplete penetrance). During the course of my Ph.D. work, both of these topics were investigated using theoretical population genetics and *Drosophila* genetics. One unifying theme was that genes do not exist in isolation: they occur within genetic and environmental contexts. In turn, these contexts were found to affect the evolution of populations.

The above considerations were addressed via four interlocking studies. First, we studied epistatic interactions between naturally segregating chromosomes in *D. melanogaster*. We

placed a number of X chromosomes into multiple genetic backgrounds, finding that many X-autosome combinations were lethal or sterile. X-autosome incompatibilities also exhibited sex-specific and geographic patterns. Second, population genetics theory was extended to encompass X-autosome interactions. Allele frequency trajectories were calculated and found to be consistent with relatively high levels of segregating variation. This also led to testable predictions about the early stages of Haldane's rule. Third, we studied the incomplete penetrance of a naturally occurring mutation affecting wing development in *D. melanogaster*. This mutation was found to be allelic with the *vesiculated* gene. Subsequent tests examined how genetic background, temperature, sex, and maternal effects affected the presence and severity of wing vesiculation. Fourth, the theoretical population genetics of incomplete penetrance was developed. Mean fitness accurately predicted the evolutionary trajectories of beneficial alleles. However, fitness variance and maternal effects greatly affected the evolution of incompletely penetrant alleles that were neutral on average. Together, these four studies illustrate the complexity of natural genetic variation and suggest how it might have evolved.

TABLE OF CONTENTS

List of figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of tables

Acknowledgements

First, I would like to thank Alison Onstine. Because of her, I'm a better scientist and a better person. I would also like to thank my parents. Their support and willingness to hear me rattle on about population genetics has been much appreciated.

John True has been an exemplary mentor, colleague, and friend. His tolerance for my love of side projects has not gone unnoticed. John and Shian-Ren have done an amazing job of making the lab feel like home, and I feel very fortunate to have learned evolutionary genetics from both of them. Each of the members of the True Lab (Roman, Rocio, Shu-Dan, Lawrence, and Erin) has added to this familial feeling and enriched the quality of my science.

Discussions with the members of my PhD committee (Walt Eanes, Dan Dykhuizen, Josh Rest, and Steve Munch) have furthered my growth. Thanks again for making me feel like a peer. Kate, Jerry, Nadine and the rest of the Genetics Program have made my time at Stony Brook enjoyable. I'll also miss the collegiality of the Department of Ecology of Evolution.

I thank the following people for insightful discussions and helpful comments: R. Cogni, M. Cressy, K. Dyer, J. Flowers, R. Geeta, E. Hill-Burns, X. Hua, N. Johnson, M. Jonas, L. Jung, M. Kernan, S. Koury, S. Kumagai, E. Lavington, S.R. Liou, C. Morales, R. Ng, D. Opulente, S. Proulx, J. Reinitz, S. Salinas, S. Sabatino, M. Trotter, O. Warsi, S.D. Yeh, and R. Yukilevich. Additional thanks are directed towards S. Ahn, J. Hyder, L. Jung, and J. Lenci for assistance in some of the fly work. I am grateful to W. Eanes, P. Gergen, T. Merritt, J. Pool, and R. Yukilevich for providing *Drosophila* stocks. This work was supported in part by Stony Brook University, a National Institutes of Health grant (1RO1GM07171001) to John True, and a National Institutes of Health predoctoral training grant (5 T32 GM007964-24).

Two chapters in this dissertation were the result of collaborations with researchers other than my advisor, John True. Chapter 3 was completed with the assistance of Norman A. Johnson (University of Massachusetts, Amherst). His knowledge of speciation genetics and epistasis greatly improved the theory of sex-specific synthetic incompatibilities. Chapter 4 was completed with the assistance of Lawrence Jung. As an undergraduate in the True Lab, he helped determine the chromosomal basis of *vesiculated* suppression.

Additional publications

The following additional research publications were completed during my PhD:

Yukilevich, R., **J. Lachance**, F. Aoki, and J.R. True. 2008. Long-term adaptation of epistatic genetic networks. Evolution 62:2215-2235.

Lachance, J. 2008. A fundamental relationship between genotype frequencies and fitnesses. Genetics 180:1087-93.

Lachance, J. 2009. Detecting selection-induced departures from Hardy-Weinberg proportions. Genetics Selection Evolution 41:15.

Lachance, J. 2009. Inbreeding, pedigree size, and the most recent common ancestor of humanity. Journal of Theoretical Biology 261: 238-247.

Lachance, J. 2010. Genome-wide association studies reveal that disease associated alleles are enriched for derived low frequency alleles relative to HapMap and neutral expectations. BMC Medical Genomics (*under revision*)

In addition, Chapter 2 exists in a published form. As the purpose of this dissertation is non-commercial, the inclusion of Chapter 2 satisfies the copyright requirements of Wiley-Blackwell (see http://authorservices.wiley.com/bauthor/faqs_copyright.asp). The relevant citation is:

Lachance, J., and J.R. True. 2010. X-autosome incompatibilities in *Drosophila melanogaster:* tests of Haldane's rule and geographic patterns within species. Evolution 64: 3035-3046.

Chapter 1

Introduction

Genes exist within multiple contexts; they are found within organisms and populations. This latter level of organization is examined in population genetics, where genetic variation is represented as a set of allele frequencies. The metaphor of a "gene pool" enables mathematical treatments of evolutionary questions (Crow 2001). For example, in small populations one expects allele frequencies to vary stochastically and in large populations one expects allele frequency trajectories to be driven by natural selection. For diploid organisms, an additional level of complexity can be incorporated: populations can be represented as a set of genotype frequencies. While this pays lip-service to the fact that alleles are found within organisms, important aspects of biological reality are typically ignored. In response to this, Ernst Mayr's criticism of "beanbag genetics" is both a call for the inclusion of population thinking and an awareness that genes appear in different contexts (Mayr 1959). Not every organism with the same genotype will have the same fitness. In addition, the effects of genes can be buffered by other loci. One can then ask: What are the evolutionary ramifications of the fact that genes occur within organisms?

Drosophila melanogaster

Two of the chapters in this dissertation involve *D. melanogaster* as a model system. This species has a long history in both genetics and evolutionary biology. The first major research program involving *Drosophila* was founded by Thomas Hunt Morgan at Columbia University in 1907 (Kohler 1994). The short generation time and ease of maintaining stocks gave flies a comparative advantage as an early laboratory organism. Also, the existence of visual mutants (such as white eyes) lent empirical support for such fundamental genetic concepts as linkage and the fact that chromosomes contain genetic material (Carlson 2004). Studies of *Drosophila* soon left the laboratory, and researchers such as Timofeeff-Ressovsky and Dobzhansky pioneered the study of flies in natural populations during the 1920s and '30s (Powell 1997). As molecular techniques advanced, early evidence of rampant natural variation was detected in flies via the use of allozyme polymorphism and DNA sequence information. More recently, *Drosophila* research has entered the genomic age. The first *Drosophila melanogaster* genome was sequenced in 2000 (Adams et al. 2000), and the genomes of 12 different *Drosophila* species were sequenced in 2007 (Stark et al. 2007). Now we see the beginnings of population genomics, as multiple *D. melanogaster* and *D. simulans* individuals are being sequenced (Begun et al. 2007). Much of what we know about speciation genetics stems from studies of *Drosophila* (Coyne and Orr 2004), and there is evidence that pre- and post-zygotic incompatibilities segregate in divergent populations of *D. melanogaster* (Lachance and True 2010; Wu et al. 1995; Yukilevich and True 2008).

In addition to classical Mendelian crosses, a number of genetic tools exist for *D. melanogaster* that make it particularly suitable for this dissertation. One particularly useful tool

involves balancer chromosomes, which contain inversions, recessive lethal mutations, and visual markers (Greenspan 2004). The net effect of balancer chromosomes is to reduce genomes to a number of small units. This is aided by the fact that *D. melanogaster* has a relatively small number of chromosomes (one sex chromosome, two large autosomes, and a small dot chromosome). Through the use of balancer chromosomes, it is possible to generate homozygous genetic backgrounds, thereby eliminating possible sources of variation and experimental error. In addition, balancer chromosomes allow researchers to indefinitely maintain stocks containing recessive lethal mutations.

Additional resources include FlyBase and the Bloomington Stock Center. FlyBase is an online database that integrates over a century's worth of classical and molecular data (Tweedie et al. 2009). In addition to the annotated genome of *D. melanogaster*, FlyBase contains phenotypic information about known mutants, expression data, and a searchable database of *Drosophila* papers. FlyBase also contains lists of fly stocks that can be obtained from the Bloomington Stock Center. Continuing a long history of openness and resource sharing, lines can be ordered at a nominal fee. At present, over 27,000 different stocks are maintained at the Bloomington Stock Center. In addition to a number of lines collected from natural populations, multiple lines used in this dissertation were obtained from the Bloomington Stock Center.

Theoretical population genetics

Broadly speaking, theoretical population genetics involves the study of how gene pools evolve. The field of population genetics was largely pioneered by R. A. Fisher, J. B. S. Haldane, and Sewall Wright. Their work bridged the gap between Mendelian genetics and Darwinian

evolution, leading to the modern evolutionary synthesis (Provine 1971). Such concepts as fitness, effective population size, and genotype frequency space have been used to explain the dynamics of alleles in populations, and over the past few decades the theoretical toolkit available to population genetics has expanded to include diffusion approximations, coalescent theory, and computer simulations (Kimura 1964; Wakeley 2009).

In 1959 at Cold Spring Harbor, Ernst Mayr described population genetics as "beanbag genetics" because population states were reduced to a set of genotype frequencies (Mayr 1959). To Mayr, these simplified models were poor representations of biological phenomena in natural populations. Haldane wrote an entertaining rebuttal shortly after Mayr's comments (Haldane 1964), and a wealth of studies have validated the approaches of population genetics (Crow 2001). However, the stigma of beanbag genetics persisted for many years (Rao and Nanjundiah 2010). Empirical chapters in this dissertation provide examples of biological complexity, and theoretical chapters show how these details can be incorporated into advanced models of beanbag genetics.

Two metaphors aid in understanding beanbag genetics and Mayr's objection to this approach: Wright's fitness landscape and Waddington's epigenetic landscape. Wright envisioned a fitness surface where X and Y-axes refer to allele frequencies and the Z-axis corresponds to the mean fitness of a population (Wright 1932). Natural selection involves hill climbing on this fitness landscape. Waddington envisioned the process of development as a ball rolling down an epigenetic landscape (Waddington 1957). Epigenetic landscapes can vary in the extent in which they are canalized. Both of these metaphors can incorporate epistasis and incomplete penetrance (Figure 1). In Wright's framework, epistasis is visualized as a rugged

fitness landscape, where fitness depends on allele frequencies at multiple loci. Incomplete penetrance is visualized as a fitness seascape, where combinations of genotypes can map to multiple fitnesses. In Waddington's framework, epistasis can be visualized as a number of guy-wires tugging upon the epigenetic landscape, modifying the path of a canal. Genotypes are not canalized when penetrance is incomplete, resulting in multiple paths for the ball rolling down the epigenetic landscape.

Epistasis

Epistasis refers to genetic interactions between two or more alleles at different loci. A major distinction exists between functional and statistical epistasis (Phillips 1998; Phillips 2008). Functional epistasis involves alleles that interact with each other, either directly through physical interactions or indirectly via pathways. Conversely, statistical epistasis involves the partitioning of genetic variance in a quantitative genetics framework. This dissertation is largely concerned with functional epistasis. Two important types of epistatic interactions are synthetic lethality and synthetic sterility. Synthetic incompatibilities involve pairs of alleles that are neutral when found singly but are deleterious when both alleles are present in the same individual. This dissertation examines naturally segregating synthetic incompatibilities in *D. melanogaster* and the population genetic theory behind these interactions.

Multiple examples of fitness epistasis exist in natural populations. Experimental evolution of *Escherichia coli* populations has revealed that fitnesses of mutants are contingent upon genetic background (Lunzer et al. 2005; Weinreich et al. 2006). In *Tribolium castaneum* it was observed that X-autosome interactions directly influence fitness (Demuth and Wade 2007).

Synthetic lethal alleles have also been found segregate among common haplotypes of *Caenorhabitis elegans* (Seidel et al. 2008) and in populations of the copepod *Tigriopus californicus* (Harrison and Edmands 2006).

Epistasis influences the evolutionary trajectories of alleles. When alleles interact with other alleles, fitness becomes context-dependent (Eldar and Elowitz 2010). Theory indicates that some forms of epistasis cause evolutionary trajectories to be highly dependent on the order of mutations (Weinreich et al. 2005). Using a combination of mathematical modeling and individual-based simulations, we also found evidence that epistasis can result in saltational evolution (Yukilevich et al. 2008).

Penetrance

Penetrance can be defined as the proportion of individuals of a given genotype that show the expected phenotype. Expressivity, on the other hand, quantifies the degree that a phenotype is expressed in an individual or the average of this measure in a population of individuals. These terms were coined by the German neurologist Oskar Vogt subsequent to visiting Timofeeff-Ressovsky's lab in Russia during the mid-1920s (Laubichler and Sarkar 2002). Incomplete penetrance is ubiquitous in natural populations, and empirical examples have been found in a wide variety of species, including *Drosophila mercatorum*, *Arabidopsis thaliana*, and *Homo sapiens* (Hollocher et al. 1992; Ogas et al. 1997; Shields and Harris 2000). When fitnesses are considered, incomplete penetrance yields a one-to-two genotype-to-fitness mapping. Thus, the genotype-to-fitness map can be viewed as a simple form of stochastic fitness. Incomplete penetrance resembles recessivity in that phenotypes are masked. In addition, incomplete

penetrance involves an additional level of non-genetic drift stochasticity. Incomplete penetrance can also be viewed as increasing the number of alleles in a genetic system (with conversion between penetrant and non-penetrant forms analogous to bidirectional mutation). A compelling question that is addressed in this dissertation is how these aspects of incomplete penetrance influence the fates of alleles.

Multiple environmental and genetic causes underlie incomplete penetrance. Environmental or developmental stochasticity can cause individuals with the same genotype to have different phenotypes. This occurs either because phenotypes are determined by a small number of cells or because there are critical times in development (Eldar and Elowitz 2010; Raser and O'Shea 2005). A recent study demonstrated that variability in gene expression is a cause of incomplete penetrance. Individual transcripts were tracked in *C. elegans*, and epistatic interactions modified the penetrance of intestinal cell fate (Raj et al. 2010).

More generally, incomplete penetrance can be explained by the combination of threshold traits, developmental noise, and expression levels that vary by genotype (Figure 2). Consider a single diploid locus with two segregating alleles. Due to cis-regulatory differences, allele *A* is expressed at higher levels than allele *B*. Flies have defective wings when overall levels of expression are below an arbitrary threshold (the solid black line in Figure 2), and flies have wild-type wings if overall levels of expression are above the threshold. In Figure 2A, expression levels are highly canalized. Every *AA* individual has defective wings, and every *AB* and *BB* individual has wild-type wings. In Figure 2B, expression levels are not canalized (individuals with the same genotype have different levels of expression). However because variable levels of expression do not overlap the threshold, penetrance of each genotype is complete. In figure 2C, expression levels are not canalized but the presence of a different genetic background results in

higher levels of expression for all three genotypes. Because variable levels of expression overlap the threshold in this scenario, penetrance of wild-type wings in *AA* individuals is incomplete. This illustrates one way that epistasis can modify penetrance. Note that the independent variable need not be expression, and this concept also applies to soft-thresholds. Subsequent theory in this dissertation examines the ramifications of incomplete penetrance on a population scale.

Outline

Subsequent chapters contain empirical evidence of epistasis and incomplete penetrance. In addition, population genetics theory is extended to cover both of these topics.

Chapter 2: Levels of synthetic incompatibilities between naturally segregating X chromosomes and autosomes were assessed using *D. melanogaster* as a model system.

Chapter 3: Theoretical population genetics was extended to include X-autosome incompatibilities. Both synthetic lethality and synthetic sterility were considered.

Chapter 4: Studying natural variation at the *vesiculated* locus in *D. melanogaster*, we found that complex interactions underlie incomplete penetrance.

Chapter 5: Theoretical population genetics was extended to include incompletely penetrant alleles.

References

Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne, P. G. Amanatides, S. E. Scherer et al. 2000. The genome sequence of Drosophila melanogaster. Science 287:2185-2195.

Begun, D. J., A. K. Holloway, K. Stevens, L. W. Hillier, Y. P. Poh, M. W. Hahn, P. M. Nista et al. 2007. Population genomics: whole-genome analysis of polymorphism and divergence in Drosophila simulans. PLoS Biol 5:e310.

Carlson, E. A. 2004, Mendel's legacy: the origin of classical genetics. Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press.

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

Crow, J. F. 2001. The beanbag lives on. Nature 409:771.

Demuth, J. P., and M. J. Wade. 2007. Population differentiation in the beetle Tribolium castaneum. II. Haldane'S rule and incipient speciation. Evolution Int J Org Evolution 61:694-699.

Eldar, A., and M. B. Elowitz. 2010. Functional roles for noise in genetic circuits. Nature 467:167-173.

Greenspan, R. J. 2004, Fly Pushing: The Theory and Practice of *Drosophila* Genetics. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.

Haldane, J. B. 1964. A Defense of Beanbag Genetics. Perspect Biol Med 7:343-359.

Harrison, J. S., and S. Edmands. 2006. Chromosomal basis of viability differences in *Tigropus californicus* interpopulation hybrids. Journal of Evolutionary Biology 19:2040-2051.

Hollocher, H., A. R. Templeton, R. DeSalle, and J. S. Johnston. 1992. The molecular through ecological genetics of abnormal abdomen. IV. Components of genetic variation in a natural population of Drosophila mercatorum. Genetics 130:355-366.

Kimura, M. 1964. Diffusion models in population genetics. J. Appl. Prob. 1:177-232.

Kohler, R. E. 1994, Lords of the Fly: Drosophila Genetics and the Experimental Life. Chicago, University of Chicago Press.

Lachance, J., and J. R. True. 2010. X-autosome incompatibilities in *Drosophila melanogaster*: tests of Haldane's rule and geographic patterns within species. Evolution 64:3035-3046.

Laubichler, M. D., and S. Sarkar. 2002. Flies, genes, and brains: Oskar Vogt, Nikolai Timofeeff-Ressovsky, and the origin of the concepts of penetrance and expressivity *in* L. S. Parker, and R. A. Ankeny, eds. Mutating concepts, evolving disciplines: genes, medicine, and society. Philosophy and Medicine. Dordrecht, Kluvier.

Lunzer, M., S. P. Miller, R. Felsheim, and A. M. Dean. 2005. The biochemical architecture of an ancient adaptive landscape. Science 310:499-501.

Mayr, E. 1959. Where Are We? Cold Spring Harbor Symposium on Quantitative Biology 24:409-440.

Ogas, J., J. C. Cheng, Z. R. Sung, and C. Somerville. 1997. Cellular differentiation regulated by gibberellin in the Arabidopsis thaliana pickle mutant. Science 277:91-94.

Phillips, P. C. 1998. The language of gene interaction. Genetics 149:1167-1171.

—. 2008. Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet 9:855-867.

Powell, J. R. 1997, Progress and Prospects in Evolutionary Biology: The Drosophila Model. New York, Oxford University Press.

Provine, W. B. 1971, The Origins of Theoretical Population Genetics. Chicago, University of Chicago Press.

Raj, A., S. A. Rifkin, E. Andersen, and A. van Oudenaarden. 2010. Variability in gene expression underlies incomplete penetrance. Nature 463:913-918.

Rao, V., and V. Nanjundiah. 2010. J. B. S. Haldane, Ernst Mayr and the Beanbag Genetics Dispute. J Hist Biol.

Raser, J. M., and E. K. O'Shea. 2005. Noise in gene expression: origins, consequences, and control. Science 309:2010-2013.

Seidel, H. S., M. V. Rockman, and L. Kruglyak. 2008. Widespread Genetic Incompatibility in *C. elegans* Maintained by Balancing Selection. Science 319:589-594.

Shields, P. G., and C. C. Harris. 2000. Cancer risk and low-penetrance susceptibility genes in gene-environment interactions. J Clin Oncol 18:2309-2315.

Stark, A., M. F. Lin, P. Kheradpour, J. S. Pedersen, L. Parts, J. W. Carlson, M. A. Crosby et al. 2007. Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures. Nature 450:219-232.

Tweedie, S., M. Ashburner, K. Falls, P. Leyland, P. McQuilton, S. Marygold, G. Millburn et al. 2009. FlyBase: enhancing Drosophila Gene Ontology annotations. Nucleic Acids Res 37:D555-559.

Waddington, C. H. 1957, The Strategy of Genes. London, Allen & Unwin.

Wakeley, J. 2009, Coalescent theory : an introduction. Greenwood Village, Colo., Roberts & Co. Publishers.

Weinreich, D. M., N. F. Delaney, M. A. Depristo, and D. L. Hartl. 2006. Darwinian evolution can follow only very few mutational paths to fitter proteins. Science 312:111-114.

Weinreich, D. M., R. A. Watson, and L. Chao. 2005. Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. Evolution Int J Org Evolution 59:1165-1174.

Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proc. 6th International Congress of Genetics 1:356-366.

Wu, C. I., H. Hollocher, D. J. Begun, C. F. Aquadro, Y. Xu, and M. L. Wu. 1995. Sexual isolation in Drosophila melanogaster: a possible case of incipient speciation. Proc Natl Acad Sci U S A 92:2519-2523.

Yukilevich, R., J. Lachance, F. Aoki, and J. R. True. 2008. Long-term adaptation of epistatic genetic networks. Evolution 62:2215-2235.

Yukilevich, R., and J. R. True. 2008. Incipient sexual isolation among cosmopolitan Drosophila melanogaster populations. Evolution 62:2112-2121.

Figure legends

Figure 1. Four landscapes. In this version of Wright's fitness landscape, X and Y-axes refer to allele frequencies and the Z-axis corresponds to the mean fitness of a population. In Waddington's epigenetic landscape, the process of development is viewed as a ball rolling down a hill.

Figure 2. Expression thresholds, genetic backgrounds, and penetrance. Expression below a threshold (thick black line) results in flies with a wing defect, and expression above a threshold results in wild-type wings. Expression patterns vary by genotype, and thin gray lines denote the probability that an individual with a particular genotype has a particular level of expression. Expression patterns can either be canalized or non-canalized, and overall levels of expression are modified by different genetic backgrounds. A) Expression is highly canalized and penetrance is complete. B) Expression is not canalized and penetrance is complete. C) Expression is not canalized and flies with an *AA* genotype have incomplete penetrance. This occurs because genetic background #2 causes expression levels to straddle the threshold.

Figures

Figure 1. Four landscapes

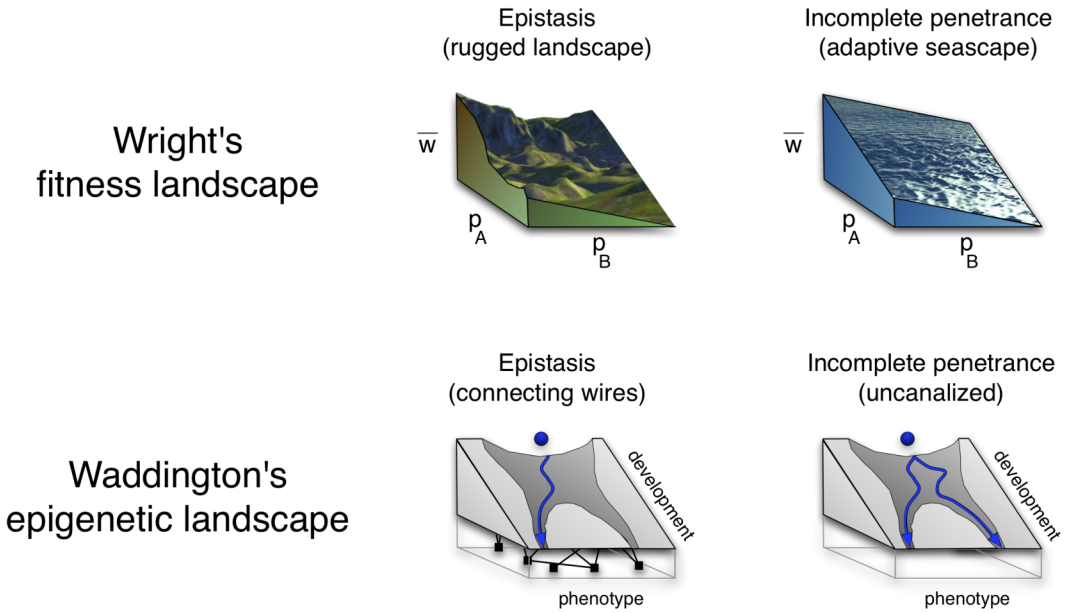
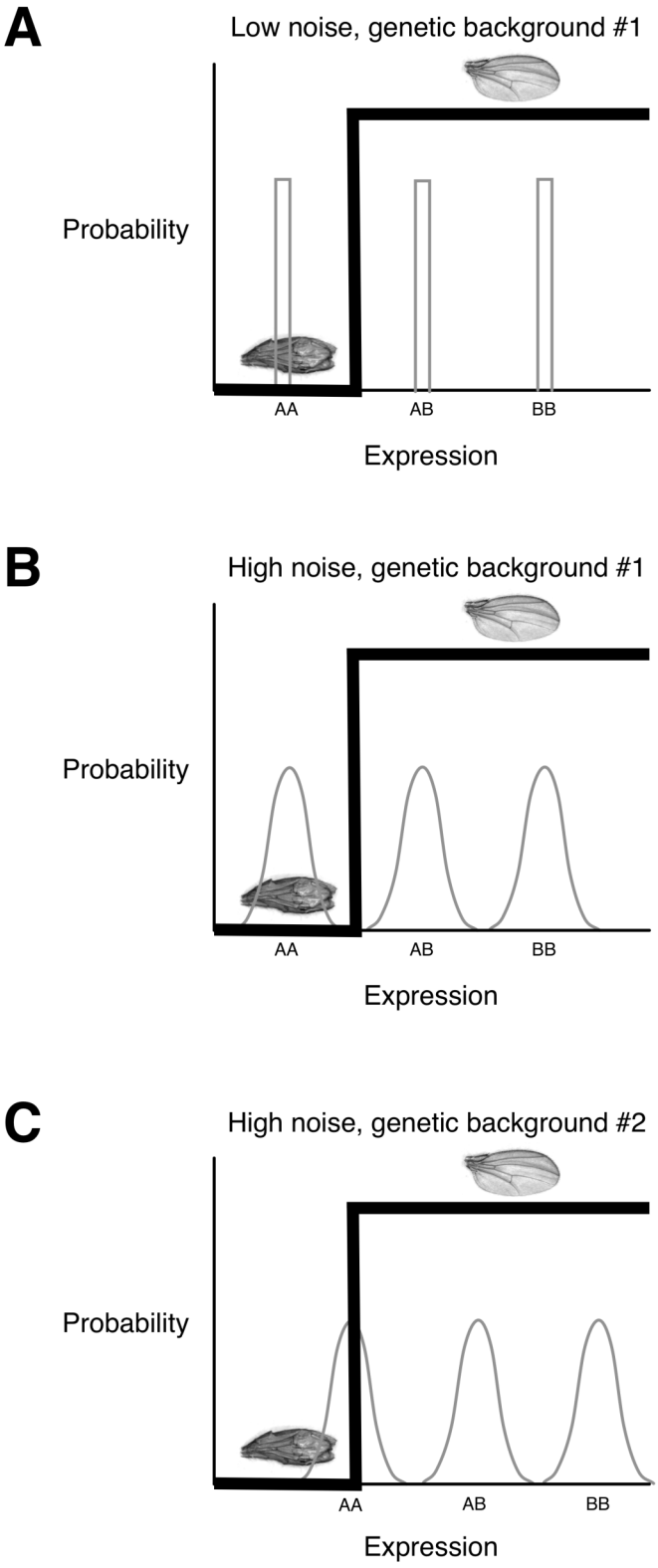


Figure 2. Expression thresholds, genetic backgrounds, and penetrance



Chapter 2

X-autosome incompatibilities in *Drosophila melanogaster*: tests of Haldane's rule and geographic patterns within species

Joseph Lachance and John R. True

Abstract

Substantial genetic variation exists in natural populations of *Drosophila melanogaster*. This segregating variation includes alleles at different loci that interact to cause lethality or sterility (synthetic incompatibilities). Fitness epistasis in natural populations has important implications for speciation and the rate of adaptive evolution. To assess the prevalence of epistatic fitness interactions, we placed naturally occurring X chromosomes into genetic backgrounds derived from different geographic locations. Considerable amounts of synthetic incompatibilities were observed between X chromosomes and autosomes: greater than 44% of all combinations were either lethal or sterile. Sex-specific lethality and sterility were also tested to determine whether Haldane's rule holds for within-species variation. Surprisingly, we observed an excess of female sterility in genotypes that were homozygous, but not heterozygous, for the X chromosome. The recessive nature of these incompatibilities is similar to that predicted for incompatibilities underlying Haldane's rule. Our study also found higher levels of sterility and lethality for genomes that contain chromosomes from different geographical regions. These findings are consistent with the view that genomes are co-adapted gene complexes and that geography affects the likelihood of epistatic fitness interactions.

Introduction

Genes act in a genomic context. Their epistatic interactions affect the evolution of natural populations (Wolf et al. 2000). Many different types of epistasis exist (Phillips 2008), but one unifying theme is that the effects of genes often depend upon genetic background. Some notable examples of epistatic interactions include olfactory behavior in *Drosophila melanogaster* (Anholt et al. 2003), plant growth in *Arabidopsis thaliana* (Alcazar et al. 2009), and blood and bone traits in *Mus musculus* and *Rattus rattus* (Shao et al. 2008). Regardless of the phenotype affected, epistasis must modify fitness to be evolutionarily relevant. Fitness epistasis lies at the core of evolutionary genetic dynamics, influencing both the rate of adaptation (Griswold and Masel 2009; Kim 2007; Sanjuan et al. 2005; Yukilevich et al. 2008) and the genetic architecture of speciation (Johnson 2000; Presgraves 2007).

Synthetic incompatibilities are an important type of fitness epistasis. These interactions take place between alleles at different loci causing fitness to be reduced when both alleles are present in the same individual. Synthetic incompatibilities can cause either lethality or sterility. Population genetics theory suggests that synthetic lethal alleles can segregate at relatively high frequencies at mutation-selection balance (Phillips and Johnson 1998). This is because single mutations are masked and only individuals containing multiple mutations are exposed to selection. The mechanistic basis of specific synthetic incompatibilities can be conserved over evolutionary time: >20% of gene combinations that lead to incompatibilities in *Saccharomyces pombe* also lead to incompatibilities in *Saccharomyces cerevisiae* (Dixon et al. 2008). Synthetic lethal alleles have been found segregating in common haplotypes of *Caenorhabditis elegans* (Seidel *et al*, 2008) and in populations of the copepod *Tigriopus californicus* (Harrison and

Edmands 2006). Numerous studies of natural variation in the genus *Drosophila* have detected synthetic incompatibilities, particularly those involving interactions between alleles on different chromosomes (Krimbas 1960; Lucchesi 1968; Powell 1997; Temin et al. 1969; Thompson 1986). Most of the chromosome pairs tested in these studies did not result in large numbers of synthetic incompatibilities (on the order of zero to 5%). However, most of these studies only examined synthetic lethality. Much less is known about the synthetic sterility of naturally segregating alleles, especially in females.

Because synthetic incompatibilities can form the basis of Dobzhansky-Bateson-Muller (DBM) incompatibilities, they are relevant to speciation. DBM incompatibilities result in postzygotic reproductive isolation without either diverging population having to cross a fitness valley, and introgression studies indicate that many small genomic regions from one species cause DBM incompatibilities when placed in the genetic background of a sister species (Masly and Presgraves 2007; Moyle and Nakazato 2009; Palopoli and Wu 1994; Tao et al. 2003; True et al. 1996). Another example involves the *Lethal hybrid rescue* gene in *D. simulans* and the *Hybrid male rescue* gene in *D. melanogaster*. These genes interact to cause lethality in hybrid F1 males (Brideau et al. 2006). Within-species DBM incompatibilities have also been observed in the autoimmune response of *Arabidopsis thaliana* (Bomblies et al. 2007). Some types of synthetic incompatibilities do not result in DBM incompatibilities. For example, synergistic epistasis can occur between two deleterious alleles, resulting in synthetic lethality or sterility. Since each of these alleles is slightly deleterious by itself, this scenario is not an example of a DBM incompatibility.

Theoretical models such as Orr's "snowball effect" suggest that the accumulation of DBM incompatibilities is contingent on the number of divergent epistatic loci (Orr 1995). In this

model, substitutions that occur in each of two divergent populations can potentially result in deleterious interactions. There currently is some debate whether the "snowball effect" best describes the accumulation of reproductive incompatibilities in nature (Gourbiere and Mallet 2009). There is evidence that standing genetic variation can greatly affect time to speciation (Schluter and Conte 2009). It is therefore of great interest to assess whether natural populations harbor large amounts of synthetic lethal and synthetic sterile alleles.

The chromosomes involved in DBM incompatibilities are important to Haldane's rule. This rule states that if only one sex of a hybrid is sterile or inviable, it tends to be the heterogametic sex (Haldane 1922). Four alternative hypotheses that explain Haldane's rule are: faster-X, faster male, meiotic drive, and dominance (Coyne and Orr 2004). The faster-X hypothesis is sensitive to demography and it posits that rates of adaptive change differ between X-linked and autosomal loci (Mank et al. 2010). The faster male hypothesis posits that male traits may evolve faster due to sexual selection and the sensitivity of spermatogenesis to new mutations (Wu and Davis 1993). Divergence of meiotic drive suppression systems can cause hybrid sterility and lethality, and segregation distortion of sex chromosomes can distort sex ratios away from 50:50 (Frank 1991; Henikoff et al. 2001; Hurst and Pomiankowski 1991; Tao et al. 2001). Finally, the dominance hypothesis states that if alleles causing interspecific incompatibilities behave recessively in hybrids, then the heterogametic sex will be more likely to be affected (Turelli and Orr 1995). Data from a wide variety of taxa support the dominance hypothesis (Bierne et al. 2006; Carling and Brumfield 2008; Coyne and Orr 2004), as do theoretical models with a firm grounding in DBM incompatibilities (Turelli and Orr 2000). *Drosophila* have an XY system in which males are the heterogametic sex. Because recessive X-linked alleles cannot be masked in *Drosophila* males, sex-specific patterns can arise when

epistasis involves the X chromosome. While Haldane's rule is known to apply to between-species incompatibilities, it is unknown whether it applies to within-species sterility and lethality. Note, however, that the dominance theory does not make any predictions about the dominance of incompatible alleles within species. Recent work in *Tribolium castaneum* suggests that Haldane's rule may apply to within-species incompatibilities between different populations (Demuth and Wade 2007). However, sterility in other heterogametic-male species (such as *Homo sapiens*) seems to be female-biased (Thonneau et al. 1991).

An additional consideration is that genomes can be viewed as locally co-adapted gene complexes (Dobzhansky 1970). If this is the case, then we would expect genes derived from different geographic regions to be more likely to exhibit synthetic incompatibilities. Natural selection may be unable to remove deleterious allelic combinations if populations are spatially structured. This is consistent with the phenomenon of outbreeding depression (Edmands 2007). *Drosophila melanogaster* is a human commensal that has a worldwide distribution, having recently expanded out of Africa (Keller 2007). Clines exist for multiple traits and isolation by distance has led to genetic differentiation among populations of *D. melanogaster* (Pool and Aquadro 2006; Sezgin et al. 2004; Umina et al. 2005). In addition, flies derived from Zimbabwe are sexually isolated from cosmopolitan populations (Wu et al. 1995). There is also evidence of phenotypic differentiation and partial prezygotic reproductive isolation between United States and Caribbean populations of *D. melanogaster* (Yukilevich and True 2008a; Yukilevich and True 2008b). While previous studies detected only low levels of synthetic lethality in *D. melanogaster*, these studies did not test chromosomes derived from multiple geographic locations (Thompson 1986).

In this study we investigated synthetic incompatibilities in *D. melanogaster* on a chromosomal scale. Extracted-X lines were constructed by placing naturally segregating X chromosomes into multiple genetic backgrounds. Levels and types of synthetic incompatibilities were assessed (including lethality vs. sterility, recessive vs. dominant, and male vs. female incompatibilities). This allowed us to address the following questions: 1) How common are synthetic incompatibilities among naturally segregating chromosomes? 2) Does Haldane's rule hold for within-species synthetic lethality and sterility? 3) Are synthetic incompatibilities more common when chromosomes are derived from different geographic locations?

Materials and Methods

Stocks and construction of lines

We constructed two sets of extracted-X lines using balancer chromosomes. Set 1 contains 118 different X chromosomes bred into a single autosomal background. Set 2 contains 52 different X chromosomes bred into three different autosomal backgrounds. Two of the lines in Set 2 were unable to be maintained, resulting in a total of 154 lines. X chromosomes tested in Set 2 overlap with X chromosomes tested in Set 1. Extracted-X lines were constructed via three stages (Figure 1). First, wild-caught isofemale lines were isogenized (either via balancer stocks or ten generations of sib-mating). This resulted in lines with genotypes of $+^{x}$ or $+^{a}$ (superscripts indicate whether a line serves as a source of extracted-X chromosomes or autosomes). Secondly, intermediate stocks containing markers and either an extracted X chromosome or an autosomal background were constructed. These crosses involved w^{1118};*T(2;3)ap*^{Xa}/*CyO:TM3* flies (Bloomington stock 2475), and resulted in $+^{x}$;*T(2;3)ap*^{Xa}/*CyO:TM3* and w^{1118};$+^{a}$;$+^{a}$ lines. Third,

homozygous extracted-X lines were constructed. The final cross of this scheme involved crossing $+^{x};+^{a}/CyO;+^{a}/TM3$ flies and selecting wild-type offspring. Each of the resulting extracted-X lines has a genotype of $+^{x};+^{a};+^{a}$. The fourth chromosome (~2% of the genome) was not monitored.

X chromosomes used in these lines were derived from wild-caught isofemale lines, while autosomal backgrounds were derived from both wild caught and laboratory lines. Geographic origins of X chromosomes were: Sudbury, Ontario, Canada and Long Island, New York (collected by T. Merritt in 2005), Southern United States and the Caribbean (collected by R. Yukilevich in 2004-5), and Cameroon and Zimbabwe (collected by J. Pool and C. Aquadro in 1990, 1994 and 2004). See supplemental information for a full list of X chromosomes. The autosomal background used in Set 1 was derived from a mapping stock (Bloomington stock 6326). The three autosomal backgrounds used in Set 2 were: 6326, Sudbury (latitude: 46.49, longitude: -81.01), and Rum Cay, Bahamas (latitude: 23.38, longitude: -74.5). Note that Sudbury, Ontario is near the northern range limit of *D. melanogaster*, and Rum Cay is a southern location in the Bahamas. By contrast, the geographic origins of 6326 are unknown (R. Hoskins and A. Phan, personal communication). Each source line (X and autosomal) could be maintained as an isofemale line, indicating the absence of single-locus lethality or sterility. The 6326 and Sudbury lines used to provide autosomal backgrounds are isogenic, whereas the Rum Cay line was produced by ten generations of sib-mating.

Flies were cultured on standard corn meal/molasses/agar medium supplemented with antibiotics (either penicillin at 40 μg/ml or a mix of tetracycline and streptomycin at 63 μg/ml and 19 μg/ml, respectively). All crosses were performed at 25°C with a 12 hour light:dark cycle.

Lethality assays

The crossing scheme used to generate extracted-X lines allowed different types of synthetic lethality to be distinguished. In particular, when we were unable to construct homozygous lines the stage at which crosses were unable to proceed was recorded. This allowed us to determine the dominance of the fitness interactions, which chromosomes were involved, and whether synthetic lethality was male or female-specific (see bottom part of Figure 1). Set 1 of the extracted-X lines was tested for recessive synthetic lethality, and Set 2 was tested for dominant and recessive synthetic lethality. Dominant male lethality was assessed by crossing $+^{x}$;*T(2;3)ap*^{Xa}/*CyO:TM3* females and *w*^{1118};$+^{a}$;$+^{a}$ males. When this cross did not result in male offspring, dominant X-autosome interactions were implicated. Dominant female lethality was assessed by crossing $+^{x}$;*T(2;3)ap*^{Xa}/*CyO:TM3* females and $+^{x}$;$+^{a}$;$+^{a}$ males. When this cross did not result in female offspring with curly wings and stubble bristles (i.e. flies with a $+^{x}$;$+^{a}$/*CyO*;$+^{a}$/*TM3* genotype), dominant X-autosome interactions were implicated. Note that dominance and recessivity in this case refers to the number of autosomal copies required for synthetic lethality.

Recessive lethal interactions were assessed by intercrossing $+^{x}$;$+^{a}$/*CyO*;$+^{a}$/*TM3* flies. Wild-type offspring from this cross can only appear in the absence of recessive lethal interactions. Recessive X-2nd lethal interactions cause all offspring from this cross to have curly wings (i.e. flies with $+^{x}$;$+^{a}$/*CyO*;$+^{a}$/$+^{a}$ or $+^{x}$;$+^{a}$/*CyO*;$+^{a}$/*TM3* genotypes). Recessive X-3rd lethal interactions cause all offspring from this cross to have stubble bristles (i.e. flies with $+^{x}$;$+^{a}$/$+^{a}$;$+^{a}$/*TM3* or $+^{x}$;$+^{a}$/*CyO*;$+^{a}$/*TM3* genotypes). Each X-autosome combination was replicated twice, and at least 60 offspring were genotyped for each combination of X chromosome and autosomal background.

Sterility assays

X-autosome synthetic sterility was detected by crossing synthetic genotypes to wild-type flies and looking for the presence of offspring. For each X-autosome combination, flies were mass mated with other flies belonging to the same extracted-X line. Newly emerged flies were aged three to five days prior to each cross. Three virgin $+^{x};+^{a};+^{a}$ females were then placed into a vial with three $+^{x};+^{a};+^{a}$ males. After seven days, adults were removed and vials were inspected for developing offspring. If larvae or pupae were observed, genotypes were considered to be fertile. Each of these crosses was replicated three times. Note that Y chromosomes in these sterility assays were derived from $+^{x}$;*T(2;3)ap*^{Xa}/*CyO:TM3* balancer stocks.

The absence of viable offspring can be due to either male or female sterility. To determine if sterility was male or female-specific, extracted-X flies were outcrossed to SBU1, a wild-caught isofemale line from Stony Brook, NY (collected by J. R. True in 2005). Three males and females from each extracted-X line were mated with SBU1 flies of the opposite sex. Each of these crosses was replicated three times. If no offspring resulted from crossing female SBU1 flies with extracted-X males ($+^{x};+^{a};+^{a}$), the line was considered male sterile. If no offspring resulted from crossing extracted-X females ($+^{x};+^{a};+^{a}$) with SBU1 males, the line was considered female sterile. This scheme also allowed us to infer whether sterility was an organismal property or the property of a pair of mating individuals. With organismal sterility, flies of a particular sex and genotype are unable to produce offspring regardless of the genotype of their mating partner. With mating pair sterility, flies are only sterile when they are paired with flies of a particular genotype. If a particular genotype was unable to produce offspring when

mated with either its own genotype or SBU1 flies, then the synthetic incompatibility was classified as organismal sterility.

Sperm motility was assayed for lines that contained males that were sterile when crossed to females of two different genotypes (organismal male sterile lines). After developing at 25°C, newly emerging males were separated by genotype and aged 4-5 days without access to females. Testes of individual males were dissected in a drop of Ringer's solution, gently squashed under a coverslip, and examined under a stereomicroscope. If a single motile sperm was observed, males were classified as possessing motile sperm.

Dominance tests for female sterility

Because the lines used in the above sterility tests were homozygous for the X chromosome, it was unclear whether female sterility was due to dominant or recessive X-autosome interactions (i.e. is a single copy of an X chromosome sufficient to confer synthetic sterility?). To assess this, we controlled for genetic background and generated females heterozygous for a putatively sterile X chromosome. $+^{x1};+^{a}/CyO;+^{a}/TM3$ females were crossed with $+^{x2};+^{a}/CyO;+^{a}/TM3$ males (and vice versa) to generate $+^{x1}/+^{x2};+^{a};+^{a}$ females (where $+^{x1}$ and $+^{x2}$ are two different X chromosomes). As per the above sterility assays, the resulting wild-type females were mass mated with $+^{x1};+^{a};+^{a}$, $+^{x2};+^{a};+^{a}$, and SBU1 males. Each putatively sterile X chromosome was tested with four other X chromosomes (two putatively sterile and two fertile X chromosomes). If all four combinations failed to generate offspring, female-specific synthetic sterility was classified as dominant. Otherwise, synthetic sterility was classified as (partially) recessive. A total of 14 different synthetic sterile X chromosomes were tested, and each cross

was replicated twice. Note that these crosses are also complementation tests for X-linked sterility factors.

Sex ratio assays

Sex ratio data were recorded for 49 extracted-X lines from Set 1. Each of these lines was able to be maintained as a homozygous stock. For each X-autosome combination, three replicate vials were checked. Numbers of newly emerging males and females were recorded five different days for each vial. Emerging flies were counted no later than 17 days after the original cross was set up. Samples sizes were too small for 13 of the lines (less than 30 flies emerged), leaving a total of 36 lines. For each of the 36 remaining lines, a mean number of 127 flies were counted.

Non-parametric test of geographical patterns

A non-parametric test was used to determine whether synthetic incompatibilities were more likely when X chromosomes and autosomes were derived from different geographical locations. First, X chromosomes were classified as either northern or southern. Northern X chromosomes were derived from populations found at latitudes above 40 (approximately the Mason-Dixon line, see Figure 5A), and southern X chromosomes were derived from populations found at latitudes below 40. Under this formulation, Set 2 contains 18 northern X lines and 34 southern X lines. Note that the set of southern X chromosomes contains some lines with African X chromosomes. The Sudbury autosomal background has a northern origin, and the Rum Cay autosomal background has a southern origin. The following test statistic (γ) was calculated:

$$\gamma = (x_{NN} - x_{SN}) + (x_{SS} - x_{NS}). \quad (1)$$

In this equation x refers to the proportion of X chromosomes that are incompatible with a particular autosomal background. Subscripts indicate the geographic origins of the X and autosomal chromosomes, with X chromosomal origin listed first. For example, x_{SN} refers to the proportion of southern X chromosomes that are incompatible with a northern autosomal background. γ is negative if synthetic incompatibilities occur more often when X chromosomes and autosomes are mismatched (i.e. chromosomes are derived from different regions).

Given a null hypothesis that synthetic incompatibilities are independent of the geographic origin of chromosomes, Monte-Carlo simulations were run to determine the distribution of γ. The probability of synthetic incompatibility varies by autosomal background (38.5% of tested X chromosomes were incompatible with a northern autosomal background, and 73.1% were incompatible with a southern autosomal background). Using these probabilities, simulated datasets of 52 X chromosomes (18 northern and 34 southern) were generated for each autosomal background. Thus, the number of incompatible X-autosome combinations for each of the four possibilities (north-north, south-north, south-south, and north-south) follows a binomial distribution, and null expectations are equal proportions of incompatible northern and southern X chromosomes. γ was calculated for each simulation run. Monte Carlo simulations were run 100,000 times and the distribution of the test statistic γ was compared to the observed data.

Results

Overall levels of synthetic incompatibility

Substantial levels of synthetic incompatibilities were observed for both sets of extracted-X lines (Table 1). In Set 1, 43 of 118 X chromosomes (36.4%) resulted in synthetic sterility or lethality when placed on a 6326 genetic background. Similarly, 50.0% of the X-autosome combinations tested in Set 2 resulted in synthetic sterility or lethality. The numbers of synthetic incompatibilities varied by genetic background in Set 2: 20 of 51 X chromosomes (39.2%) were incompatible with a 6326 background, 19 of 52 X chromosomes (36.5 %) were incompatible with a Sudbury background, and 38 of 51 X chromosomes (74.5%) were incompatible with a Rum Cay background. The excess proportion of deleterious interactions on a Rum Cay background relative to other backgrounds was highly significant (two-tailed p-value < 0.001 for each comparison; Fisher's exact test). Note that two lines were unable to be tested in this second data set, and levels of incompatibilities on a 6326 background were similar for Set 1 and Set 2. The proportion of incompatibilities that involved lethality or sterility also varied by genetic background. A larger percentage of incompatible X-autosome combinations involving the 6326 background were sterile (as opposed to lethal) compared to Sudbury and Rum Cay backgrounds (35.0% vs. 15.0% and 13.2%, single-tailed p-values of 0.137 and 0.056 for each comparison; Fisher's exact test). As the presence of only a few larvae or pupae was sufficient to classify a line as fertile, many of the X-autosome combinations labeled as lacking synthetic incompatibilities actually exhibited semi-sterility and were difficult to maintain as a homozygous stocks. This detail was corroborated by the observation that 13 of the 49 lines assayed in the sex ratio experiment had insufficient statistical power due to low numbers of offspring. The relatively low level of synthetic lethality observed in Set 1 as opposed to Set 2 is due to the fact that Set 1 lines were only tested for recessive lethal interactions. In addition, Set 1 involved X

chromosomes over a 6326 background (the background that was more likely to result in synthetic sterility). The net result of the data in Table 1 was that many otherwise viable and fertile X chromosomes interacted deleteriously with novel autosomal backgrounds.

Does knowledge that a particular X-autosome combination is incompatible tell us anything about the lethality or sterility of another X-autosome combination? Set 2 of the extracted X lines was used for a test of independence, as Set 1 involved only a single autosomal background. Inspection of Figure 2 suggests the absence of any pattern: X chromosomes that were incompatible with a 6326 background (alternatively Sudbury or Rum Cay) were not any more or less likely to be incompatible with another autosomal background. Indeed, independence of the incompatibilities found in each background could not be rejected when the data in Figure 2 were converted into a 2x2x2 contingency table and a log-linear model was tested (p-value = 0.2742, $\chi^2 = 5.13$, d.f. = 4). Seven X chromosomes were incompatible with all three genetic backgrounds and eight X chromosomes were compatible with every tested genetic background. However, this was consistent with what one would expect from multiplying background-specific probabilities of synthetic lethality and sterility.

Chromosome-specific lethality patterns

A majority of X-autosome interactions were recessive (requiring homozygous autosomes) and both X-2nd and X-3rd interactions were observed. Of the 154 lines tested in Set 2, 92 did not exhibit any synthetic lethality (Figure 3). Of the remaining synthetic lethal lines, 22 (35.5%) exhibited dominant synthetic lethality and 40 (64.5%) exhibited recessive lethality. Here, dominance and recessivity refers to the number of autosomal copies required for synthetic lethality (all lines tested were homozygous for an extracted X chromosome). One caveat is that

our methodology may overestimate the frequency of synthetic lethals. This is because Mendelian segregation alone can cause a genotype to be absent (even if 60 flies were assayed per cross). Three times as many X-autosome combinations involving the 2nd chromosome resulted in recessive synthetic lethality relative to combinations involving the 3rd chromosome (two-tailed p-value < 0.001; Fisher's exact test). While the *Drosophila melanogaster* second chromosome is slightly larger than the third chromosome (and thus a larger target for epistatic interactions), this alone is an insufficient explanation for the observed differences in X-2nd vs. X-3rd synthetic lethality. Finally, an appreciable number of lines (16) exhibited both X-2nd and X-3rd interactions, suggesting that complex epistasis may underlie synthetic lethality in these lines.

Sterility tests

Sex-specific patterns of synthetic incompatibilities were observed. These patterns involved a very slight excess of male lethality over female lethality and a greater than three-fold excess of female sterility over male sterility (Table 1). Comparisons between female-specific and male-specific sterility revealed a statistically significant difference for Set 1, but not Set 2 (one-tailed p-value = 0.0086 for Set 1, one-tailed p-value = 0.3112 for Set 2, Fisher's exact test). The excess of female sterility was less striking when sex-specific sterility and both-sexes-sterile data were pooled (one-tailed p-value = 0.0891 for Set 1, one-tailed p-value = 0.4190 for Set 2; Fisher's exact test). Regardless of statistical significance, these patterns would not be expected if Haldane's rule holds within species. While slight differences were observed for different autosomal backgrounds (greater levels of female sterility for lines containing a 6326 background), sample sizes were too small to say anything definitive about background-specific female sterility.

There was evidence for both organismal and mating pair specific sterility. Of 33 synthetic sterile lines tested in Set 1, 11 were unable produce offspring when mated with flies containing either the same combination of X and autosomes or SBU1 flies (i.e. sterility was an organismal property). The other 22 lines were able to produce offspring when mated with one, but not the other, genotype (i.e. sterility was a property of a mating pair). Of 15 synthetic sterile lines tested in Set 2, six exhibited organismal sterility and nine exhibited mating pair sterility. Note that the proportion of synthetic sterile lines exhibiting organismal sterility might be an overestimate. This is because flies were only tested against two genotypes of the opposite sex, and we cannot formally rule out the possibility that they might be able to produce viable offspring when mated with flies that contain a third, untested genotype. Males that were unable to sire offspring with females of multiple genotypes were assayed for sperm motility. Motile sperm were not observed for any of the six lines tested, suggesting that these incidences of male-specific organismal sterility involved defective spermatogenesis.

Tests of recessivity of female-sterile lines

Lines that exhibited female sterility were tested for recessivity of X effects. Hybrid flies used in these tests were homozygous for their respective autosomal backgrounds. When hybrid females containing a single putatively sterile X chromosome over a fertile X chromosome were tested for sterility, viable offspring resulted for all 14 of the lines tested. This indicates that X-autosome interactions causing female sterility in our study required flies to be homozygous for the same X chromosome (i.e. they were recessive). When hybrid females containing X chromosomes derived from two different female-sterile lines were tested for sterility, 13 of 14 cases resulted in viable offspring. The exception involved a line with an X chromosome derived

from Cameroon, Africa over a 6326 background (MD 16). The MD 16 X chromosome was incompatible with Sud 24 (Sudbury, Ontario) and 18 26 (Montgomery, Alabama) X chromosomes. As this line was able to produce viable offspring when mated with fertile lines, there was still evidence that MD 16 contains a recessive X-linked female sterility factor. The failure of MD 16 to complement two other female sterile X chromosomes suggests that the same loci may be implicated in multiple cases of synthetic sterility.

Skewed sex ratios

Additional evidence of sex-specific effects was observed in the form of skewed sex ratios. While 1:1 sex ratios were expected, we observed an excess of males for a number of extracted-X lines (Figure 4). Six of 36 lines tested from Set 1 had a significant excess of males (p-value < 0.05 after correcting for multiple tests using the Benjamini and Hochberg false discovery rate). These male-biased sex ratios are inconsistent with Haldane's rule expectations. Lines were successfully maintained and sex ratios were reasonably stable over time, suggesting that Y-linked meiotic drive was not a cause of the observed patterns. As we observed an excess of males (as opposed to females) and our media contained antibiotics, *Wolbachia* can also be ruled out as a cause of unequal sex ratios. It is possible that female inviability could explain the male-biased sex ratios, but this was not assayed in our sex ratio tests.

Geographic patterns

Interesting patterns arose when the geographic origins of chromosomes were considered (Figure 5). A general trend was that southern X chromosomes were more likely to result in synthetic incompatibilities than northern X chromosomes (two-tailed p-value = 0.013 for Set 1,

two-tailed p-value = 0.259 for pooled data from Set 2; Fisher's exact test). In addition, synthetic incompatibilities were much more common when X chromosomes were placed into the southern autosomal background (Rum Cay). More importantly, levels of synthetic incompatibilities depended on the *combination* of X chromosomes and autosomes. Southern X chromosomes were two and a half times more likely than northern X chromosomes to result in synthetic lethality or sterility when placed on a 6326 autosomal background (Set 2, single-tailed p-value = 0.081; Fisher's exact test). When X chromosomes were placed into a northern autosomal background, southern X chromosomes were more likely to result in synthetic incompatibilities than northern X chromosomes (44.1% vs. 27.8%, single-tailed p-value = 0.198; Fisher's exact test). When X chromosomes were placed into a southern autosomal background, northern X chromosomes were more likely to result in synthetic incompatibilities than southern X chromosomes (83.3% vs. 67.6%, single-tailed p-value = 0.190; Fisher's exact test). While each of these geographical trends was not significant by itself, the data were what one would expect if local populations contain coadapted gene complexes. Sudbury and Rum Cay background data were combined to calculate the test statistic described in Equation 1, yielding $\gamma = -0.320$. This value of γ was statistically significant (p-value = 0.0445; only 4.45% of all Monte Carlo runs resulted in $\gamma < -0.320$). When African X chromosomes were omitted from Set 2, $\gamma = -0.337$ (p-value = 0.0397). Correlations between synthetic incompatibility and the great-circle distance between the geographic origin of X chromosomes and autosomes were weakly positive ($\rho =$ 0.1461 for the Sudbury autosomal background and $\rho = 0.0723$ for the Rum Cay autosomal background). When African X chromosomes were omitted from Set 2, the correlations were slightly different (0.1366 for Sudbury and 0.1217 for Rum Cay). X-autosome combinations

were more likely to result in incompatibilities when chromosomes were derived from different geographical regions.

Discussion

We observed substantial levels of synthetic incompatibility. These findings are consistent with theoretical models that predict recessive synthetic incompatibilities segregating at moderately high frequencies (Phillips and Johnson 1998; Proulx and Phillips 2005). Because each of the X chromosomes and autosomal backgrounds tested can be maintained indefinitely as an isofemale line, we were able to infer that observed incompatibilities are due to epistatic interactions involving X chromosomes and autosomes rather than single locus effects. Note that cytonuclear incompatibilities could be ruled out by our crossing scheme. The independence of X-autosome combinations and the failure of most female sterile X chromosomes to complement each other suggest that multiple genetic factors may be involved. Assuming that the incompatibilities observed in our study are the "stuff" of speciation, this perspective is consistent with the view that speciation can be due to genes at many different loci (Wu and Ting 2004). In addition, our data support past findings that the genetic basis of hybrid incompatibility is complex even at early stages of divergence (Good et al. 2008).

There are three likely reasons why our study detected much higher levels of synthetic incompatibilities than classic studies (Powell 1997; Thompson 1986). First, we assayed levels of synthetic sterility and lethality, as opposed to just synthetic lethality. Second, chromosomes assayed in our study differed in their geographic origins. Third, larger regions of the genome were tested in our study (instead of detecting incompatibilities between a single pair of

chromosomes we were able to detect incompatibilities between X chromosomes and either autosome). Many of the lines constructed in this study are double handicapped: they contain homozygous variants for each chromosome (inbreeding depression) *and* X-autosome combinations from different geographic locations (outbreeding depression).

We did not observe increased levels of synthetic incompatibilities for the heterogametic sex, suggesting that Haldane's rule may not apply for within-species variation of our study species. However, when we examined the genetic basis of female sterility we found evidence consistent with the dominance hypothesis of Haldane's rule. This is because female-specific synthetic sterility involved recessive alleles in each case. As there is evidence of female-biased expression patterns on the X chromosome (Ranz et al. 2003), our findings are also consistent with the faster-X hypothesis. Had the faster male hypothesis held, we would have expected to observe greater levels of male sterility. Note that while our study reveals the recessivity of standing epistatic fitness variation, it does not directly explain what causes Haldane's rule.

What can cause within species patterns of sterility and lethality to differ from between species patterns? One key difference between these two situations is that natural selection is able to eliminate deleterious combinations within species, but it is unable to eliminate deleterious combinations between species (barring reinforcement). Because X-linked sterility factors are unable to be masked in the heterogametic sex, natural selection is more effective at eliminating X-linked male sterility factors than recessive X-linked female sterility factors (Vicoso and Charlesworth 2006). Thus, within a single species the frequencies of naturally segregating X-linked alleles are likely to be greater for female synthetic sterile alleles than male synthetic sterile alleles. By contrast, sex-linked DBM incompatibilities between species have not had the chance to be filtered by natural selection. Note that a comprehensive mathematical treatment of

synthetic incompatibilities at mutation-selection equilibrium only exists for autosomal loci at present (Phillips and Johnson 1998).

Another possible explanation for the absence of Haldane's rule in our study is that we assayed homozygous flies. These are genotypes that can occur in the F2 generation of hybridizing populations. However, the original formulation and subsequent discussion of Haldane's rule has largely focused on the F1 generation (Laurie 1997; Wu et al. 1996). It is unknown whether sex-biased patterns of synthetic incompatibilities are expected to differ between the F1 and F2 generations. For Set 2, 7.8% of X-autosome combinations resulted in male sterility and 9.1% of X-autosome combinations resulted in female sterility (Table 1, pooled sex-specific and both-sterile data). In natural populations this would actually result in higher frequencies of hemizygous males compared to homozygous females (assuming Hardy-Weinberg proportions).

Our data contained both sex-specific lethality and sex-specific sterility. While it is likely that genes affecting viability will have the same effect in both sexes, different physiological processes underlie female and male sterility. The *D. melanogaster* X chromosome is enriched for genes with female-biased expression and deficient for genes with male-biased expression (Ranz et al. 2003). Misregulation of X-linked genes (due to trans effects from the autosomal background) may affect each sex differently. Female-biased expression patterns of X-linked genes may also explain why we observed greater levels of female sterility for homozygous flies. However, mutation studies suggest that the number of X-linked genes involved in male and female fertility are approximately the same (Kaplan et al. 1970; Watanabe and Lee 1977).

Our data are consistent with the idea that genomes are locally coadapted gene complexes. X chromosomes and autosomes derived from the same geographic region were usually compatible. Conversely, northern X chromosomes were more likely to be incompatible with southern autosomes, and vice versa. In addition, there were weak positive correlations between synthetic incompatibility and the geographic distance between the origin of X chromosomes and autosomes. The existence of these incompatibilities may be due to the demographic history of this species. The spread of *D. melanogaster* into the New World likely arose via two separate routes: a northern route from Africa via Europe and a southern route involving direct immigration from Africa. North America and the Caribbean thus appear to be zones of secondary contact where potentially incompatible alleles can interact. Additional support for this hypothesis comes from the fact that Caribbean populations have phenotypes that are more similar to African than United States populations (Caracristi and Schlotterer 2003; Yukilevich and True 2008a). Synthetic incompatibilities can also be a byproduct of local selection pressures if locally adaptive alleles have pleiotropic effects. Despite the intriguing geographic patterns in our study, generalizations should be taken with caution. This is because only three autosomal backgrounds were tested and it is possible that the observed patterns are due to the specific backgrounds tested rather than geography. More data are needed, as are additional theoretical models of synthetic incompatibility that incorporate spatial population structure. Our findings indicate that levels of synthetic incompatibility may be underestimated if chromosomes from only a single location are tested. Natural populations may already contain the genetic potential for speciation in the form of cryptic DBM incompatibilities.

References

Alcazar, R., A. V. Garcia, J. E. Parker, and M. Reymond. 2009. Incremental steps toward incompatibility revealed by Arabidopsis epistatic interactions modulating salicylic acid pathway activation. Proc Natl Acad Sci U S A 106:334-339.

Anholt, R. R., C. L. Dilda, S. Chang, J. J. Fanara, N. H. Kulkarni, I. Ganguly, S. M. Rollmann et al. 2003. The genetic architecture of odor-guided behavior in Drosophila: epistasis and the transcriptome. Nat Genet 35:180-184.

Bierne, N., F. Bonhomme, P. Boudry, M. Szulkin, and P. David. 2006. Fitness landscapes support the dominance theory of post-zygotic isolation in the mussels Mytilus edulis and M. galloprovincialis. Proc Biol Sci 273:1253-1260.

Bomblies, K., J. Lempe, P. Epple, N. Warthmann, C. Lanz, J. L. Dangl, and D. Weigel. 2007. Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. PLoS Biol 5:e236.

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

Caracristi, G., and C. Schlotterer. 2003. Genetic differentiation between American and European Drosophila melanogaster populations could be attributed to admixture of African alleles. Mol Biol Evol 20:792-799.

Carling, M. D., and R. T. Brumfield. 2008. Haldane's rule in an avian system: using cline theory and divergence population genetics to test for differential introgression of mitochondrial, autosomal, and sex-linked loci across the Passerina bunting hybrid zone. Evolution 62:2600-2615.

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

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

Dixon, S. J., Y. Fedyshyn, J. L. Koh, T. S. Prasad, C. Chahwan, G. Chua, K. Toufighi et al. 2008. Significant conservation of synthetic lethal genetic interaction networks between distantly related eukaryotes. Proc Natl Acad Sci U S A 105:16653-16658.

Dobzhansky, T. 1970, Genetics of the Evolutionary Process. New York, Columbia University Press.

Edmands, S. 2007. Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. Mol Ecol 16:463-475.

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

Good, J. M., M. A. Handel, and M. W. Nachman. 2008. Asymmetry and polymorphism of hybrid male sterility during the early stages of speciation in house mice. Evolution 62:50-65.

Gourbiere, S., and J. Mallet. 2009. Are Species Real? The Shape of the Species Boundary with Exponential Failure, Reinforcement, and the 'Missing Snowball'. Evolution.

Griswold, C. K., and J. Masel. 2009. Complex adaptations can drive the evolution of the capacitor [PSI], even with realistic rates of yeast sex. PLoS Genet 5:e1000517.

Haldane, J. B. S. 1922. Sex ratio and unisexual sterility in hybrid animals. J. Genet. 12:101-109.

Harrison, J. S., and S. Edmands. 2006. Chromosomal basis of viability differences in Tigriopus californicus interpopulation hybrids. J Evol Biol 19:2040-2051.

Henikoff, S., K. Ahmad, and H. S. Malik. 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. Science 293:1098-1102.

Hurst, L. D., and A. Pomiankowski. 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.

Johnson, N. 2000. Gene Interactions and the Origin of Species, Pages 197-212 *in* J. Wolf, E. Brodie, and M. Wade, eds. Epistasis and the Evolutionary Process. New York, Oxford University Press.

Kaplan, W. D., R. L. Seecof, W. E. Trout, and M. E. Pasternack. 1970. Production and Relative Frequency of Maternally Influenced Lethals in Drosophila melanogaster. American Naturalist 104:261-271.

Keller, A. 2007. Drosophila melanogaster's history as a human commensal. Curr Biol 17:R77-81.

Kim, Y. 2007. Rate of adaptive peak shifts with partial genetic robustness. Evolution 61:1847-1856.

Krimbas, C. B. 1960. Synthetic Sterility in Drosophila Willistoni. Proc Natl Acad Sci U S A 46:832-833.

Laurie, C. C. 1997. The weaker sex is heterogametic: 75 years of Haldane's rule. Genetics 147:937-951.

Lucchesi, J. C. 1968. Synthetic lethality and semi-lethality among functionally related mutants of Drosophila melanfgaster. Genetics 59:37-44.

Mank, J. E., B. Vicoso, S. Berlin, and B. Charlesworth. 2010. Effective population size and the faster-X effect: empirical results and their interpretation. Evolution 64:663-674.

Masly, J. P., and D. C. Presgraves. 2007. High-resolution genome-wide dissection of the two rules of speciation in Drosophila. PLoS Biol 5:e243.

Moyle, L. C., and T. Nakazato. 2009. Complex epistasis for Dobzhansky-Muller hybrid incompatibility in solanum. Genetics 181:347-351.

Orr, H. A. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics 139:1805-1813.

Palopoli, M. F., and C. I. Wu. 1994. Genetics of hybrid male sterility between drosophila sibling species: a complex web of epistasis is revealed in interspecific studies. Genetics 138:329-341.

Phillips, P. C. 2008. Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet 9:855-867.

Phillips, P. C., and N. A. Johnson. 1998. The population genetics of synthetic lethals. Genetics 150:449-458.

Pool, J. E., and C. F. Aquadro. 2006. History and structure of sub-Saharan populations of Drosophila melanogaster. Genetics 174:915-929.

Powell, J. 1997, Progress and Prospects in Evolutionary Biology: The Drosophila Model: Oxford Series in Ecology and Evolution. New York, Oxford University Press.

Presgraves, D. C. 2007. Speciation genetics: epistasis, conflict and the origin of species. Curr Biol 17:R125-127.

Proulx, S. R., and P. C. Phillips. 2005. The Opportunity for Canalization and the Evolution of Genetic Networks. American Naturalist 165:147-162.

Ranz, J. M., C. I. Castillo-Davis, C. D. Meiklejohn, and D. L. Hartl. 2003. Sex-dependent gene expression and evolution of the Drosophila transcriptome. Science 300:1742-1745.

Sanjuan, R., J. M. Cuevas, A. Moya, and S. F. Elena. 2005. Epistasis and the adaptability of an RNA virus. Genetics 170:1001-1008.

Schluter, D., and G. L. Conte. 2009. Genetics and ecological speciation. Proc Natl Acad Sci U S A 106 Suppl 1:9955-9962.

Sezgin, E., D. D. Duvernell, L. M. Matzkin, Y. Duan, C. T. Zhu, B. C. Verrelli, and W. F. Eanes. 2004. Single-locus latitudinal clines and their relationship to temperate adaptation in metabolic genes and derived alleles in Drosophila melanogaster. Genetics 168:923-931.

Shao, H., L. C. Burrage, D. S. Sinasac, A. E. Hill, S. R. Ernest, W. O'Brien, H. W. Courtland et al. 2008. Genetic architecture of complex traits: large phenotypic effects and pervasive epistasis. Proc Natl Acad Sci U S A 105:19910-19914.

Tao, Y., S. Chen, D. L. Hartl, and C. C. Laurie. 2003. Genetic dissection of hybrid incompatibilities between Drosophila simulans and D. mauritiana. I. Differential accumulation of hybrid male sterility effects on the X and autosomes. Genetics 164:1383-1397.

Tao, Y., D. L. Hartl, and C. C. Laurie. 2001. Sex-ratio segregation distortion associated with reproductive isolation in Drosophila. Proc Natl Acad Sci U S A 98:13183-13188.

Temin, R. G., H. U. Meyer, P. S. Dawson, and J. F. Crow. 1969. The influence of epistasis on homozygous viability depression in Drosophila melanogaster. Genetics 61:497-519.

Thompson, V. 1986. Synthetic lethals: a critical review. Evolutionary Theory 8:1-13.

Thonneau, P., S. Marchand, A. Tallec, M. L. Ferial, B. Ducot, J. Lansac, P. Lopes et al. 1991. Incidence and main causes of infertility in a resident population (1,850,000) of three French regions (1988-1989). Hum Reprod 6:811-816.

True, J. R., B. S. Weir, and C. C. Laurie. 1996. A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of Drosophila mauritiana chromosomes into Drosophila simulans. Genetics 142:819-837.

Turelli, M., and H. A. Orr. 1995. The dominance theory of Haldane's rule. Genetics 140:389-402.

—. 2000. Dominance, epistasis and the genetics of postzygotic isolation. Genetics 154:1663-1679.

Umina, P. A., A. R. Weeks, M. R. Kearney, S. W. McKechnie, and A. A. Hoffmann. 2005. A rapid shift in a classic clinal pattern in Drosophila reflecting climate change. Science 308:691-693.

Vicoso, B., and B. Charlesworth. 2006. Evolution on the X chromosome: unusual patterns and processes. Nat Rev Genet 7:645-653.

Watanabe, T. K., and W. H. Lee. 1977. Sterile mutation in *Drosophila melanogaster*. Genetical Research 30:107-113.

Wolf, J., E. Brodie, and M. J. Wade. 2000. Epistasis and the Evolutionary Process. New York, Oxford University Press.

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

Wu, C. I., H. Hollocher, D. J. Begun, C. F. Aquadro, Y. Xu, and M. L. Wu. 1995. Sexual isolation in Drosophila melanogaster: a possible case of incipient speciation. Proc Natl Acad Sci U S A 92:2519-2523.

Wu, C. I., N. A. Johnson, and M. F. Palopoli. 1996. Haldane's rule and its legacy: Why are there so many sterile males? TREE 11:281-284.

Wu, C. I., and C. T. Ting. 2004. Genes and speciation. Nat Rev Genet 5:114-122.

Yukilevich, R., J. Lachance, F. Aoki, and J. R. True. 2008. Long-term adaptation of epistatic genetic networks. Evolution 62:2215-2235.

Yukilevich, R., and J. R. True. 2008a. African morphology, behavior and phermones underlie incipient sexual isolation between us and Caribbean Drosophila melanogaster. Evolution 62:2807-2828.

—. 2008b. Incipient sexual isolation among cosmopolitan Drosophila melanogaster populations. Evolution 62:2112-2121.

Figure legends

Figure 1. Construction of extracted-X lines. Color scheme is as follows: balancer chromosomes are labeled gray, w^{1118} X chromosomes are cross-hatched, genetic backgrounds of balancer stocks used in Stage 1 are labeled with dots, source lines of extracted-X chromosomes are labeled white ($+^{x}$), and source lines of autosomal chromosomes are labeled black ($+^{a}$). The series of crosses depicted here result in extracted-X lines that have a genotype of $+^{x};+^{a};+^{a}$. Different types of synthetic incompatibilities are distinguishable by the offspring generated by each of the crosses in Stage 3.

Figure 2. X-autosome incompatibilities for Set 2. Each row corresponds to a different autosomal background, and each column corresponds to a different X chromosome. X-chromosomes are ordered by increasing latitude (i.e. the rightmost lines have X chromosomes that are derived from northern populations). Incompatible combinations are labeled gray and fertile combinations are labeled white. Two combinations were unable to be tested, and are labeled with cross-hatching. A log-linear test of independence yields a *p*-value of 0.2742 (χ^2 = 5.13, d.f. = 4).

Figure 3. Chromosomal patterns of synthetic lethality. Data are from Set 2 and involve pooled autosomal backgrounds (6326, Sudbury, Rum Cay). The color scheme is the same as Figure 1. There is a significant excess of X-2nd lethality relative to or X-3rd lethality (two-tailed p-value < 0.001; Fisher's exact test).

Figure 4. Biased sex ratios. Error bars are +/- one standard error. Lines are force ranked by % female. Lines with a significant excess of males are labeled in gray (p-value < 0.05 after correcting for multiple tests using the Benjamini and Hochberg false discovery rate). Autosomal background for all lines is 6326 (Set 1).

Figure 5. Geographic patterns of sterility and lethality for different autosomal backgrounds. Lines are classified as northern or southern according to the geographic origins of X chromosomes, and the proportion of northern or southern X-autosome combinations that result in sterility or lethality are depicted. 118 different X chromosomes were tested in Set 1, and 52 different X chromosomes were tested for each autosomal background in Set 2. Northern X chromosomes are labeled white and southern X chromosomes are labeled gray. Note that the Sudbury autosomal background is derived from a northern population and the Rum Cay autosomal background is derived from a southern population. Error bars are +/- one standard error.

Table 1. Levels of synthetic incompatibility

Incompatible sex	Females only	Males only	Both	Neither
Lethality				
Set 1	1	3	1	113
Set 2	7	7	48	92
Sterility				
Set 1*	13	3	22	80
Set 2	3	1	11	139
Lethality or sterility				
Set 1	14	6	23	75
Set 2	10	8	59	77

Table 1. Levels of synthetic incompatibility. Values indicate the number of lines that exhibit a particular type of incompatibility. All lines in Set 1 have a 6326 autosomal background. Data from multiple autosomal backgrounds (6326, Sudbury, and Rum Cay) are pooled for Set 2. * indicates a significant excess of female-specific sterility (one-tailed p-value < 0.05; Fisher's exact test).

Table 2. Chromosomal data

X	X location	X lat.	X long.	Autosome	Set	Fertile/sterile/lethal?
CWP 1	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 2	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 3	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 4	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 5	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 6	Long Island, New York	40.54	-77.23	6326	1	female sterile (regardless of mate)
CWP 7	Long Island, New York	40.54	-77.23	6326	1	female sterile (regardless of mate)
CWP 8	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 9	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 10	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 11	Long Island, New York	40.54	-77.23	6326	1	synthetic sterility (mating pair only, fertile with SBU1)
CWP 12	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 13	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 14	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 15	Long Island, New York	40.54	-77.23	6326	1	female lethal (X-3rd interaction)
CWP 16	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 17	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 18	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 19	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 20	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 21	Long Island, New York	40.54	-77.23	6326	1	fertile

CWP 22	Long Island, New York	40.54	-77.23	6326	1	synthetic sterility (mating pair only, fertile with SBU1)
CWP 23	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 24	Long Island, New York	40.54	-77.23	6326	1	fertile
CWP 25	Long Island, New York	40.54	-77.23	6326	1	fertile
Sudbury 1	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 2	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 3	Sudbury, Ontario	46.31	-81.02	6326	1	female sterile (regardless of mate)
Sudbury 4	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 5	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 6	Sudbury, Ontario	46.31	-81.02	6326	1	female sterile (only sterile w/ SBU1)
Sudbury 7	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 8	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 9	Sudbury, Ontario	46.31	-81.02	6326	1	female sterile (regardless of mate)
Sudbury 10	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 11	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 12	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 13	Sudbury, Ontario	46.31	-81.02	6326	1	female sterile (regardless of mate)
Sudbury 14	Sudbury, Ontario	46.31	-81.02	6326	1	female sterile (regardless of mate)
Sudbury 15	Sudbury, Ontario	46.31	-81.02	6326	1	synthetic sterility (mating pair only, fertile with SBU1)
Sudbury 16	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 17	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 18	Sudbury, Ontario	46.31	-81.02	6326	1	both sexes sterile (regardless of mate)
Sudbury 19	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 20	Sudbury, Ontario	46.31	-81.02	6326	1	fertile

Sudbury 21	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 22	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 23	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 24	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
Sudbury 25	Sudbury, Ontario	46.31	-81.02	6326	1	fertile
23 3	Columbus, Mississippi	33.29	-88.25	6326	1	female sterile (only sterile w/ SBU1)
23 13	Columbus, Mississippi	33.29	-88.25	6326	1	both sexes lethal (X-3rd interaction)
21 29	Birmington, Alabama	33.31	-86.48	6326	1	fertile
21 3	Birmington, Alabama	33.31	-86.48	6326	1	female sterile (only sterile w/ SBU1)
22 21	Tuscaloosa, Alabama	33.12	-87.34	6326	1	both sexes sterile (regardless of mate)
22 14	Tuscaloosa, Alabama	33.12	-87.34	6326	1	fertile
22 10B	Tuscaloosa, Alabama	33.12	-87.34	6326	1	both sexes sterile (regardless of mate)
16 9	Ozark, Alabama	31.27	-85.38	6326	1	both sexes sterile (regardless of mate)
16 15	Ozark, Alabama	31.27	-85.38	6326	1	male sterile (only sterile w/ SBU1)
17 23	Greenville, Alabama	31.49	-86.38	6326	1	male lethal (X-3rd interactions)
17 7	Greenville, Alabama	31.49	-86.38	6326	1	male lethal (X-2nd interactions)
20 5	Selma, Alabama	33.25	-86.53	6326	1	fertile
20 26	Selma, Alabama	33.25	-86.53	6326	1	fertile
18 21	Montgomery, Alabama	32.22	-86.18	6326	1	fertile
18 20	Montgomery, Alabama	32.22	-86.18	6326	1	fertile
18 26	Montgomery, Alabama	32.22	-86.18	6326	1	fertile
18 19	Montgomery, Alabama	32.22	-86.18	6326	1	fertile
4 27	Tampa Bay, Florida	27.56	-82.07	6326	1	synthetic sterility (mating pair only, fertile with SBU1)
28 8	Sebastian, Florida	27.48	-80.28	6326	1	female sterile (regardless of mate)

28 3	Sebastian, Florida	27.48	-80.28	6326	1	fertile
28 5	Sebastian, Florida	27.48	-80.28	6326	1	male sterile (only sterile w/ SBU1)
4 22	Tampa Bay, Florida	27.56	-82.07	6326	1	male lethal
23 15	Columbus, Mississippi	33.29	-88.25	6326	1	fertile
37 11	Long Island, Bahamas	23.1	-75.06	6326	1	fertile
37 22	Long Island, Bahamas	23.1	-75.06	6326	1	both sexes sterile (regardless of mate)
41 24	Rum Cay, Bahamas	23.38	-74.5	6326	1	fertile
41 15	Rum Cay, Bahamas	23.38	-74.5	6326	1	synthetic sterility (mating pair only, fertile with SBU1)
41 18	Rum Cay, Bahamas	23.38	-74.5	6326	1	fertile
45 5	Acklins Island, Bahamas	22.26	-73.59	6326	1	synthetic sterility (mating pair only, fertile with SBU1)
48 7	Lesser Antilles, St. Lucia	13.54	-60.58	6326	1	fertile
45 9	Acklins Island, Bahamas	22.26	-73.59	6326	1	fertile
36 11	George Town, Bahamas	23.31	-75.47	6326	1	fertile
36 7	George Town, Bahamas	23.31	-75.47	6326	1	fertile
43 17	Mayaguana, Bahamas	22.22	-72.54	6326	1	fertile
43 1	Mayaguana, Bahamas	22.22	-72.54	6326	1	both sexes sterile (regardless of mate)
48 13	Lesser Antilles, St. Lucia	13.54	-60.58	6326	1	synthetic sterility (mating pair only, fertile with SBU1)
30 9	Eleuthera, Bahamas	25.23	-76.33	6326	1	fertile
30 4	Eleuthera, Bahamas	25.23	-76.33	6326	1	fertile
35 11	High Rock, Bahamas	25.07	-77.33	6326	1	fertile
35 6	High Rock, Bahamas	25.07	-77.33	6326	1	fertile
29 1	Governor's Harbour, Bahamas	25.15	-76.18	6326	1	fertile
48 1	Lesser Antilles, St. Lucia	13.54	-60.58	6326	1	fertile
48 8	Lesser Antilles, St. Lucia	13.54	-60.58	6326	1	fertile

48 14	Lesser Antilles, St. Lucia	13.54	-60.58	6326	1	both sexes sterile (regardless of mate)
45 11	Acklins Island, Bahamas	22.26	-73.59	6326	1	male sterile (only sterile w/ SBU1)
34 19	McLean's Town, Bahamas	26.38	-77.56	6326	1	fertile
34 18	McLean's Town, Bahamas	26.38	-77.56	6326	1	both sexes sterile (regardless of mate)
33 15	Freeport, Bahamas	26.3	-78.38	6326	1	fertile
33 11	Freeport, Bahamas	26.3	-78.38	6326	1	synthetic sterility (mating pair only, fertile with SBU1)
33 12	Freeport, Bahamas	26.3	-78.38	6326	1	fertile
32 7	Andros Town, Bahamas	24.42	-77.46	6326	1	fertile
35 23	High Rock, Bahamas	25.07	-77.33	6326	1	synthetic sterility (mating pair only, fertile with SBU1)
ZK 184	Lake Kariba, Zimbabwe	-17	27.59	6326	1	female sterile (only sterile w/ SBU1)
ZK 58	Lake Kariba, Zimbabwe	-17	27.59	6326	1	fertile
ZK 159	Lake Kariba, Zimbabwe	-17	27.59	6326	1	fertile
ZS 53B(A)	Sengwa, Zimbabwe	-16.5	28.34	6326	1	female sterile (regardless of mate)
ZS 53B(B)	Sengwa, Zimbabwe	-16.5	28.34	6326	1	female sterile (regardless of mate)
ZS 11	Sengwa, Zimbabwe	-16.5	28.34	6326	1	both sexes sterile (regardless of mate)
MD 16	Mbalang-Djalingo, Cameroon	-5.23	10.05	6326	1	synthetic sterility (mating pair only, fertile with SBU1)
MD 14	Mbalang-Djalingo, Cameroon	-5.23	10.05	6326	1	both sexes sterile (regardless of mate)
MD 34	Mbalang-Djalingo, Cameroon	-5.23	10.05	6326	1	both sexes sterile (regardless of mate)
MD 42	Mbalang-Djalingo, Cameroon	-5.23	10.05	6326	1	fertile
MD 13	Mbalang-Djalingo, Cameroon	-5.23	10.05	6326	1	fertile
ZK 191	Lake Kariba, Zimbabwe	-17	27.59	6326	1	fertile
MD 35	Mbalang-Djalingo, Cameroon	-5.23	10.05	6326	1	fertile
MD 39	Mbalang-Djalingo, Cameroon	-5.23	10.05	6326	1	fertile
MD 26	Mbalang-Djalingo, Cameroon	-5.23	10.05	6326	1	fertile

MD 30	Mbalang-Djalingo, Cameroon	-5.23	10.05	6326	1	synthetic sterility (mating pair only, fertile with SBU1)
37 11	Long Island, Bahamas	23.1	-75.06	Sudbury	2	lethal (X-3rd interaction, both sexes)
37 22	Long Island, Bahamas	23.1	-75.06	Sudbury	2	fertile
41 18	Rum Cay, Bahamas	23.38	-74.5	Sudbury	2	fertile
45 5	Acklins Island, Bahamas	22.26	-73.59	Sudbury	2	male lethal (X-2nd interactions)
45 9	Acklins Island, Bahamas	22.26	-73.59	Sudbury	2	fertile
36 11	George Town, Bahamas	23.31	-75.47	Sudbury	2	lethal (X-2nd interaction, both sexes)
29 1	Governor's Harbour, Bahamas	25.15	-76.18	Sudbury	2	lethal (X-2nd interaction, both sexes)
35 6	High Rock, Bahamas	25.07	-77.33	Sudbury	2	fertile
48 7	Lesser Antilles, St. Lucia	13.54	-60.58	Sudbury	2	lethal (X-2nd interaction, both sexes)
48 8	Lesser Antilles, St. Lucia	13.54	-60.58	Sudbury	2	lethal (X-2nd interaction, both sexes)
48 1	Lesser Antilles, St. Lucia	13.54	-60.58	Sudbury	2	fertile
48 14	Lesser Antilles, St. Lucia	13.54	-60.58	Sudbury	2	fertile
48 13	Lesser Antilles, St. Lucia	13.54	-60.58	Sudbury	2	fertile
34 18	McLean's Town, Bahamas	26.38	-77.56	Sudbury	2	fertile
33 15	Freeport, Bahamas	26.3	-78.38	Sudbury	2	lethal (X-2nd interaction, both sexes)
33 12	Freeport, Bahamas	26.3	-78.38	Sudbury	2	female lethal (X-2nd interaction)
23 13	Columbus, Mississippi	33.29	-88.25	Sudbury	2	lethal (X-autosome dominant)
21 29	Birmington, Alabama	33.31	-86.48	Sudbury	2	fertile
22 21	Tuscaloosa, Alabama	33.12	-87.34	Sudbury	2	female lethal (X-2nd interaction)
22 14	Tuscaloosa, Alabama	33.12	-87.34	Sudbury	2	fertile
22 10B	Tuscaloosa, Alabama	33.12	-87.34	Sudbury	2	fertile
16 9	Ozark, Alabama	31.27	-85.38	Sudbury	2	fertile
16 15	Ozark, Alabama	31.27	-85.38	Sudbury	2	fertile

20 5	Selma, Alabama	33.25	-86.53	Sudbury	2	fertile
20 26	Selma, Alabama	33.25	-86.53	Sudbury	2	sterile
18 21	Montgomery, Alabama	32.22	-86.18	Sudbury	2	fertile
18 20	Montgomery, Alabama	32.22	-86.18	Sudbury	2	fertile
18 26	Montgomery, Alabama	32.22	-86.18	Sudbury	2	fertile
18 19	Montgomery, Alabama	32.22	-86.18	Sudbury	2	fertile
4 27	Tampa Bay, Florida	27.56	-82.07	Sudbury	2	lethal
ZK 58	Lake Kariba, Zimbabwe	-17	27.59	Sudbury	2	lethal (X-2nd interaction, both sexes)
MD 16	Mbalang-Djalingo, Cameroon	-5.23	10.05	Sudbury	2	fertile
MD 14	Mbalang-Djalingo, Cameroon	-5.23	10.05	Sudbury	2	fertile
MD 13	Mbalang-Djalingo, Cameroon	-5.23	10.05	Sudbury	2	lethal (X-autosome dominant)
Sud 2	Sudbury, Ontario	46.31	-81.02	Sudbury	2	lethal (X-2nd interaction, both sexes)
Sud 3	Sudbury, Ontario	46.31	-81.02	Sudbury	2	fertile
Sud 4	Sudbury, Ontario	46.31	-81.02	Sudbury	2	fertile
Sud 5	Sudbury, Ontario	46.31	-81.02	Sudbury	2	fertile
Sud 8	Sudbury, Ontario	46.31	-81.02	Sudbury	2	sterile
Sud 11	Sudbury, Ontario	46.31	-81.02	Sudbury	2	fertile
Sud 15	Sudbury, Ontario	46.31	-81.02	Sudbury	2	lethal (X-2nd interaction, both sexes)
Sud 20	Sudbury, Ontario	46.31	-81.02	Sudbury	2	fertile
Sud 21	Sudbury, Ontario	46.31	-81.02	Sudbury	2	lethal (X-2nd interaction, both sexes)
Sud 22	Sudbury, Ontario	46.31	-81.02	Sudbury	2	fertile
Sud 23	Sudbury, Ontario	46.31	-81.02	Sudbury	2	fertile
Sud 24	Sudbury, Ontario	46.31	-81.02	Sudbury	2	fertile
Cwp 9	Long Island, New York	40.54	-77.23	Sudbury	2	fertile

Cwp 15	Long Island, New York	40.54	-77.23	Sudbury	2	lethal (X-autosome dominant)
Cwp 16	Long Island, New York	40.54	-77.23	Sudbury	2	fertile
Cwp 20	Long Island, New York	40.54	-77.23	Sudbury	2	fertile
Cwp 22	Long Island, New York	40.54	-77.23	Sudbury	2	fertile
Cwp 23	Long Island, New York	40.54	-77.23	Sudbury	2	fertile
37 11	Long Island, Bahamas	23.1	-75.06	Rum Cay	2	male lethal (X lethal only if both autosomes homozygous)
37 22	Long Island, Bahamas	23.1	-75.06	Rum Cay	2	male lethal (X-2nd and X-3rd interactions)
41 18	Rum Cay, Bahamas	23.38	-74.5	Rum Cay	2	lethal (X-2nd and X-3rd interactions, both sexes)
45 5	Acklins Island, Bahamas	22.26	-73.59	Rum Cay	2	lethal (X lethal if both autosomes homozygous, both sexes)
45 9	Acklins Island, Bahamas	22.26	-73.59	Rum Cay	2	fertile
36 11	George Town, Bahamas	23.31	-75.47	Rum Cay	2	lethal (X-2nd and X-3rd interactions, both sexes)
29 1	Governor's Harbour, Bahamas	25.15	-76.18	Rum Cay	2	fertile
35 6	High Rock, Bahamas	25.07	-77.33	Rum Cay	2	sterile
48 7	Lesser Antilles, St. Lucia	13.54	-60.58	Rum Cay	2	lethal (X lethal if both autosomes homozygous, both sexes)
48 8	Lesser Antilles, St. Lucia	13.54	-60.58	Rum Cay	2	lethal (X-3rd interaction, both sexes)
48 1	Lesser Antilles, St. Lucia	13.54	-60.58	Rum Cay	2	lethal (X-autosome dominant)
48 14	Lesser Antilles, St. Lucia	13.54	-60.58	Rum Cay	2	fertile
48 13	Lesser Antilles, St. Lucia	13.54	-60.58	Rum Cay	2	lethal (X-2nd interaction, both sexes)
34 18	McLean's Town, Bahamas	26.38	-77.56	Rum Cay	2	fertile
33 15	Freeport, Bahamas	26.3	-78.38	Rum Cay	2	female lethal (X-2nd interaction)
33 12	Freeport, Bahamas	26.3	-78.38	Rum Cay	2	fertile
23 13	Columbus, Mississippi	33.29	-88.25	Rum Cay	2	lethal (X-autosome dominant)
21 29	Birmington, Alabama	33.31	-86.48	Rum Cay	2	lethal (X-autosome dominant)
22 21	Tuscaloosa, Alabama	33.12	-87.34	Rum Cay	2	female lethal (X-3rd interaction)

22 14	Tuscaloosa, Alabama	33.12	-87.34	Rum Cay	2	fertile
22 10B	Tuscaloosa, Alabama	33.12	-87.34	Rum Cay	2	male lethal (X lethal only if both autosomes homozygous)
16 9	Ozark, Alabama	31.27	-85.38	Rum Cay	2	fertile
16 15	Ozark, Alabama	31.27	-85.38	Rum Cay	2	lethal (X-3rd interaction, both sexes)
20 5	Selma, Alabama	33.25	-86.53	Rum Cay	2	female lethal (X lethal only if both autosomes homozygous)
20 26	Selma, Alabama	33.25	-86.53	Rum Cay	2	lethal (male X lethal if both aut. hom., female X-3rd interaction)
18 21	Montgomery, Alabama	32.22	-86.18	Rum Cay	2	fertile
18 20	Montgomery, Alabama	32.22	-86.18	Rum Cay	2	male lethal (X-3rd interaction)
18 26	Montgomery, Alabama	32.22	-86.18	Rum Cay	2	fertile
18 19	Montgomery, Alabama	32.22	-86.18	Rum Cay	2	NOT TESTED
4 27	Tampa Bay, Florida	27.56	-82.07	Rum Cay	2	female lethal (X-2nd interaction)
ZK 58	Lake Kariba, Zimbabwe	-17	27.59	Rum Cay	2	lethal (X-3rd interaction, both sexes)
MD 16	Mbalang-Djalingo, Cameroon	-5.23	10.05	Rum Cay	2	lethal (male X lethal if both aut. hom., female X-2nd interaction)
MD 14	Mbalang-Djalingo, Cameroon	-5.23	10.05	Rum Cay	2	fertile
MD 13	Mbalang-Djalingo, Cameroon	-5.23	10.05	Rum Cay	2	lethal (X-autosome dominant)
Sud 2	Sudbury, Ontario	46.31	-81.02	Rum Cay	2	lethal (X-autosome dominant)
Sud 3	Sudbury, Ontario	46.31	-81.02	Rum Cay	2	lethal (X-2nd interaction, both sexes)
Sud 4	Sudbury, Ontario	46.31	-81.02	Rum Cay	2	lethal (X-autosome dominant)
Sud 5	Sudbury, Ontario	46.31	-81.02	Rum Cay	2	lethal (X-autosome dominant)
Sud 8	Sudbury, Ontario	46.31	-81.02	Rum Cay	2	lethal (X-autosome dominant)
Sud 11	Sudbury, Ontario	46.31	-81.02	Rum Cay	2	male lethal (X lethal only if both autosomes homozygous)
Sud 15	Sudbury, Ontario	46.31	-81.02	Rum Cay	2	male lethal (X lethal only if both autosomes homozygous)
Sud 20	Sudbury, Ontario	46.31	-81.02	Rum Cay	2	synthetic sterility (mating pair only, fertile with SBU1)
Sud 21	Sudbury, Ontario	46.31	-81.02	Rum Cay	2	synthetic sterility (mating pair only, fertile with SBU1)

Sud 22	Sudbury, Ontario	46.31	-81.02	Rum Cay	2	sterile
Sud 23	Sudbury, Ontario	46.31	-81.02	Rum Cay	2	synthetic sterility (mating pair only, fertile with SBU1)
Sud 24	Sudbury, Ontario	46.31	-81.02	Rum Cay	2	fertile
Cwp 9	Long Island, New York	40.54	-77.23	Rum Cay	2	lethal (X-2nd interaction, both sexes)
Cwp 15	Long Island, New York	40.54	-77.23	Rum Cay	2	fertile
Cwp 16	Long Island, New York	40.54	-77.23	Rum Cay	2	lethal (X-autosome dominant)
Cwp 20	Long Island, New York	40.54	-77.23	Rum Cay	2	fertile
Cwp 22	Long Island, New York	40.54	-77.23	Rum Cay	2	lethal (X-autosome dominant)
Cwp 23	Long Island, New York	40.54	-77.23	Rum Cay	2	lethal (X lethal only if both autosomes homozygous, both sexes)
37 11	Long Island, Bahamas	23.1	-75.06	6326	2	synthetic sterility (mating pair only, fertile with SBU1)
37 22	Long Island, Bahamas	23.1	-75.06	6326	2	lethal (X lethal if both autosomes homozygous, both sexes)
41 18	Rum Cay, Bahamas	23.38	-74.5	6326	2	fertile
45 5	Acklins Island, Bahamas	22.26	-73.59	6326	2	female lethal (X lethal if both autosomes homozygous)
45 9	Acklins Island, Bahamas	22.26	-73.59	6326	2	fertile
36 11	George Town, Bahamas	23.31	-75.47	6326	2	fertile
29 1	Governor's Harbour, Bahamas	25.15	-76.18	6326	2	fertile
35 6	High Rock, Bahamas	25.07	-77.33	6326	2	fertile
48 7	Lesser Antilles, St. Lucia	13.54	-60.58	6326	2	fertile
48 8	Lesser Antilles, St. Lucia	13.54	-60.58	6326	2	male sterile (regardless of mate)
48 1	Lesser Antilles, St. Lucia	13.54	-60.58	6326	2	lethal (X lethal if both autosomes homozygous, both sexes)
48 14	Lesser Antilles, St. Lucia	13.54	-60.58	6326	2	synthetic sterility (mating pair only, fertile with SBU1)
48 13	Lesser Antilles, St. Lucia	13.54	-60.58	6326	2	female sterile (regardless of mate)
34 18	McLean's Town, Bahamas	26.38	-77.56	6326	2	lethal (X-autosome dominant)
33 15	Freeport, Bahamas	26.3	-78.38	6326	2	fertile

33 12	Freeport, Bahamas	26.3	-78.38	6326	2	female sterile (regardless of mate)
23 13	Columbus, Mississippi	33.29	-88.25	6326	2	fertile
21 29	Birmingham, Alabama	33.31	-86.48	6326	2	fertile
22 21	Tuscaloosa, Alabama	33.12	-87.34	6326	2	fertile
22 14	Tuscaloosa, Alabama	33.12	-87.34	6326	2	fertile
22 10B	Tuscaloosa, Alabama	33.12	-87.34	6326	2	fertile
16 9	Ozark, Alabama	31.27	-85.38	6326	2	lethal (X-autosome dominant)
16 15	Ozark, Alabama	31.27	-85.38	6326	2	lethal (X-autosome dominant)
20 5	Selma, Alabama	33.25	-86.53	6326	2	lethal (X-2nd interaction, both sexes)
20 26	Selma, Alabama	33.25	-86.53	6326	2	sterile
18 21	Montgomery, Alabama	32.22	-86.18	6326	2	fertile
18 20	Montgomery, Alabama	32.22	-86.18	6326	2	lethal (X-autosome dominant)
18 26	Montgomery, Alabama	32.22	-86.18	6326	2	fertile
18 19	Montgomery, Alabama	32.22	-86.18	6326	2	fertile
4 27	Tampa Bay, Florida	27.56	-82.07	6326	2	lethal (X-autosome dominant)
ZK 58	Lake Kariba, Zimbabwe	-17	27.59	6326	2	fertile
MD 16	Mbalang-Djalingo, Cameroon	-5.23	10.05	6326	2	both sexes sterile (regardless of mate)
MD 14	Mbalang-Djalingo, Cameroon	-5.23	10.05	6326	2	fertile
MD 13	Mbalang-Djalingo, Cameroon	-5.23	10.05	6326	2	lethal
Sud 2	Sudbury, Ontario	46.31	-81.02	6326	2	fertile
Sud 3	Sudbury, Ontario	46.31	-81.02	6326	2	female sterile (regardless of mate)
Sud 4	Sudbury, Ontario	46.31	-81.02	6326	2	fertile
Sud 5	Sudbury, Ontario	46.31	-81.02	6326	2	fertile
Sud 8	Sudbury, Ontario	46.31	-81.02	6326	2	fertile

Sud 11	Sudbury, Ontario	46.31	-81.02	6326	2	fertile
Sud 15	Sudbury, Ontario	46.31	-81.02	6326	2	lethal (X-autosome dominant)
Sud 20	Sudbury, Ontario	46.31	-81.02	6326	2	fertile
Sud 21	Sudbury, Ontario	46.31	-81.02	6326	2	fertile
Sud 22	Sudbury, Ontario	46.31	-81.02	6326	2	fertile
Sud 23	Sudbury, Ontario	46.31	-81.02	6326	2	fertile
Sud 24	Sudbury, Ontario	46.31	-81.02	6326	2	fertile
Cwp 9	Long Island, New York	40.54	-77.23	6326	2	fertile
Cwp 15	Long Island, New York	40.54	-77.23	6326	2	fertile
Cwp 16	Long Island, New York	40.54	-77.23	6326	2	lethal
Cwp 20	Long Island, New York	40.54	-77.23	6326	2	fertile
Cwp 22	Long Island, New York	40.54	-77.23	6326	2	NOT TESTED
Cwp 23	Long Island, New York	40.54	-77.23	6326	2	fertile

Table 2. Chromosomal data. Location data (latitude and longitude) are included for each extracted-X line. Sterility and lethality data for each line are also included.

Figures

Figure 1. Construction of extracted-X lines

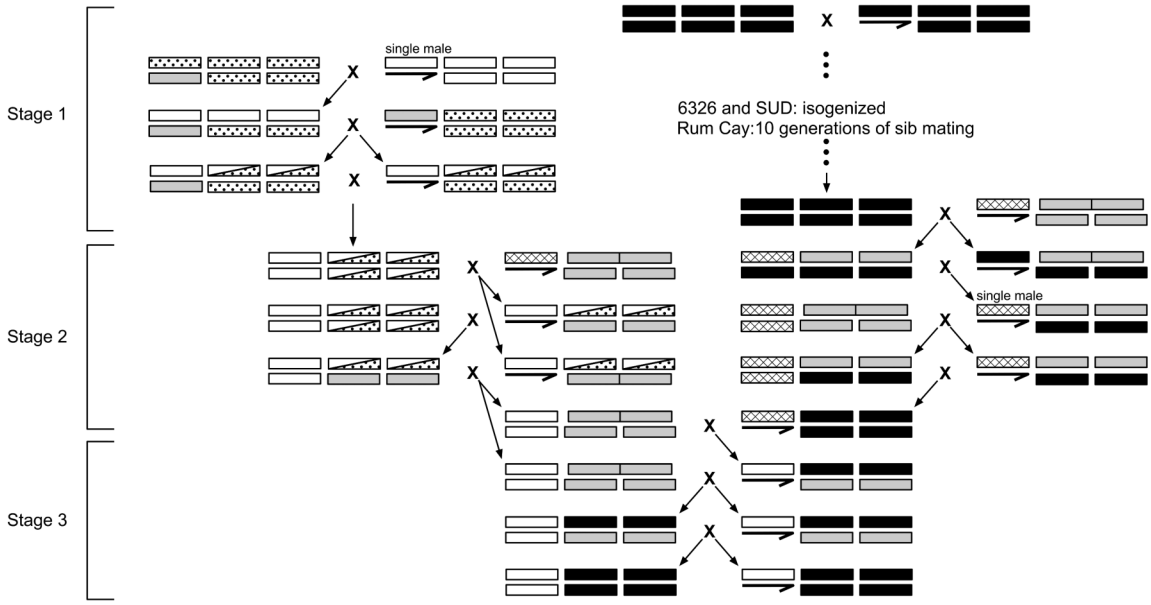


Figure 2. X-autosome incompatibilities for Set 2

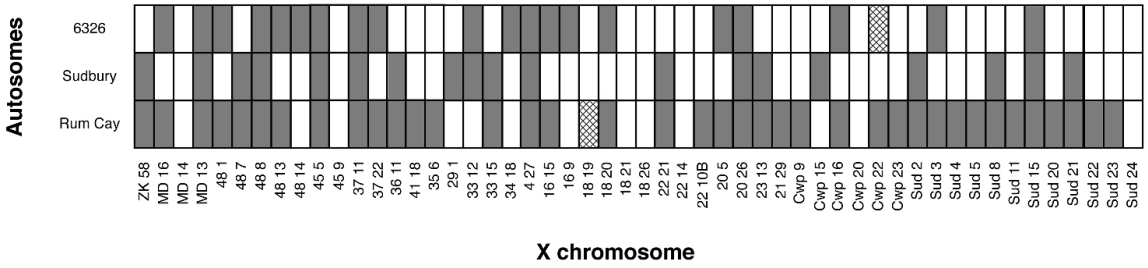


Figure 3. Chromosomal patterns of synthetic lethality

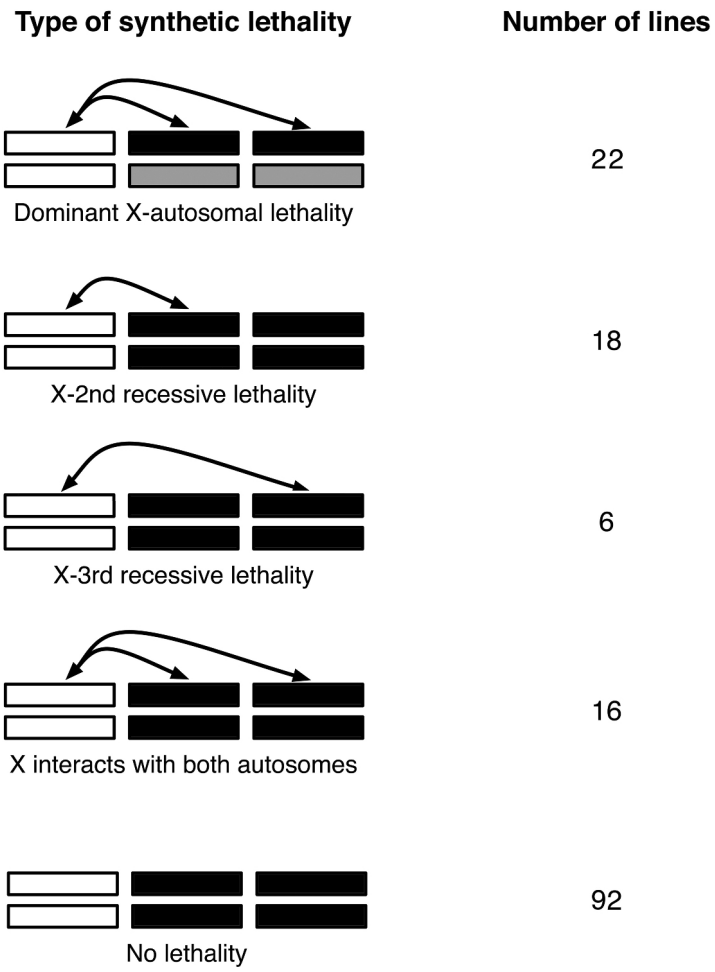


Figure 4. Biased sex ratios

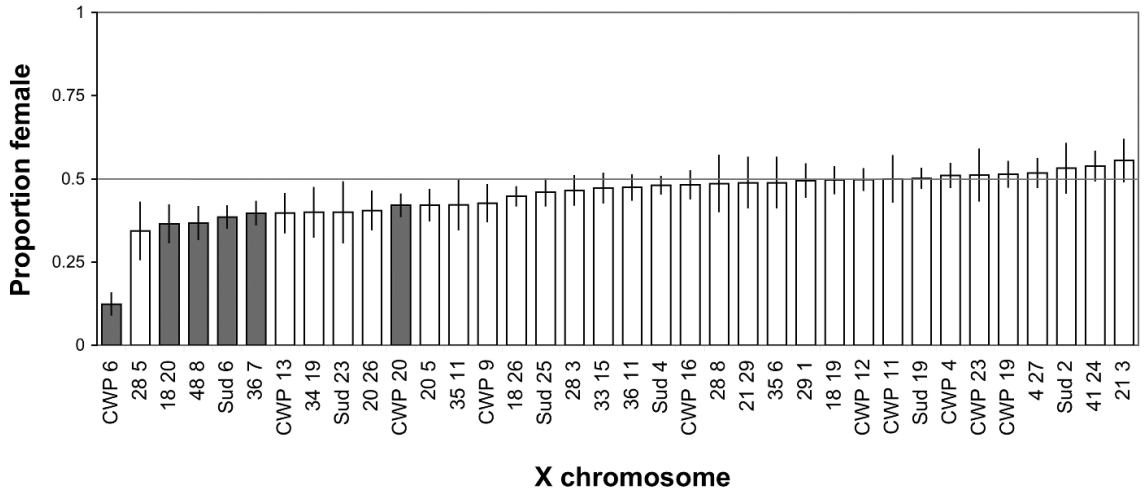
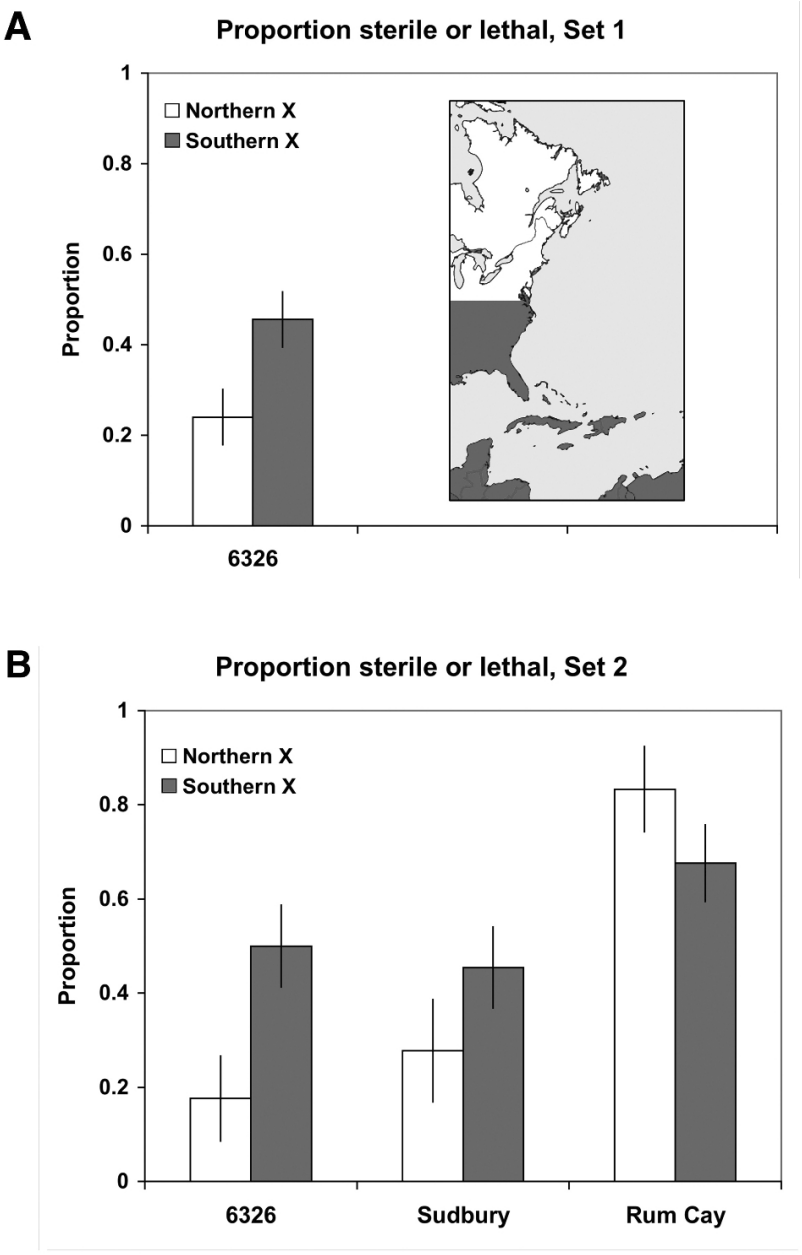


Figure 5. Geographic patterns of sterility and lethality for different autosomal backgrounds



Chapter 3

The population genetics of X-autosome synthetic incompatibilities, with implications for Haldane's Rule

Joseph Lachance, Norman A Johnson, and John R. True

Abstract

Epistatic interactions are widespread, and many of these interactions involve combinations of alleles at different loci that are deleterious when present in the same individual. The genetic environment of sex-linked genes differs from autosomal genes, suggesting that the population genetics of interacting X-linked and autosomal alleles may be complex. Using both analytical theory and computer simulations, we analyzed the evolutionary trajectories and mutation-selection balance conditions for X-autosome synthetic lethals and steriles. Allele frequencies of X-autosome incompatibilities follow a set of fundamental trajectories. Stable equilibria exist, and they can involve either fixation of autosomal or X-linked alleles. The exact equilibrium depends on whether synthetic alleles are dominant or recessive, and whether fitness effects are seen in males, females, or both sexes. Many X-autosome incompatibilities in natural populations may be cryptic, appearing to be single locus effects because one locus is fixed. If multiple allopatric populations are considered, allele frequencies can diverge. This results in reduced fitness upon secondary contact. When coupled with male hemizygosity of X chromosomes, divergent frequencies can cause Haldane's rule-like patterns to arise prior to speciation.

Introduction

Non-additive (epistatic) genetic interactions are ubiquitous in biological organisms (Mackay 2004; Moore 2003; Phillips 2008; Wolf et al. 2000). The nature and intensity of such epistasis has important implications for several areas of inquiry in evolutionary biology, most notably the evolution of sex (de Visser and Elena 2007; Otto and Lenormand 2002),and the evolution of hybrid incompatibility (speciation) (Coyne and Orr 2004; Dobzhansky 1937; Gavrilets 2004; Johnson 2000; Muller 1942; Orr 1995). Epistasis is also important in mapping genes involved in quantitative traits; it may lead to the identification of different QTLs in different populations (Liao et al. 2001; Wade 2001), and it can also explain some of the "missing heritability" of genome-wide association studies in human genetics (Manolio et al. 2009; Moore et al. 2010).

An important subset of epistatic interactions involves deleterious phenotypes that occur only in the presence of a combination of two or more alleles at different loci. Theodosius Dobzhansky first described synthetic lethals as interacting chromosomes whose effects were released through recombination (Dobzhansky 1946). A modern definition of synthetic lethality is "any combination of two separately non-lethal mutations that leads to inviability" ((Ooi et al. 2006) p. 57). On a broader scale, synthetic incompatibilities can involve either synthetic lethality or synthetic sterility. As synthetic incompatibilities need not lead to complete lethality or sterility, Phillips and Johnson used the term "synthetic deleterious loci" (SDL) (Phillips and Johnson 1998).

From a theoretical perspective, synthetic incompatibilities are important because they influence allele frequency trajectories and patterns of standing genetic variation. Because genotypes at one locus can mask the effects genotypes at another locus, synthetic

incompatibilities are able to segregate at higher frequencies than non-epistatic deleterious alleles. For example, at mutation-selection balance, the expected frequencies of double recessives are on the order of the quartic root of (μ / s), where μ and s are mutation rates and selection coefficients, respectively (Phillips and Johnson 1998). This is in contrast to single locus expectations, which are on the order of the square root of (μ / s) (Charlesworth and Charlesworth 2010). If mutation rates differ between loci involved in synthetic lethality, mutation pressure can slowly cause one locus to fix. Interestingly, genetic systems during this convergence process may appear to have single locus mutation-selection balance with incomplete penetrance (Christiansen and Frydenberg 1977). Dominance complicates the population genetics of autosomal synthetic lethals, as does linkage (Bengtsson and Christiansen 1983; Phillips and Johnson 1998). Genomic data reveals that DNA sequences evolve at different rates for X chromosomes and autosomes (Vicoso and Charlesworth 2006). Important differences exist between X-linked and autosomal variation: males are hemizygous for most X-linked genes and X chromosomes spend two-thirds of their time in females. Asymmetry between X-linked and autosomal genes can influence evolutionary trajectories (Charlesworth et al. 1987). Are incompatible alleles more likely to segregate at higher frequencies at X-linked loci or autosomal loci? Because existing theory has focused on autosomal synthetic incompatibilities, there is a clear need to determine how X-autosome interactions affect the genetics of populations.

Ample evidence of synthetic incompatibilities exists in natural and experimental populations. Many early examples of synthetic lethality were found in a variety of *Drosophila* species (Dobzhansky 1946; Lucchesi 1968; Temin et al. 1969; Thompson 1986). More recently, synthetic lethality and sterility was found between naturally segregating X chromosomes and autosomes in *D. melanogaster* (Lachance and True 2010). Interestingly, this study found higher

levels of incompatibility between chromosomes that were derived from different geographical locations. Synthetic incompatibilities have been used to map genetic networks in *C. elegans* and *S. cerevisiae* (Lehner et al. 2006; Ooi et al. 2006; Tong et al. 2004). Evidence from these studies suggests that less than four percent of all pairwise combinations are synthetic lethal or sublethal and most genes are synthetic lethal with less than ten other genes (Davierwala et al. 2005; Hartman et al. 2001). Experiments using RNA viruses also indicate that many synthetic interactions are deleterious (Sanjuan et al. 2004).

Synthetic incompatibilities within species share features with the hybrid incompatibility seen between species. Hybrid incompatibility generally involves at least two interacting loci, wherein alleles at one locus from one of the species are incompatible with alleles at another locus from the other species (Bateson 1909; Coyne and Orr 2004; Johnson 2010; Muller 1942). Such hybrid incompatibility interactions are known as Bateson-Dobzhansky-Muller (BDM) incompatibilities, named for the three pioneers who formulated this model (Bateson 1909; Dobzhansky 1936; Muller 1942). BDM incompatibilities involve alleles that are not deleterious in their normal genetic background but have reduced fitness when brought together in hybrids (Coyne and Orr 2004; Orr and Turelli 2001). DBM incompatibilities can accumulate either via single substitutions in different populations or via multiple substitution events in a single population (Johnson 2010; Presgraves 2010). X-autosome incompatibilities appear to be disproportionately common in *Drosophila* (Coyne and Orr 2004).

X-autosome incompatibilities also may explain a long-standing empirical generalization known as Haldane's rule (Coyne and Orr 2004). This pattern, first noted by J.B.S. Haldane (Haldane 1922), states that if only one sex of a hybrid has reduced fitness, it tends to be the heterogametic sex (Laurie 1997; Orr 1995; Wu and Davis 1993). One hypothesis for Haldane's

rule proposes that recessive alleles are more likely to cause reduced fitness in the heterogametic sex (Turelli and Orr 1995; Turelli and Orr 2000). Alternatively, another hypothesis is that incompatibilities may accumulate faster on the X chromosome (Lehner et al. 2006; Tao et al. 2003). Empirical evidence for both hypotheses has been found in multiple taxa (Coyne and Orr 2004; Laurie 1997; Turelli and Begun 1997; Wu and Davis 1993).

Because X-autosome synthetic lethality manifests in aberrant sex ratios, these types of interactions are easier to detect than autosome-autosome synthetic lethality. Numerous X-autosome interactions have been implicated in post-zygotic reproductive isolation between closely related species of *Drosophila* (Coyne and Orr 2004; Davis et al. 1994; Presgraves 2003; Tao et al. 2003; True et al. 1996). How likely are X-autosome incompatibilities to arise between allopatric populations and what theoretical conditions favor the presence of Haldane's rule early in the speciation process?

This paper contains a theoretical treatment of the dynamics of X-autosome synthetic incompatibilities. First, a general model is developed that includes sex-specific allele frequencies and fitnesses. We then consider synthetic lethality (where both sexes have reduced fitness) and synthetic sterility (where only a single sex has reduced fitness). In each of these scenarios we explore allele frequency trajectories and mutation-selection balance at both loci. We also examine the theoretical implications of secondary contact between allopatric populations that have diverged in allele frequencies at X-linked and autosomal loci. Testable predictions arise, including the expectation that autosomal alleles are more likely to segregate at high frequencies if X-autosome incompatibilities are dominant, and X-linked alleles are more likely to segregate at high frequencies if X-autosome incompatibilities are recessive.

General model

A two-locus population genetics model is used. Generations are discrete and population size is assumed to be infinite. Equal sex ratios prior to selection are assumed. One locus is autosomal (segregating alleles: A and a), and the other locus is X-linked (segregating alleles: X and x). Synthetic incompatibilities exist between A and X alleles. We allow these incompatibilities to be sex-specific and involve various types of dominance. Each generation, natural selection occurs subsequent to forward mutation (mutations rates from a to A are given by μ_A, and mutations from x to X are given by μ_X). For mathematical simplicity, we ignore back mutation. Throughout this paper, figures use high mutation rates ($\mu = 0.01$) to assist in visualizing general patterns.

Because X-linked and autosomal alleles are on different chromosomes, only allele frequencies need to be tracked in our study (haplotype frequencies do not need to be tracked). However, sex-linkage requires tracking both male and female allele frequencies. Allele frequencies in males are given by A_{male} and X_{male}, and allele frequencies in females are given by A_{fem} and X_{fem} (Table 1). These are post-selection allele frequencies. Selection coefficients in males and females are s_{male} and s_{fem}, respectively. Dominance terms of autosomal and sex-linked loci are h_A and h_X, respectively. Dominance terms equal to zero imply that the deleterious allele is completely recessive to the non-deleterious allele; dominance terms equal to one imply that the deleterious allele is completely dominant. Individuals containing A and X alleles have reduced fitness (Table 2). For example, $AAXx$ females have a fitness of $1 - s_{fem}h_X$. Reduced fitness in double heterozygotes is assumed to be multiplicative. In the absence of mutation, populations will become fixed for a and x alleles.

The algebra of subsequent equations is substantially more tractable if the following expressions are used:

$$\overline{A} = \frac{A_{male} + A_{fem}}{2} \tag{1}$$

$$\overline{X} = \frac{X_{male} + X_{fem}}{2} \tag{2}$$

$$A^* = A_{male} A_{fem} \tag{3}$$

$$X^* = X_{male} X_{fem} \tag{4}$$

$\overline{A}$ and $\overline{X}$ are sex-averaged allele frequencies (unweighted by the probability that X chromosomes appear in males or females). A^* and X^* are the products of sex-specific allele frequencies. When sex-specific frequencies are the same: $\overline{A} - A^* = A(1 - A)$ and $\overline{X} - X^* = X(1 - X)$.

Allele frequency changes due to selection

As per classical population genetics, allele frequencies in subsequent generations are determined by present allele frequencies and the ratio of average excess to mean fitness (Crow and Kimura 1970). The average excess is equal to the difference between the marginal fitness of an allele and the mean fitness of a population. A general equation for allele frequency change is given by:

$$\Delta p = \frac{p(\dot{w} - \overline{w})}{\overline{w}} \tag{5}$$

In Equation 5: p is allele frequency, $\dot{w}$ is the marginal fitness of an allele, and $\overline{w}$ is mean fitness.

In a genetic system with synthetic incompatibilities changes in allele frequency at one locus depend on allele frequencies at the other locus. Such dependence affects both marginal fitnesses and mean fitnesses. Males can be one of six different genotypes, and females can be one of nine different genotypes. Allele frequency trajectories also reflect the fact that X chromosomes in males are maternally derived. As each sex can be viewed as a different selective environment, equations for allele frequency change incorporate sex-specific mean fitnesses. Mean fitnesses incorporate both the fitnesses of each genotype and the probability of observing each genotype (assuming independent assortment). Mean fitnesses of males and females are:

$$\overline{w}_{male} = 1 - s_{male} A^* X_{fem} - 2 s_{male} h_A \left(\overline{A} - A^* \right) X_{fem} \tag{6}$$

$$\overline{w}_{fem} = 1 - s_{fem} A^* X^* - 2 s_{fem} h_A \left(\overline{A} - A^* \right) X^* - 2 s_{fem} h_X A^* \left(\overline{X} - X^* \right) - 4 s_{fem} h_A h_x \left(\overline{A} - A^* \right) \left(\overline{X} - X^* \right) \tag{7}$$

Male and female changes in autosomal allele frequencies follow from the general equation for allele frequency change (Equation 5) and the assumptions of the general model in this paper. In both sexes, the marginal fitness of *A* alleles is weighted by the probably that an *A* allele is found in an *AA* homozygote or *Aa* heterozygote. In males, the marginal fitness is

weighted by the probably that an *A* allele is found in an *Xy* or *xy* individual. In females, the marginal fitness is weighted by the probably that an *A* allele is found in an *XX*, *Xx*, or *xx* individual. After extensive algebra, allele frequency changes due to selection are given by:

$$\Delta A_{male,sel} = \frac{\left(\overline{A} - A_{male}\right) - s_{male}\left[A^*(1 - A_{male}) + h_A(\overline{A} - A^*)(1 - 2A_{male})\right]X_{fem}}{1 - s_{male}\left(A^* + 2h_A\left(\overline{A} - A^*\right)\right)X_{fem}} \tag{8}$$

$$\Delta A_{fem,sel} = \frac{(\overline{A} - A_{fem}) - s_{fem}\left[A^*(1 - A_{fem}) + h_A(\overline{A} - A^*)(1 - 2A_{fem})\right]\left(X^* + 2h_X(\overline{X} - X^*)\right)}{1 - s_{fem}\left(A^* + 2h_A\left(\overline{A} - A^*\right)\right)\left(X^* + 2h_X\left(\overline{X} - X^*\right)\right)} \tag{9}$$

Autosomes spend half their time in males and half in females. This means that the overall change in autosomal frequency due to selection is simply the average of the sex-specific changes.

$$\Delta A_{sel} = \frac{\Delta A_{male,sel} + \Delta A_{fem,sel}}{2} \tag{10}$$

Male and female changes in sex-linked allele frequencies follow from Equation 5 and the assumptions of the general model in this paper. In both sexes, the marginal fitness of *X* alleles is weighted by the probably of observing different autosomal genotypes. However, the marginal fitness in females is also weighted by the probably that an *X* allele is found in an *XX* homozygote or *Xx* heterozygote. After extensive algebra, changes in sex-linked allele frequencies are given by:

$$\Delta X_{male,sel} = \frac{\left(X_{fem} - X_{male}\right) - s_{male}\left[A^*(1 - X_{male}) + h_A(\overline{A} - A^*)(1 - 2X_{male})\right]X_{fem}}{1 - s_{male}\left(A^* + 2h_A\left(\overline{A} - A^*\right)\right)X_{fem}} \tag{11}$$

$$\Delta X_{fem,sel} = \frac{(\overline{X} - X_{fem}) - s_{fem}\left(A^* + 2h_A(\overline{A} - A^*)\right)\left[X^*(1 - X_{fem}) + h_X(\overline{X} - X^*)(1 - 2X_{fem})\right]}{1 - s_{fem}\left(A^* + 2h_A\left(\overline{A} - A^*\right)\right)\left(X^* + 2h_X\left(\overline{X} - X^*\right)\right)} \tag{12}$$

X chromosomes spend one-third their time in males and two-thirds of the time in females. The overall change in X frequencies due to selection incorporates this weighting.

$$\Delta X_{sel} = \frac{\Delta X_{male,sel} + 2\Delta X_{fem,sel}}{3} \tag{13}$$

Allele frequency changes due to mutation

Mutation increases the frequency of *A* and *X* alleles, while selection decreases the frequency of *A* and *X* alleles. This implies that a balance between mutation and selection exists for X-autosome synthetic incompatibilities. Mutation pressure is stronger when alleles are rare, leading to a form of negative frequency dependence. Allele frequency changes due to mutation are:

$$\Delta A_{male,mut} = (1 - A_{male})\mu_A \tag{14}$$

$$\Delta A_{fem,mut} = (1 - A_{fem})\mu_A \tag{15}$$

$$\Delta X_{male,mut} = (1 - X_{male})\mu_X \tag{16}$$

$$\Delta X_{fem,mut} = (1 - X_{fem})\mu_X \tag{17}$$

Overall changes in allele frequencies due to mutation incorporate the relative amount of time that alleles spend in males or females.

$$\Delta A_{mut} = \frac{\Delta A_{male,mut} + \Delta A_{fem,mut}}{2} \tag{18}$$

$$\Delta X_{mut} = \frac{\Delta X_{male,mut} + 2\Delta X_{fem,mut}}{3} \tag{19}$$

Allele frequency trajectories and mutation-selection balance equations in subsequent sections were tested by numerical iteration of the equations for allele frequencies (Equations 8, 9, 11, 12, 14, 15, 16, and 17).

Allele frequencies can be represented in a Cartesian plane where the x-axis refers to X-linked frequency and the y-axis refers to autosomal frequency. In contrast to single locus dynamics, an X-autosome population has multiple dimensions of movement. Parameter values that satisfy $\Delta A_{sel} + \Delta A_{mut} = 0$ need not be parameter values that satisfy $\Delta X_{sel} + \Delta X_{mut} = 0$. This implies that the presence of internal equilibria, where both loci are polymorphic, is not guaranteed. Instead, mutation-selection balance in X-autosome systems may involve fixation of one locus. Autosomal frequencies do not change when the sum of the right hand sides of

Equations 10 and 18 is zero, and X-linked frequencies do not change when the sum of the right hand sides of Equations 13 and 19 is zero. Mutation-selection equilibria in X-autosome systems result when both of these conditions are met. Simultaneously solving the non-linear allele frequency change equations for autosomal and X-linked alleles yields equilibria conditions. However, many of these solutions do no involve valid parameter values (allele frequencies must range between 0 and 1 and individuals cannot have fitnesses that are less than zero). Because of this, the number of possible equilibria varies for different levels of dominance and selection coefficients.

Damped oscillations of sex-specific allele frequencies

When sex-linked allele frequencies differ for males and females, evolutionary trajectories can be complex. Existing single locus theory indicates that sex-linked allele frequencies approach equilibrium via a series of damped oscillations, whereby the sex with the higher allele frequency alternates each generation ((Crow and Kimura 1970), pp. 44-47). Because X-autosome genetic systems contain X chromosomes, this suggests that damped oscillations may also occur when there are X-autosome synthetic incompatibilities.

An initial understanding of sex-specific oscillations can be obtained by inspecting the leading terms in the numerators of Equations 8, 9, 11, and 12. These terms are $\overline{A} - A_{male}$, $\overline{A} - A_{fem}$, $X_{fem} - X_{male}$, and $\overline{X} - X_{fem}$, respectively. When male allele frequencies are less than female allele frequencies, the leading terms suggest that male frequencies will increase and female frequencies will decrease next generation. Conversely, when male frequencies are

greater than female frequencies, the leading terms suggest that male frequencies will decrease and female frequencies will increase next generation. Selection also modifies sex-specific allele frequencies, and the numerators of Equations 8, 9, 11, and 12 contain more than just leading terms. The complexity of these equations indicates that numerical iteration of allele frequency change equations is required.

Using the equations for allele frequency change, allele frequencies were iterated for a range of fitnesses (Figure 1). Initial conditions were: $A_{male} = 1,\ A_{fem} = 0,\ X_{male} = 1$, and $X_{fem} = 0$. Mutation rates were the same at both loci ($\mu_A = 0.0001$ and $\mu_X = 0.0001$). Fitnesses scenarios included: neutrality ($s_{male} = 0$ and $s_{fem} = 0$), complete synthetic lethality ($s_{male} = 1$ and $s_{fem} = 1$), complete synthetic male sterility ($s_{male} = 1$ and $s_{fem} = 0$), and complete synthetic female sterility ($s_{male} = 0$ and $s_{fem} = 1$). Technically speaking, male sterility and female sterility scenarios also apply to situations where there is sex-specific lethality. Synthetic incompatibilities were assumed to be recessive ($h_A = 0$ and $h_X = 0$).

Figure 1 indicates that differences between male and female frequencies were minimal after ten generations. Sex-linked allele frequencies converged via a series of damped oscillations under each fitness scenario. Under neutrality, autosomal frequencies converged after a single generation. When males were under selection, autosomal frequency changes were non-monotonic. Sex-specific differences in autosomal frequencies persisted longer when males were under selection (compare panels in Figure 1). Because sex-specific differences in allele frequency were minimal after ten generations for a range of fitness scenarios, subsequent subsections of this paper assume that male and female allele frequencies are equal.

Synthetic lethality

Under synthetic lethality, selection acts in both sexes. The relative strength of selection depends on autosomal and sex-linked allele frequencies and dominance coefficients at both loci. For mathematical simplicity, we assume that the strength of synthetic lethality is the same for both sexes ($s_{male} = s_{fem} = s$), and mutation rates are the same at both loci ($\mu_A = \mu_X = \mu$). Allele frequency trajectories for synthetic lethals are shown in Figure 3. Each panel in this figure was generated by numerical iteration of the equations for allele frequency change. In Figure 3, selection can be viewed as a force pushing to the lower left of each panel, while mutation pushes populations to the upper right. For each set of dominance coefficients, populations were iterated 20,000 generations from eight different start states.

Genetic systems containing synthetic lethality evolve via a two-part process. First, there is an approach to a *fundamental trajectory* (using the terminology of (Christiansen and Frydenberg 1977)). Next, the system moves along this trajectory towards stable equilibria for both loci. These equilibria involve single locus mutation-selection balance frequencies. If *A* is completely dominant and *X* is completely recessive, the population will ultimately fix the *A* allele (Figure 3A). If both incompatible alleles are dominant, the ultimate fate of the population depends on starting allele frequencies (Figure 3B). An internal unstable equilibrium exists when *A* and *X* are dominant. If *A* is completely recessive, the population will ultimately fix the *X* allele (Figure 3C-D). While along the fundamental trajectory, X-autosome incompatibilities are able to segregate at moderate frequencies. For example, if both loci have the same frequency, completely recessive synthetic lethal alleles will have allele frequencies above 2% ($s = 1$ and $\mu = 0.00001$). Inspection of Figure 3 reveals that fundamental trajectories are asymmetric with respect to X chromosome and autosomal allele frequencies. Retrograde movement of allele

frequencies can also occur. For example, consider a double recessive genetic system where a population starts at $A = 0$ and $X = 0$. Under this scenario, *X* alleles initially increase in frequency only to subsequently decrease in frequency (Figure 3C).

Approach to equilibrium is slow

Because increases in allele frequencies are due to mutation pressure, the approach to the fundamental trajectory and stable equilibria can be quite slow. For example, a population initially lacking *A* and *X* alleles will have allele frequencies of $A = 0.0202$ and $X = 0.3574$ after 20,000 generations of evolution ($s = 1,\ \mu = 0.0001,\ h_A = 0, h_X = 0$). Slow approaches have also been observed for synthetic incompatibilities involving pairs of autosomal loci (Christiansen and Frydenberg 1977). This suggests that large populations are likely to contain incompatibilities that still segregate at both loci, despite the presence of two-locus equilibria that involve fixation of one locus. Finite populations are less likely to contain polymorphisms at both loci, as populations are able to drift along the fundamental trajectory. However, finite population sizes also open up the possibility of alternative outcomes (perhaps fixing alternative alleles than those expected from the stable equilibria for both loci).

Mutation rates and strength of selection

Lower mutation rates result in fundamental trajectories with low allele frequencies (Figure 4A). Similarly, stronger selection also results in fundamental trajectories with low allele frequencies (Figure 4B). As in the one-locus case, allele frequency trajectories depend more on the ratio of mutation rate to selection coefficients (μ/s) than either parameter by itself. Subsequent equations illustrate this point more rigorously. Stable equilibria for both loci also

depend on the ratio of mutation rates and selection coefficients. However, approach speeds to stable equilibria depend on the magnitudes of individual parameters.

Complete recessivity at both loci

Analytic expressions for mutation-selection balance of recessive synthetic lethal alleles can be derived. For mathematical simplicity, the strength of synthetic lethality is assumed to be the same for both sexes ($s_{male} = s_{fem} = s$), and mutation rates are assumed the same at both loci ($\mu_A = \mu_X = \mu$). Complete recessivity at both loci is assumed ($h_A = h_X = 0$). At equilibrium, the genetic load due to synthetic lethality is minimal, allowing us to assume that mean fitness is close to one. Mutation-selection balance conditions were calculated separately for autosomal and sex-linked alleles, and plotted in Figure 2A.

Recalling the fact that autosomal alleles spend equal times in each sex, the net change in autosome frequencies due to completely recessive synthetic lethality is:

$$\Delta A_{sel} = \frac{-sA^2(1-A)X(1+X)}{2} \tag{20}$$

Autosomal mutation-selection balance occurs when $\Delta A_{sel} + \Delta A_{mut} = 0$, and allele frequency changes due to mutation are given by Equations 14, 15, and 18. Given complete recessivity and synthetic lethality, mutation-selection balance of autosomal alleles occurs when:

$$\frac{\mu}{s} = \frac{A^2X(1+X)}{2} \tag{21}$$

Recalling the fact that one third of X chromosomes are present in males and two thirds of X chromosomes are present in females, the net change in sex-linked frequencies due to completely recessive synthetic lethality is:

$$\Delta X_{sel} = \frac{-sA^2X(1-X)(1+2X)}{3} \tag{22}$$

Sex-linked mutation-selection balance occurs when $\Delta X_{sel} + \Delta X_{mut} = 0$, and allele frequency changes due to mutation are given by Equations 16, 17, and 19. Given complete recessivity and synthetic lethality, mutation-selection balance of sex-linked alleles occurs when:

$$\frac{\mu}{s} = \frac{A^2X(1+2X)}{3} \tag{23}$$

Are there any conditions that satisfy mutation-selection balance for both autosomal and sex-linked alleles? In Figure 2A, populations below both curves increase in frequency of *A* and *X*, populations above both curves decrease in frequency of *A* and *X*, and populations between the autosomal and sex-linked curves increase in *X* and decrease in *A*. This implies that the *X* allele will ultimately be fixed. In Figure 2B, autosomal and X-linked equilibria curves cross at a single point. This unstable internal equilibrium involves a population with low frequencies of the *A* allele and moderate frequencies of the *X* allele. Discrepancies between the positions of internal equilibria in Figures 2,3, and 6 arise because of minor differences in sex-specific allele frequencies. The Transitive Probability of Equality allows Equations 21 and 23 to be combined.

$$\frac{A^2X(1+X)}{2} = \frac{A^2X(1+2X)}{3} \tag{24}$$

Equation 24 is satisfied when $X = 1$. This indicates that complete recessivity at both loci is expected to ultimately result in a population that is fixed for the X allele. This is consistent with the evolutionary trajectories of synthetic lethal alleles in Figure 3C.

Complete dominance at both loci

Analytic expressions for mutation-selection balance of dominant synthetic lethal alleles can also be derived. The strength of synthetic lethality is assumed to be the same for both sexes, and mutation rates are assumed the same at both loci. Complete dominance at both loci is assumed ($h_A = h_X = 1$). Mutation-selection balance conditions were calculated separately for autosomal and sex-linked alleles, and plotted in Figure 2B. Because synthetic lethality is observed at low frequencies when alleles are dominant, the equations in this subsection incorporate mean fitness terms.

Recalling the fact that autosomal alleles spend equal times in each sex, the net change in autosome frequencies due to completely dominant synthetic lethality is:

$$\Delta A_{sel} = \frac{-sA(1-A)^2X\left(\overline{w}_{fem} + (2-X)\overline{w}_{male}\right)}{2\overline{w}_{male}\overline{w}_{fem}} \tag{25}$$

Given complete dominance and synthetic lethality, mutation-selection balance of autosomal alleles occurs when the following non-linear equation is satisfied:

$$\frac{\mu}{s} = \frac{A(1-A)X\left(\overline{w}_{fem} + (2-X)\overline{w}_{male}\right)}{2\overline{w}_{male}\overline{w}_{fem}} \tag{26}$$

Recalling the fact that one third of X chromosomes are present in males and two thirds of X chromosomes are present in females, the net change in sex-linked frequencies due to completely recessive synthetic lethality is:

$$\Delta X_{sel} = \frac{-sAX\left((1-2X+AX)\overline{w}_{fem} + 2(2-A)(1-X)^2\overline{w}_{male}\right)}{3\overline{w}_{male}\overline{w}_{fem}} \tag{27}$$

Given complete recessivity and synthetic lethality, mutation-selection balance of sex-linked alleles occurs when a rather unwieldy nonlinear equation is satisfied.

$$\frac{\mu}{s} = \frac{AX\left((1-2X+AX)\overline{w}_{fem} + 2(2-A)(1-X)^2\overline{w}_{male}\right)}{3(1-X)\overline{w}_{male}\overline{w}_{fem}} \tag{28}$$

Synthetic male sterility

Under synthetic male sterility, selection only acts in males ($s_{fem} = 0$). The relative strength of selection depends on autosomal frequencies, sex-linked allele frequencies, and autosomal dominance coefficients. Because selection only acts in hemizygous males, sex-linked

dominance coefficients do not affect evolutionary trajectories. For mathematical simplicity, mutation rates are assumed the same at both loci ($\mu_A = \mu_X = \mu$). Allele frequency trajectories for synthetic male sterile alleles are shown in Figure 5. Each panel in this figure was generated by numerical iteration of the equations for allele frequency change. For each set of dominance coefficients, populations were iterated 20,000 generations from eight different start states.

The dynamics of synthetic male sterile alleles are broadly similar to synthetic lethal alleles: allele frequency changes involve approaches to fundamental trajectories and stable equilibria. Dominant autosomal alleles result in a stable equilibrium with fixation of the *A* allele (Figure 5A), and recessive autosomal alleles result in a stable equilibrium with fixation of the *X* allele (Figure 5B). Populations are found at single locus mutation-selection balance frequencies when at either of these stable equilibria. Because selection only acts in one sex, segregating allele frequencies of male sterile alleles are slightly higher than expected under synthetic lethality. One major difference is that X-autosome male sterility is not affected by dominance of X chromosomes. Once again, approaches to stable equilibria are slow.

Complete autosomal recessivity

Analytic expressions can be derived for mutation-selection balance of recessive synthetic male sterile alleles ($h_A = 0$). Mutation-selection balance conditions were calculated separately for autosomal and sex-linked alleles, and plotted in Figure 2C. The net change in autosome frequencies due to completely recessive synthetic male sterility is:

$$\Delta A_{sel} = \frac{-s_{male} A^2 (1 - A) X}{2} \qquad (29)$$

Given complete recessivity and synthetic male sterility, mutation-selection balance of autosomal alleles occurs when:

$$\frac{\mu}{s_{male}} = \frac{A^2 X}{2} \tag{30}$$

The net change in sex-linked frequencies due to completely recessive synthetic male sterility is:

$$\Delta X_{sel} = \frac{-s_{male} A^2 X(1-X)}{3} \tag{31}$$

Given complete recessivity and synthetic male sterility, mutation-selection balance of sex-linked alleles occurs when:

$$\frac{\mu}{s_{male}} = \frac{A^2 X}{3} \tag{32}$$

Are there any conditions that satisfy mutation-selection balance for both autosomal and sex-linked alleles? Inspection of Equations 30 and 32 indicates that autosomal mutation-selection balance frequencies are smaller than sex-linked mutation-selection balance frequencies for all values A and X (see Figure 2C). Consistent with the evolutionary trajectories in Figure 5B, this indicates that X linked alleles will ultimately fix when male sterility is recessive.

Complete autosomal dominance

Analytic expressions can also be derived for mutation-selection balance of dominant synthetic male sterile alleles ($h_A = 0$). Mutation-selection balance conditions were calculated separately for autosomal and sex-linked alleles and plotted in Figure 2D. Because synthetic male sterility is observed at low frequencies when alleles are dominant, the equations in this subsection incorporate mean fitness terms.

The net change in autosome frequencies due to completely dominant synthetic male sterility is:

$$\Delta A_{sel} = \frac{-s_{male} A(1-A)^2 X}{2\overline{w}_{male}} \tag{33}$$

Given complete dominance and synthetic male sterility, mutation-selection balance of autosomal alleles occurs when:

$$\frac{\mu}{s_{male}} = \frac{A(1-A)X}{2\overline{w}_{male}} \tag{34}$$

The net change in sex-linked frequencies due to completely dominant synthetic male sterility is:

$$\Delta X_{sel} = \frac{-s_{male} AX\left(1 + X(A-2)\right)}{3\overline{w}_{male}} \tag{35}$$

Given complete dominance and synthetic male sterility, mutation-selection balance of sex-linked alleles occurs when:

$$\frac{\mu}{s_{male}} = \frac{AX\left(1 + X(A-2)\right)}{3(1-X)\overline{w}_{male}} \tag{36}$$

Synthetic female sterility

Under synthetic female sterility, selection only acts in females ($s_{male} = 0$). The relative strength of selection depends on autosomal and sex-linked allele frequencies and dominance coefficients at both loci. Because selection only occurs in diploid individuals under this scenario, X-autosome female sterility is similar to synthetic incompatibility involving a pair of autosomal loci. However, differences in evolutionary trajectories arise because selection only acts in a single sex. Two-thirds of the X chromosomes in a population are found in females, in contrast to one half the autosomal chromosomes. This suggests that allele frequency trajectories of autosomal and sex-linked alleles will differ when female synthetic sterility is present. For mathematical simplicity, mutation rates are assumed the same at both loci ($\mu_A = \mu_X = \mu$). Allele frequency trajectories for synthetic female sterile alleles are shown in Figure 6. Each panel in this figure was generated by numerical iteration of the equations for allele frequency change. For each set of dominance coefficients, populations were iterated 20,000 generations from eight different start states.

The dynamics of synthetic female sterile alleles are broadly similar to synthetic lethal alleles and male sterile alleles. Dominance at both loci affects the ultimate fate of populations containing synthetic female sterile alleles, and populations approach stable equilibria via fundamental trajectories. Autosomal dominance and X-linked recessivity results in a stable equilibrium with fixation of *A* alleles (Figure 6A). Complete dominance at both loci results in an unstable internal equilibrium and two stable external equilibria (Figure 6B). If autosomal alleles are recessive, stable equilibria involve fixation of X-linked factors (Figure 6C-D). One difference between female sterile alleles and synthetic lethal alleles is that complete recessivity at both loci (Figure 6C) results in autosomal, as opposed to X-linked, fixation. When present at stable stable equilibria, populations are found at single locus mutation-selection balance frequencies. Because selection only acts in one sex, segregating allele frequencies of female sterile alleles are slightly higher than expected under synthetic lethality. Once again, approaches to stable equilibria are slow.

Complete recessivity at both loci

Analytic expressions can also be derived for mutation-selection balance of recessive synthetic female sterile alleles ($h_A = h_X = 0$). Mutation-selection balance conditions were calculated separately for autosomal and sex-linked alleles and plotted in Figure 2E. The net change in autosome frequencies due to completely recessive synthetic female sterility is:

$$\Delta A_{sel} = \frac{-s_{fem} A^2 (1 - A) X^2}{2} \tag{37}$$

Given complete recessivity and synthetic female sterility, mutation-selection balance of autosomal alleles occurs when:

$$\frac{\mu}{s_{fem}} = \frac{A^2X^2}{2} \tag{38}$$

The net change in sex-linked frequencies due to completely recessive synthetic female sterility is:

$$\Delta X_{sel} = \frac{-2s_{fem}A^2X^2(1-X)}{3} \tag{39}$$

Given complete recessivity and synthetic female sterility, mutation-selection balance of sex-linked alleles occurs when:

$$\frac{\mu}{s_{fem}} = \frac{2A^2X^2}{3} \tag{40}$$

Are there any conditions that satisfy mutation-selection balance for both autosomal and sex-linked alleles? Inspection of Equations 38 and 40 indicates that autosomal mutation-selection balance frequencies are larger than sex-linked mutation-selection balance frequencies for all values A and X (see Figure 2C). Consistent with the evolutionary trajectories in Figure 6C, this indicates that autosomal alleles will ultimately fix when female sterility is recessive.

Complete dominance at both loci

Analytic expressions can also be derived for mutation-selection balance of dominant synthetic female sterile alleles ($h_A = h_X = 0$). Mutation-selection balance conditions were calculated separately for autosomal and sex-linked alleles and plotted in Figure 2F. Because synthetic female sterility is observed at low frequencies when alleles are dominant, the equations in this subsection incorporate mean fitness terms.

The net change in autosome frequencies due to completely dominant synthetic female sterility is:

$$\Delta A_{sel} = \frac{-s_{fem} A(1-A)^2 X(2-X)}{2\overline{w}_{fem}} \tag{41}$$

Given complete dominance and synthetic female sterility, mutation-selection balance of autosomal alleles occurs when:

$$\frac{\mu}{s_{fem}} = \frac{A(1-A)X(2-X)}{2\overline{w}_{fem}} \tag{42}$$

The net change in sex-linked frequencies due to completely dominant synthetic female sterility is:

$$\Delta X_{sel} = \frac{-2s_{fem} A(2-A)X(1-X)^2}{3\overline{w}_{fem}} \tag{43}$$

Given complete dominance and synthetic female sterility, mutation-selection balance of sex-linked alleles occurs when:

$$\frac{\mu}{s_{fem}} = \frac{2A(2-A)X(1-X)}{3\overline{w}_{fem}} \tag{44}$$

Divergent allopatric populations with secondary contact

The outcome of secondary contact between allopatric populations depends on whether allele frequencies differ between each population. However, the allele frequency trajectories of X-autosome synthetic incompatibilities suggest that each population will evolve along the same path. Subsequent peak shifts can still occur, but they are unlikely. This process depends on chromosome-specific effective population sizes, mutation-selection balance frequencies, and whether one or both sexes are affected (Whitlock and Wade 1995). How then might divergent populations initially evolve to have different allele frequencies?

One possibility involves genetic drift. In finite populations, one incompatible allele can drift to fixation before the incompatible allele at the other locus appears. This is expected to occur more often in small populations, and can result in different populations fixing different alleles. This is particularly interesting in the case of dominant X-autosome incompatibilities, where multiple equilibria exist. The relative proportion of populations fixing a particular allele depends on population size and mutation rates. In addition, the historical order of mutations can be important.

To determine the effects of finite population size (*N*), Monte Carlo simulations were run. These simulations were programmed in MATLAB (Mathworks 2005). Simulations were run for multiple population sizes and types of fitness dominance. For each set of parameters, simulations were run 1000 times. Each simulation run began with populations that were fixed for *a* and *x* alleles. After weighting by fitness, alleles were randomly sampled every generation. Each run was stopped after 100 or 1000 generations.

Data from finite population simulations mirror the results from analytic theory (Figure 7). Consider synthetic lethality where populations initially contained only *a* and *x* alleles and were allowed to evolve for 100 generations. Autosomal dominance yielded populations with higher frequencies of incompatible *A* alleles, and autosomal recessivity resulted in populations with higher frequencies of incompatible *X* alleles (compare Figures 3 and 7). However, the stochastic nature of finite populations allowed exceptions to the general pattern: some simulation runs yielded populations where the "unexpected" allele is found at a higher frequency than the "expected" allele. The degree of spread around the fundamental trajectory also varies from panel to panel. This is because synthetic lethal alleles are nearly neutral when recessive autosomal alleles are found at low frequencies, but not when X-linked recessives are found at low frequencies. The reason for this is that males are hemizygous for the X chromosome.

Population size has important implications for the evolutionary dynamics of X-autosome incompatibilities (Figure 8). Because of genetic drift, small populations have a larger spread around the fundamental trajectory than large populations. Small populations also have a smaller influx of new mutations and reduced efficacy of selection (Ohta 1973). This is shown in Figure 8: many of the low population size runs involved populations in which incompatible alleles segregate at only a single locus. When this occurs, alleles are effectively neutral and are able to

drift to fixation, even if they are the "unexpected" allele. Also, the speed at which populations increase to high allele frequencies can be much faster in small populations. This is because neutral fixation times are on the order of $4N$ generations (Charlesworth and Charlesworth 2010). The results in Figure 8 suggest that the "unexpected" allele can become fixed if the product of population size and mutation rate ($\theta = 4N\mu$) is less than one.

Alternatively, dominance modification may enable different populations to fix different alleles. If different genetic backgrounds cause synthetic lethal alleles to be autosomal dominant in one population and autosomal recessive in another population, the first population would be expected to fix the *A* allele and the second population would be expected to fix the *X* allele. Secondary contact would result in individuals containing deleterious combinations of alleles. The evolution of dominance has a long and contentious history (Bagheri 2006), and despite multiple examples in mammalian systems (Nadeau 2001), the full extent to which genetic background can modify the dominance of autosomal and sex-linked alleles is unknown. However, there is evidence that synthetic incompatibilities depend on more than just pairwise interactions (Cabot et al. 1994; Davis et al. 1994; Orr and Irving 2001).

The fitness costs of secondary contact depend on the underlying fitness landscape and the extent of allele frequency divergence. Also, reductions in the fitness of inter-population hybrids are not seen in the F1 generation if X-autosome interactions are recessive.

Synthetic load and sex-specific reductions in fitness

Populations containing synthetic incompatibilities have lower mean fitness than populations that lack these alleles. This is a form of genetic load. *Synthetic load* is defined here

as the amount that a population's fitness is below maximum due to synthetic incompatibilities. This term can be applied to individuals or populations. If synthetic incompatibilities involve X chromosomes and autosomes, synthetic load can be unequal for males and females.

Divergent allopatric populations are expected to lie on different parts of the same fundamental trajectory. Upon secondary contact, the two populations are combined and intermediate allele frequencies result. Consider a straight line connecting the allele frequencies of both source populations. The position of a merged population along this line depends on the relative size of each source population (Figure 9). Inspection of Figures 3, 5 and 6 indicates that fundamental trajectories are concave down regardless of the type of selection. In addition, the concavity of these trajectories does not depend on whether X-autosome synthetic incompatibilities are dominant or recessive. One implication of this is that the allele frequencies of merged allopatric populations will be above the fundamental trajectory. This means levels of synthetic load will be increased compared to population states prior to secondary contact.

The fact that males contain only a single X chromosome can be costly if synthetic lethal alleles are recessive. This is because the effects of deleterious X-linked alleles are not masked in hemizygous males. However, the opposite pattern is observed for dominant synthetic lethal alleles. In this scenario, each X chromosome that a female has can potentially contain a deleterious allele. Autosomal frequencies do not influence the relative proportion of synthetic load in males or females. Substituting the equations for sex-specific mean fitnesses allows an equation for the male proportion of synthetic load to be derived.

$$\frac{1-\overline{w}_{male}}{(1-\overline{w}_{male})+(1-\overline{w}_{fem})} = \frac{1}{1+\left(X+2h_X(1-X)\right)} \tag{45}$$

Solving for $h_X = 0$ and $h_X = 1$ gives the male proportion of synthetic load for recessive and dominant alleles:

$$\text{Proportion of synthetic load in males } (h_X = 0): \ = \frac{1}{1+X} \tag{46}$$

$$\text{Proportion of synthetic load in males } (h_X = 1): \ \frac{1}{3-X} \tag{47}$$

When dominance is intermediate on an additive scale ($h_A = 0.5$), synthetic load is equal for males and females (Figure 10). Recessive synthetic lethality results in an excess of male lethality, and dominant synthetic lethality results in an excess of female lethality. These effects are magnified when frequencies of the *X* allele are low. This reliance of sex-specific effects on dominance coefficients has been noted before in the Haldane's rule literature (Orr 1993; Turelli and Orr 1995; Turelli and Orr 2000). In recently merged populations, frequencies of *X* alleles are expected to drop as populations move towards the fundamental trajectory. Interestingly, this suggests that sex-specific differences in synthetic load can increase over time. Haldane's rule-like patterns may persist for multiple generations. However, as synthetic alleles decrease in frequency, so too do overall levels of synthetic load.

Summary

Comparisons with other types of incompatibilities

X-autosome incompatibilities segregate along a fundamental trajectory before ultimately fixing one locus, the identity of which is largely determined by the type of fitness dominance. Because autosomal and X-linked alleles compete with each other, the evolution of X-autosome systems involves a form of genomic conflict . As is the case for autosome-autosome synthetic incompatibilities, X-autosome incompatibilities are able to segregate at much higher frequencies than single locus expectations. Positions along fundamental trajectories where $A = X$ gives frequencies that are on the order of $\sqrt{\mu/s}$ when both incompatible alleles are dominant, $\sqrt[3]{\mu/s}$ when one incompatible allele is dominant, and $\sqrt[4]{\mu/s}$ when both incompatible alleles are recessive. However, the fundamental trajectories of X-autosome incompatibilities differ from those of autosome-autosome incompatibilities. Also, stable equilibria for X-autosome incompatibilities are biased towards fixation of alleles at X-linked or autosomal loci (with the particular locus depending on fitness dominance). This is in contrast to autosome-autosome theory, which has symmetric expectations so long as mutation rates at both loci are equal (Christiansen and Frydenberg 1977). One reason these patterns arise is because of differences in mutation pressure between X-linked and autosomal loci. Another reason these patterns arise is that X-linkage yields greater selection in males when alleles are recessive and reduced selection in males when alleles are dominant.

Synthetic lethality vs. synthetic sterility

Technically speaking, male sterility and female sterility scenarios also apply to situations where there is sex-specific lethality. However, there are biological reasons to assume that sex-specific incompatibilities will involve sterility rather than lethality. Epistatic interactions that reduce viability are likely to have the same effect in both sexes, while incompatibilities that

disrupt the reproductive system of one sex need not disrupt the reproductive system of the other sex.

The primary difference between synthetic lethality and synthetic sterility is that the strength of selection is stronger for the former, because both sexes are affected. This allows synthetic sterile alleles to segregate at higher frequencies. Another difference is that autosomal dominance coefficients affect all types of synthetic incompatibilities, while X-linked dominance coefficients affect only synthetic lethality and synthetic female sterility. In males, each selection event removes two autosomal alleles and one X-linked allele. In females, each selective event removes two autosomal alleles and two X-linked alleles. Because of this, the relative amount that autosomal and sex-linked alleles are purged depends on whether selection acts in males, females, or both sexes. In contrast to selection, mutation pressure is independent of the type of synthetic incompatibility. This is because female sterile mutations can arise in males, only to be passed onto their daughters (and vice-versa).

Cryptic incompatibilities

Stable equilibria involve fixation at one locus and low frequency variation at the other locus. Allele frequencies at the other locus follow single locus mutation-selection balance expectations. Synthetic incompatibilities may masquerade as single locus effects. This parallels the partitioning of variance in quantitative genetics: populations fixed for one locus will fail to show statistical epistasis even though there is functional epistasis (Templeton 2000).

An additional implication of this work is that cryptic incompatibilities are likely to be biased in a particular direction. If X-autosome incompatibilities tend to involve recessive alleles, the autosomal locus is expected to be polymorphic and the X-linked locus is expected to be fixed

for the *X* allele. By contrast, dominant X-autosome incompatibilities yield fixation of *A* alleles and polymorphism at the X-linked locus.

Secondary contact and Haldane's rule

If divergent allopatric populations have different frequencies of incompatible autosomal and X-linked alleles, secondary contact can result in populations with high levels of genetic load. This is most striking if one population is fixed for the *A* allele and the other population is fixed for the *X* allele. At least three mechanisms can enable this: genetic drift coupled with low mutation rates in small populations, population-specific dominance modification, and population-specific selection. Existing theoretical work has also shown that X-autosome incompatibilities can modify clinal patterns in hybrid zones (Wang and Zhao 2008).

Even prior to divergence, single populations can have asymmetric levels of reduced fitness in males and females. The driving issue here is whether X-linked factors are recessive or dominant. The extent to which male-biased synthetic incompatibilities are present depends on two elements: the overall amount of synthetic load (which is maximized when both alleles segregate at moderate frequencies) and the proportion of load that occurs in males (which is maximized when X-linked incompatibilities are rare and recessive). With X-autosome incompatibilities, Haldane's rule-like patterns can occur prior to speciation. Theory present in this paper lends support to both the faster-X and dominance hypotheses for Haldane's rule (Coyne and Orr 2004; Mank et al. 2010; Turelli and Orr 1995).

Conclusion

The overall impact of X-autosome interactions depends on the proportion of the genome that is X-linked. For example, ~18% of the *Drosophila melanogaster* genome is X-linked compared to only ~4-5% of the *Homo sapiens* genome (Adams et al. 2000; Lander et al. 2001; Venter et al. 2001). Also, future empirical studies examining how Haldane's rule scales with divergence time can be placed into the theoretical framework of this paper. Finally, there is increasing evidence that copy number variation is ubiquitous (Freeman et al. 2006), and many taxa have a haplodiploid sex-determination system. This suggests that future population genetic models may need to depart from complete diploidy or haploidy, of which this paper is one example. Insights gleaned from X-autosome interactions may be applicable to a much larger class of situations.

References

Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne, P. G. Amanatides, S. E. Scherer et al. 2000. The genome sequence of Drosophila melanogaster. Science 287:2185-2195.

Bagheri, H. C. 2006. Unresolved boundaries of evolutionary theory and the question of how inheritance systems evolve: 75 years of debate on the evolution of dominance. J Exp Zool B Mol Dev Evol 306:329-359.

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

Bengtsson, B. O., and F. B. Christiansen. 1983. A Two-Locus Mutation-Selection Model and Some of Its Evolutionary Implications. Theoretical Population Biology 24:59-77.

Cabot, E. L., A. W. Davis, N. A. Johnson, and C. I. Wu. 1994. Genetics of reproductive isolation in the Drosophila simulans clade: complex epistasis underlying hybrid male sterility. Genetics 137:175-189.

Charlesworth, B., and D. Charlesworth. 2010, Elements of Evolutionary Genetics. Greenwood Village, Colorado, Roberts and Company Publishers.

Charlesworth, B., J. Coyne, and N. H. Barton. 1987. The Relative Rates of Evolution of Sex Chromosomes and Autosomes. Am Nat 130:113-146.

Christiansen, F. B., and O. Frydenberg. 1977. Selection-mutation balance for two nonallelic recessives producing an inferior double homozygote. Am J Hum Genet 29:195-207.

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

Crow, J. F., and M. Kimura. 1970, An introduction to populaiton genetics theory. New York, Harper and Row.

Davierwala, A. P., J. Haynes, Z. Li, R. L. Brost, M. D. Robinson, L. Yu, S. Mnaimneh et al. 2005. The synthetic genetic interaction spectrum of essential genes. Nat Genet 37:1147-1152.

Davis, A. W., E. G. Noonburg, and C. I. Wu. 1994. Evidence for complex genic interactions between conspecific chromosomes underlying hybrid female sterility in the Drosophila simulans clade. Genetics 137:191-199.

de Visser, J. A., and S. F. Elena. 2007. The evolution of sex: empirical insights into the roles of epistasis and drift. Nat Rev Genet 8:139-149.

Dobzhansky, T. 1936. Studies on Hybrid Sterility. II. Localization of Sterility Factors in Drosophila Pseudoobscura Hybrids. Genetics 21:113-135.

—. 1937, Genetics and the Origin of Species. New York, Columbia University Press.

—. 1946. Genetics of Natural Populations. Xiii. Recombination and Variability in Populations of Drosophila Pseudoobscura. Genetics 31:269-290.

Freeman, J. L., G. H. Perry, L. Feuk, R. Redon, S. A. McCarroll, D. M. Altshuler, H. Aburatani et al. 2006. Copy number variation: new insights in genome diversity. Genome Res 16:949-961.

Gavrilets, S. 2004, Fitness Landscapes and the Origin of Species. Princeton, NJ, Princeton University Press.

Haldane, J. B. S. 1922. Sex ratio and unisexual sterility in hybrid animals. J. Genet. 12:101-109.

Hartman, J. L., B. Garvick, and L. Hartwell. 2001. Principles for the Buffering of Genetic Variation. Science 291:1001-1004.

Johnson, N. A. 2000. Speciation: Dobzhansky-Muller incompatibilities, dominance and gene interactions. Trends Ecol Evol 15:480-482.

—. 2010. Hybrid incompatibility genes: remnants of a genomic battlefield? Trends Genet 26:317-325.

Lachance, J., and J. R. True. 2010. X-autosome incompatibilities in *Drosophila melanogaster*: tests of Haldane's rule and geographic patterns within species. Evolution 64:3035-3046.

Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon et al. 2001. Initial sequencing and analysis of the human genome. Nature 409:860-921.

Laurie, C. C. 1997. The weaker sex is heterogametic: 75 years of Haldane's rule. Genetics 147:937-951.

Lehner, B., C. Crombie, J. Tischler, A. Fortunato, and A. G. Fraser. 2006. Systematic mapping of genetic interactions in Caenorhabditis elegans identifies common modifiers of diverse signaling pathways. Nat Genet 38:896-903.

Liao, C. Y., P. Wu, B. Hu, and K. K. Yi. 2001. Effects of genetic background and environment on QTLs and epistasis for rice (*Oryza sativa* L.) panicle number. Theor Appl Genet 103:104-111.

Lucchesi, J. C. 1968. Synthetic lethality and semi-lethality among functionally related mutants of Drosophila melanfgaster. Genetics 59:37-44.

Mackay, T. F. 2004. The genetic architecture of quantitative traits: lessons from Drosophila. Curr Opin Genet Dev 14:253-257.

Mank, J. E., B. Vicoso, S. Berlin, and B. Charlesworth. 2010. Effective population size and the faster-X effect: empirical results and their interpretation. Evolution 64:663-674.

Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff, D. J. Hunter, M. I. McCarthy et al. 2009. Finding the missing heritability of complex diseases. Nature 461:747-753.

Mathworks. 2005.MATLAB 7.The Mathworks, Natick, MA.

Moore, J. H. 2003. The ubiquitous nature of epistasis in determining susceptibility to common human diseases. Hum Hered 56:73-82.

Moore, J. H., F. W. Asselbergs, and S. M. Williams. 2010. Bioinformatics challenges for genome-wide association studies. Bioinformatics 26:445-455.

Muller, H. J. 1942. Isolating mechanisms, evolution and temperature. Biol. Symp. 6:71-125.

Nadeau, J. H. 2001. Modifier genes in mice and humans. Nat Rev Genet 2:165-174.

Ohta, T. 1973. Slightly deleterious mutant substitutions in evolution. Nature 246:96-98.

Ooi, S. L., X. Pan, B. D. Peyser, P. Ye, P. B. Meluh, D. S. Yuan, R. A. Irizarry et al. 2006. Global synthetic-lethality analysis and yeast functional profiling. Trends Genet 22:56-63.

Orr, H. A. 1993. A Mathematical Model of Haldane's Rule. Evolution 47:1606-1611.

—. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics 139:1805-1813.

Orr, H. A., and S. Irving. 2001. Complex epistasis and the genetic basis of hybrid sterility in the Drosophila pseudoobscura Bogota-USA hybridization. Genetics 158:1089-1100.

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

Otto, S. P., and T. Lenormand. 2002. Resolving the paradox of sex and recombination. Nat Rev Genet 3:252-261.

Phillips, P. C. 2008. Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet 9:855-867.

Phillips, P. C., and N. A. Johnson. 1998. The population genetics of synthetic lethals. Genetics 150:449-458.

Presgraves, D. C. 2003. A fine-scale genetic analysis of hybrid incompatibilities in Drosophila. Genetics 163:955-972.

—. 2010. The molecular evolutionary basis of species formation. Nat Rev Genet 11:175-180.

Sanjuan, R., A. Moya, and S. F. Elena. 2004. The contribution of epistasis to the architecture of fitness in an RNA virus. Proc Natl Acad Sci U S A 101:15376-15379.

Tao, Y., S. Chen, D. L. Hartl, and C. C. Laurie. 2003. Genetic dissection of hybrid incompatibilities between Drosophila simulans and D. mauritiana. I. Differential accumulation of hybrid male sterility effects on the X and autosomes. Genetics 164:1383-1397.

Temin, R. G., H. U. Meyer, P. S. Dawson, and J. F. Crow. 1969. The influence of epistasis on homozygous viability depression in Drosophila melanogaster. Genetics 61:497-519.

Templeton, A. R. 2000. Epistasis and Complex Traits *in* J. B. Wolf, E. D. Brodie, and M. J. Wade, eds. Epistasis and the Evolutionary Process. New York, Oxford University Press.

Thompson, V. 1986. Synthetic lethals: a critical review. Evolutionary Theory 8:1-13.

Tong, A. H., G. Lesage, G. D. Bader, H. Ding, H. Xu, X. Xin, J. Young et al. 2004. Global mapping of the yeast genetic interaction network. Science 303:808-813.

True, J. R., B. S. Weir, and C. C. Laurie. 1996. A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of Drosophila mauritiana chromosomes into Drosophila simulans. Genetics 142:819-837.

Turelli, M., and D. J. Begun. 1997. Haldane's rule and X-chromosome size in Drosophila. Genetics 147:1799-1815.

Turelli, M., and H. A. Orr. 1995. The dominance theory of Haldane's rule. Genetics 140:389-402.

—. 2000. Dominance, epistasis and the genetics of postzygotic isolation. Genetics 154:1663-1679.

Venter, J. C., M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith et al. 2001. The sequence of the human genome. Science 291:1304-1351.

Vicoso, B., and B. Charlesworth. 2006. Evolution on the X chromosome: unusual patterns and processes. Nat Rev Genet 7:645-653.

Wade, M. J. 2001. Epistasis, complex traits, and mapping genes. Genetica 112-113:59-69.

Wang, R. X., and Y. L. Zhao. 2008. Differential barrier strength and allele frequencies in hybrid zones maintained by sex-biased hybrid incompatibilities. Heredity 100:326-336.

Whitlock, M. C., and M. J. Wade. 1995. Speciation: founder events and their effects on X-linked and autosomal genes. Am Nat 145:676-685.

Wolf, J., E. Brodie, and M. J. Wade. 2000. Epistasis and the Evolutionary Process. New York, Oxford University Press.

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

Tables

Table 1. List of parameters for the X-autosome model

Parameter	Meaning
X_{male}	Frequency of X allele in males
X_{fem}	Frequency of X allele in females
A_{male}	Frequency of A allele in males
A_{fem}	Frequency of A allele in females
s_{male}	Selection coefficient in males
s_{fem}	Selection coefficient in females
h_X	Sex-linked dominance coefficient
h_A	Autosomal dominance coefficient
$\overline{w}_{male}$	Mean fitness of males
$\overline{w}_{fem}$	Mean fitness of males
μ_X	Mutation rate at the X-linked locus
μ_A	Mutation rate at the autosomal locus

Table 2. Fitness matrix

		Sex-chromosome genotype				
		XY	***xY***	***XX***	***Xx***	***xx***
Autosomal genotype	***AA***	$1 - s_{male}$	1	$1 - s_{fem}$	$1 - s_{fem}h_X$	1
	Aa	$1 - s_{male}h_A$	1	$1 - s_{fem}h_A$	$1 - s_{fem}h_Ah_X$	1
	aa	1	1	1	1	1

Table 2. Fitness matrix. Rows correspond to autosomal genotypes and columns correspond to sex chromosome genotypes. *A* and *X* are incompatible alleles. Selection coefficients in males and females are s_{male} and s_{fem}, respectively. Dominance terms of autosomal and sex-linked loci are h_A and h_X, respectively. Reduced fitness in double heterozygotes is assumed to be multiplicative.

Figure legends

Figure 1. Sex-specific differences in allele frequency. Complete recessivity at both loci is assumed and mutation rates are set equal to 0.0001. Initial conditions: $A_{male} = 1, A_{fem} = 0, X_{male} = 1, X_{fem} = 0$. Panels A,C,E, and G show sex-specific differences in autosomal frequencies. Panels B, D, F, and H show sex-specific differences in X-linked frequencies. Four selective scenarios were considered: no selection ($s_{male} = 0, s_{fem} = 0$, panels A and B), synthetic lethality ($s_{male} = 1, s_{fem} = 1$, panels C and D), synthetic male sterility ($s_{male} = 1, s_{fem} = 0$, panels E and F), and synthetic female sterility ($s_{male} = 0, s_{fem} = 1$, panels G and H).

Figure 2. Autosomal and X-linked mutation-selection balance. Allele frequency combinations where mutation-selection balance occurs for one locus need not be combinations where mutation-selection balance occurs for the other locus. Autosomal mutation-selection balance curves are plotted in black, and X-linked mutation-selection balance curves are plotted in gray. The position of a population genetic state relative to each of these curves determines the direction of allele frequency change. Each panel corresponds to different types of fitness dominance and selection: A) recessive synthetic lethality, B) dominant synthetic lethality, C) recessive synthetic male sterility, D) dominant synthetic male sterility, E) recessive synthetic female sterility, F) dominant synthetic female sterility. Selection coefficients were $s_{male} = 1, s_{fem} = 1$ for synthetic lethality, $s_{male} = 1, s_{fem} = 0$ for synthetic male sterility, and $s_{male} = 0, s_{fem} = 1$ for synthetic female sterility. In all panels, mutation rates at both loci were equal ($\mu_A = 0.01, \mu_X = 0.01$).

Figure 3. Allele frequency trajectories for synthetic lethality. Both sexes have the same selection coefficient ($s_{male} = 1, s_{fem} = 1$). Each panel depicts a different type of fitness dominance: A) autosomal dominance and X-linked recessivity ($h_A = 1, h_X = 0$), B) complete dominance at both loci ($h_A = 1, h_X = 1$), C) complete recessivity at both loci ($h_A = 0, h_X = 0$), and D) autosomal recessivity and X-linked dominance ($h_A = 0, h_X = 1$). To assist in viewing patterns, large mutation rates are used ($\mu_A = 0.01, \mu_X = 0.01$). For each set of dominance coefficients, populations were iterated 20,000 generations from eight different start states. Directions of allele frequency change are denoted with filled arrowheads. In panel B, a circle denotes the unstable internal equilibrium.

Figure 4. Different mutation rates and selection coefficients. Synthetic lethality, equal mutation rates, and complete recessivity at both loci were assumed ($s_{male} = 1, s_{fem} = 1, \mu_A = \mu_X, h_A = 0, h_X = 0$). Curves were generated by iterating the equations for allele frequency change for 20,000 generations. Start conditions: $A_{initial} = 0.9, X_{initial} = \mu / s$. In panel A, different mutation rates are compared (*s* held constant at 0.1). Higher mutation rates resulted in faster evolutionary change along the fundamental trajectory. In panel B, different selection coefficients were compared (μ held constant at 0.01). Gray curves depict $\mu / s = 0.1$ and black curves depict $\mu / s = 0.01$.

Figure 5. Allele frequency trajectories for synthetic male sterility. Selection coefficients differ for each sex ($s_{male} = 1, s_{fem} = 0$). In panel A, the *A* allele is dominant ($h_A = 1$). In panel B, the *A*

allele is recessive ($h_A = 0$). To assist in viewing patterns, large mutation rates are used ($\mu_A = 0.01$, $\mu_X = 0.01$). For each set of dominance coefficients, populations were iterated for 20,000 generations from eight different start states. Directions of allele frequency change are denoted with filled arrowheads.

Figure 6. Allele frequency trajectories for synthetic female sterility. Selection coefficients differ for each sex ($s_{male} = 0$, $s_{fem} = 1$). Each panel depicts a different type of fitness dominance: A) autosomal dominance and X-linked recessivity ($h_A = 1$, $h_X = 0$), B) complete dominance at both loci ($h_A = 1$, $h_X = 1$), C) complete recessivity at both loci ($h_A = 0$, $h_X = 0$), and D) autosomal recessivity and X-linked dominance ($h_A = 0$, $h_X = 1$). To assist in viewing patterns, large mutation rates are used ($\mu_A = 0.01$, $\mu_X = 0.01$). For each set of dominance coefficients, populations were iterated for 20,000 generations from eight different start states. Directions of allele frequency change are denoted with filled arrowheads. In panel B, a circle denotes the unstable internal equilibrium.

Figure 7. Synthetic lethality simulations. Data were generated via Monte Carlo simulations in MATLAB. Initial conditions involved populations without incompatible alleles. Data points indicate allele frequencies after 100 generations of evolution, and simulations were run 1000 times for each type of fitness dominance: A) complete autosomal dominance and X-linked recessivity, B) complete dominance at both loci, C) complete recessivity at both loci, and D) complete autosomal recessivity and X-linked dominance. For comparison with earlier figures, the following parameter values were used: $s_{male} = 1$, $s_{fem} = 1$, $\mu_A = 0.01$, $\mu_X = 0.01$, $N = 100$.

Figure 8. Simulations reveal population size effects. Monte Carlo simulations were run 1000 generations for population sizes of 100 (gray circles) and 1000 (black circles). Initial conditions involved populations without incompatible alleles. Simulations were run 1000 times for each population size. Parameter values used: $s_{male} = 1,\ s_{fem} = 1,\ \mu_A = 0.001,\ \mu_X = 0.001,\ h_A = 0,$ $h_X = 0$.

Figure 9. Secondary contact on a fitness landscape. Population mean fitness is represented by shading (with darker regions of allele frequency space corresponding to higher mean fitness). Synthetic lethal alleles are recessive at both loci. A solid black line denotes the fundamental trajectory. Divergent populations are envisioned as open circles (one population is fixed for the *A* allele, and the other population is fixed for the *X* autosome). Upon secondary contact, merged populations will initially lie along the dashed line. If population sizes are equal, the merged population will lie at the midpoint of the dashed line.

Figure 10. Male proportion of synthetic load for different levels of X-linked dominance. Synthetic refers to the amount that a population's fitness is below maximum due to synthetic incompatibilities. Three levels of X-linked dominance are considered: complete recessivity (gray line), intermediate dominance (dashed black line), and complete dominance (solid black line). Autosomal dominance does not influence the proportion of synthetic load that is in each sex.

Figures

Figure 1. Sex-specific differences in allele frequency

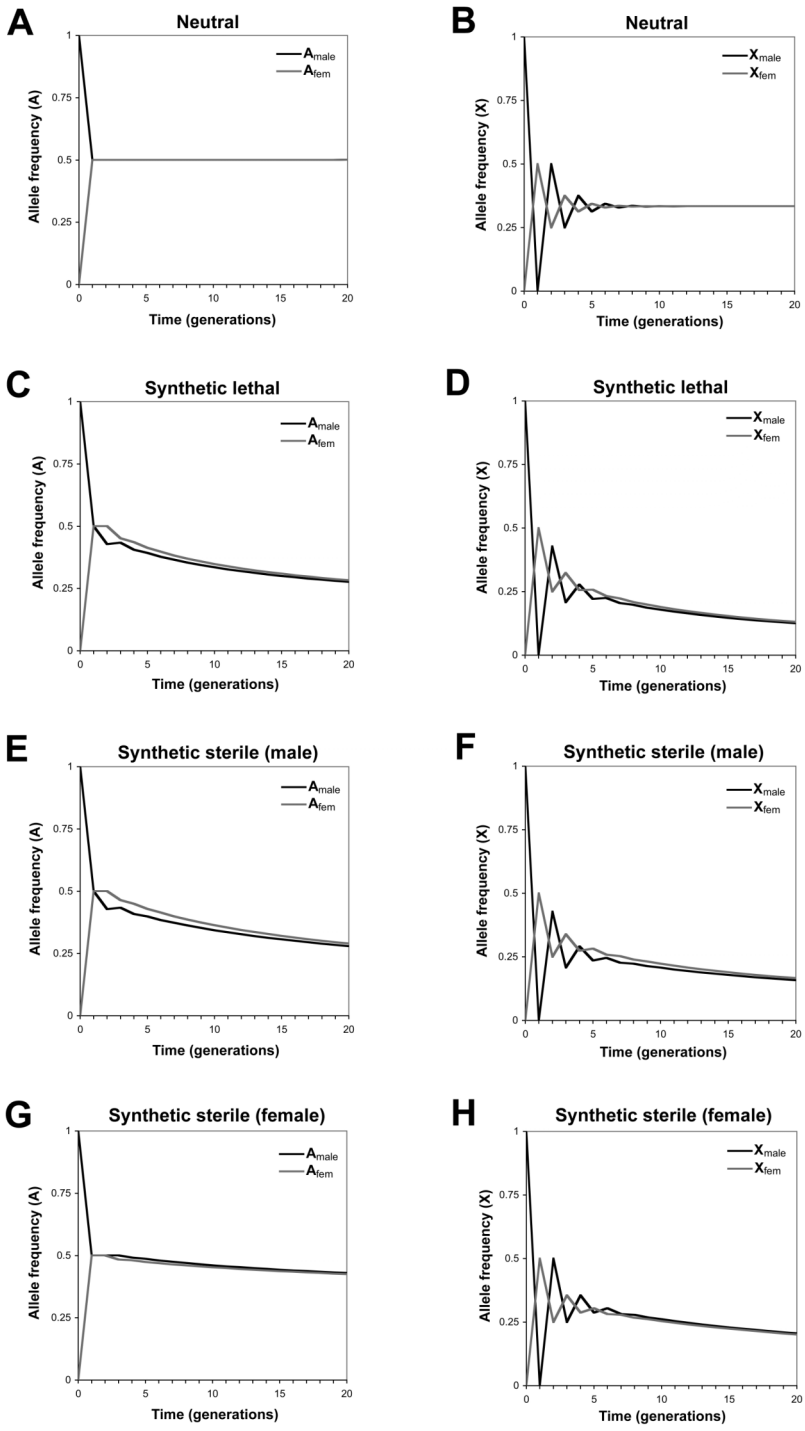


Figure 2. Mutation-selection balance

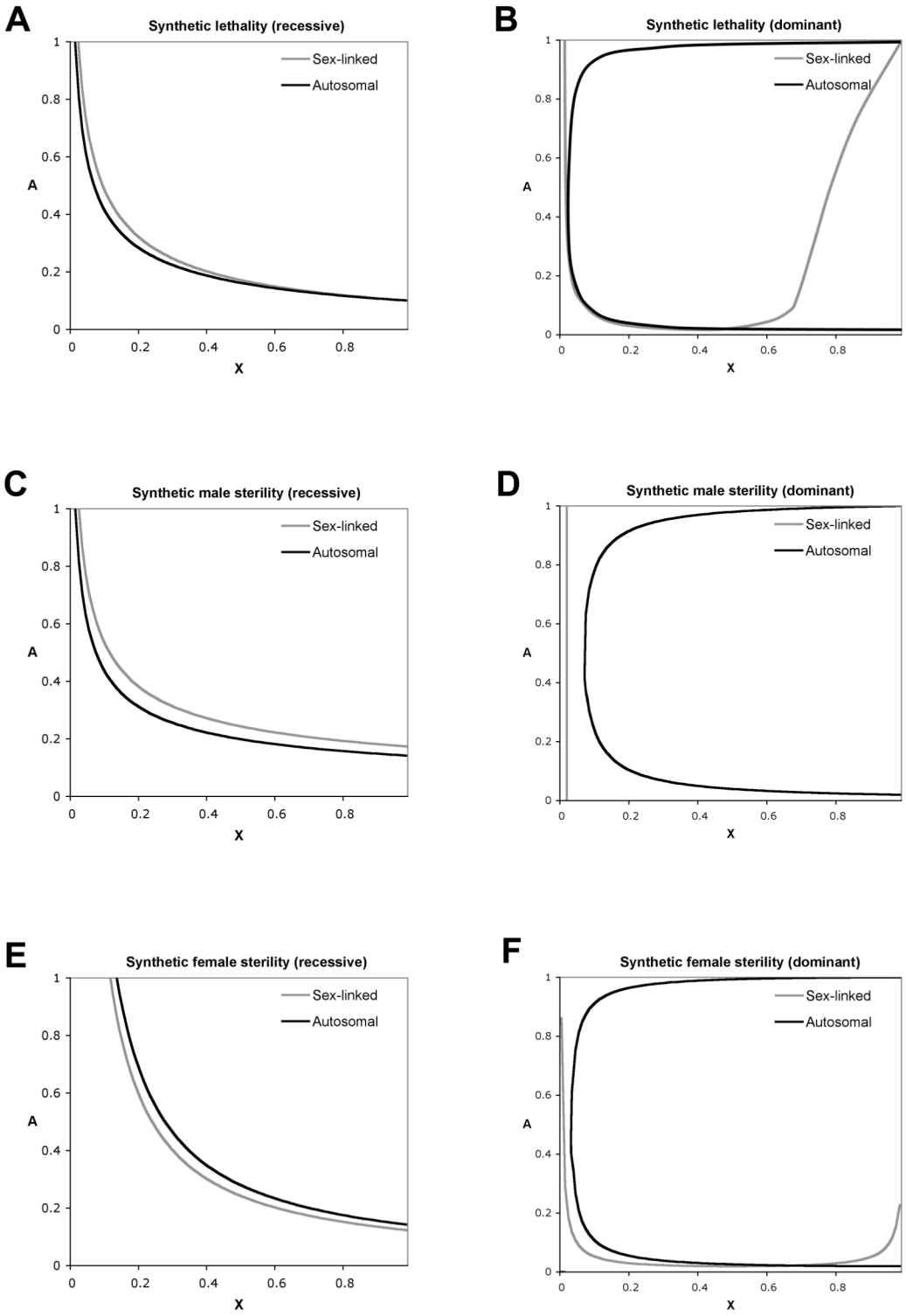


Figure 3. Allele frequency trajectories for synthetic lethality

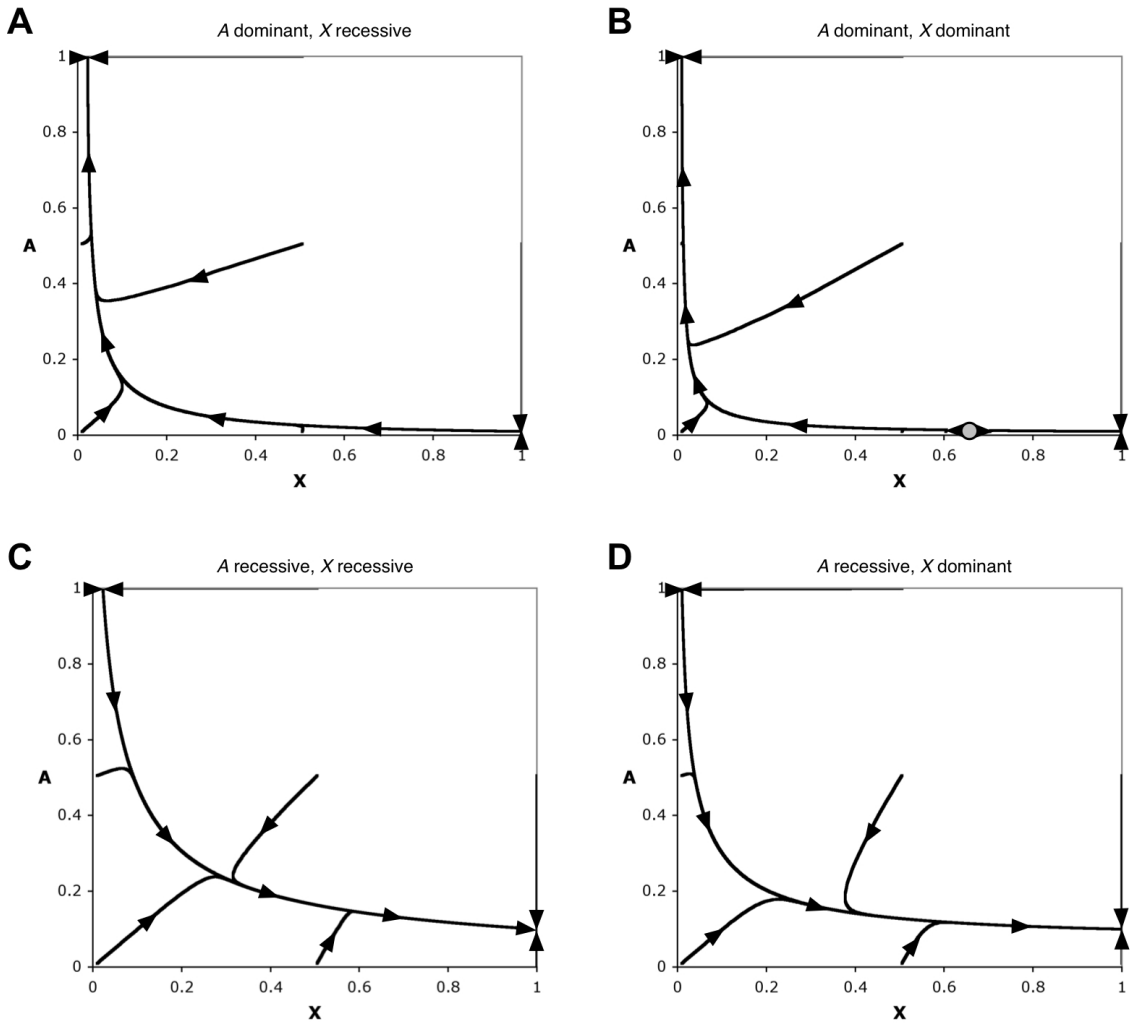


Figure 4. Different selection coefficients and mutation rates

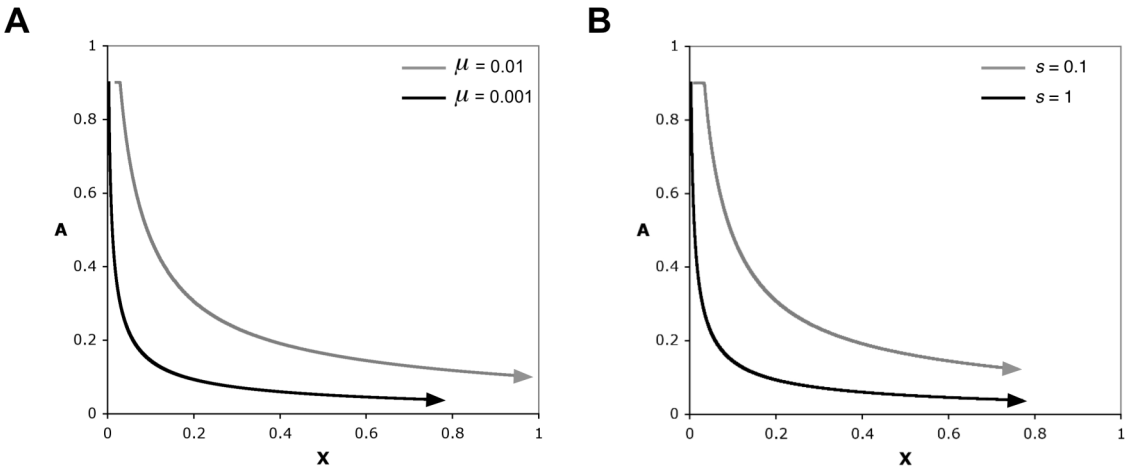


Figure 5. Allele frequency trajectories for synthetic male sterility

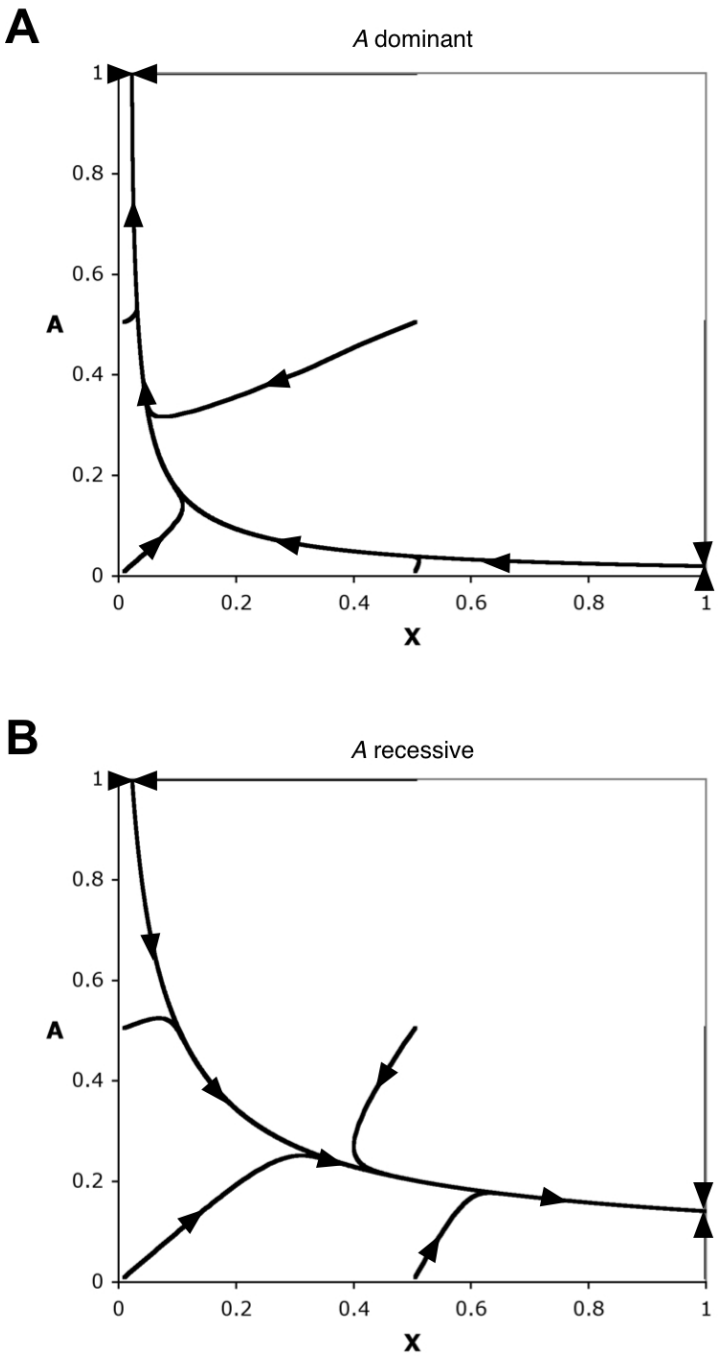


Figure 6. Allele frequency trajectories for synthetic female sterility

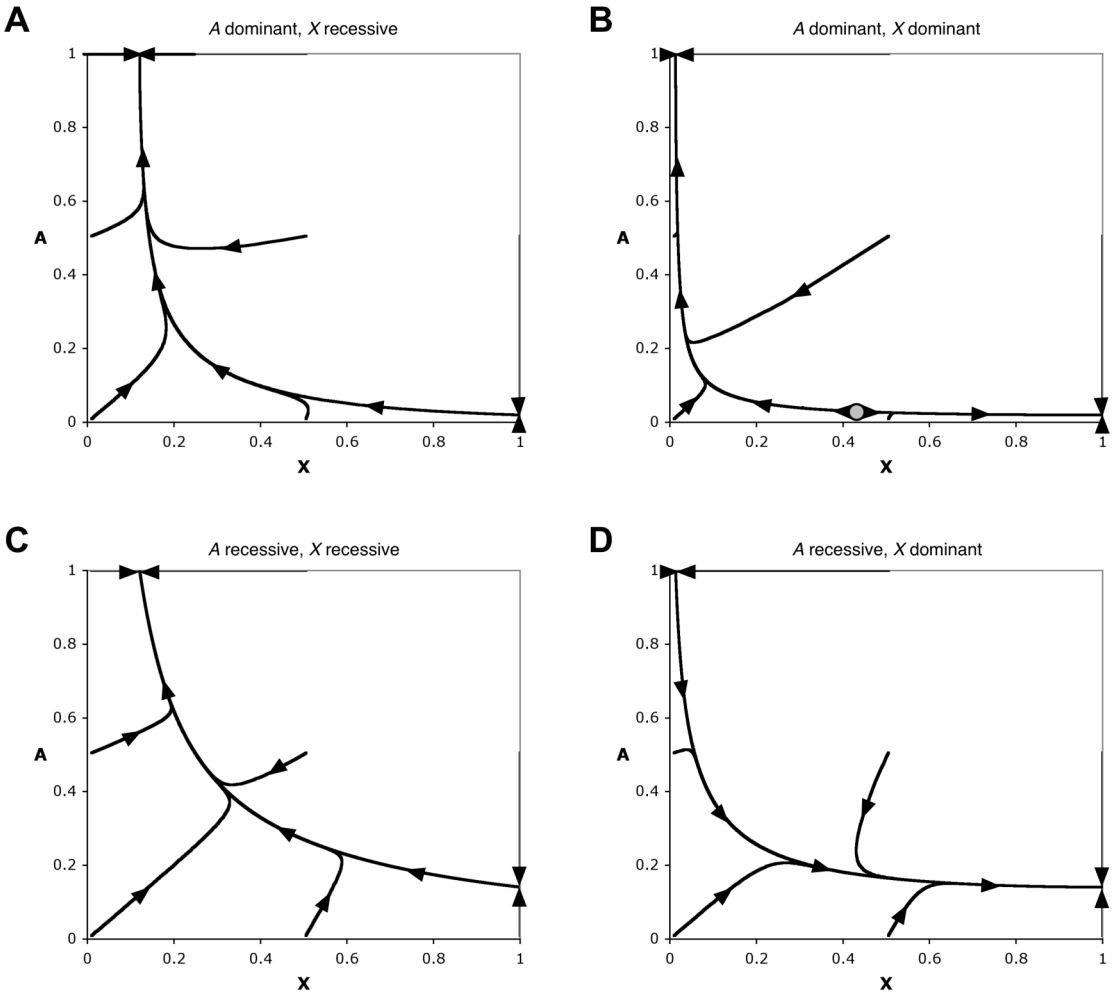


Figure 7. Synthetic lethality simulations

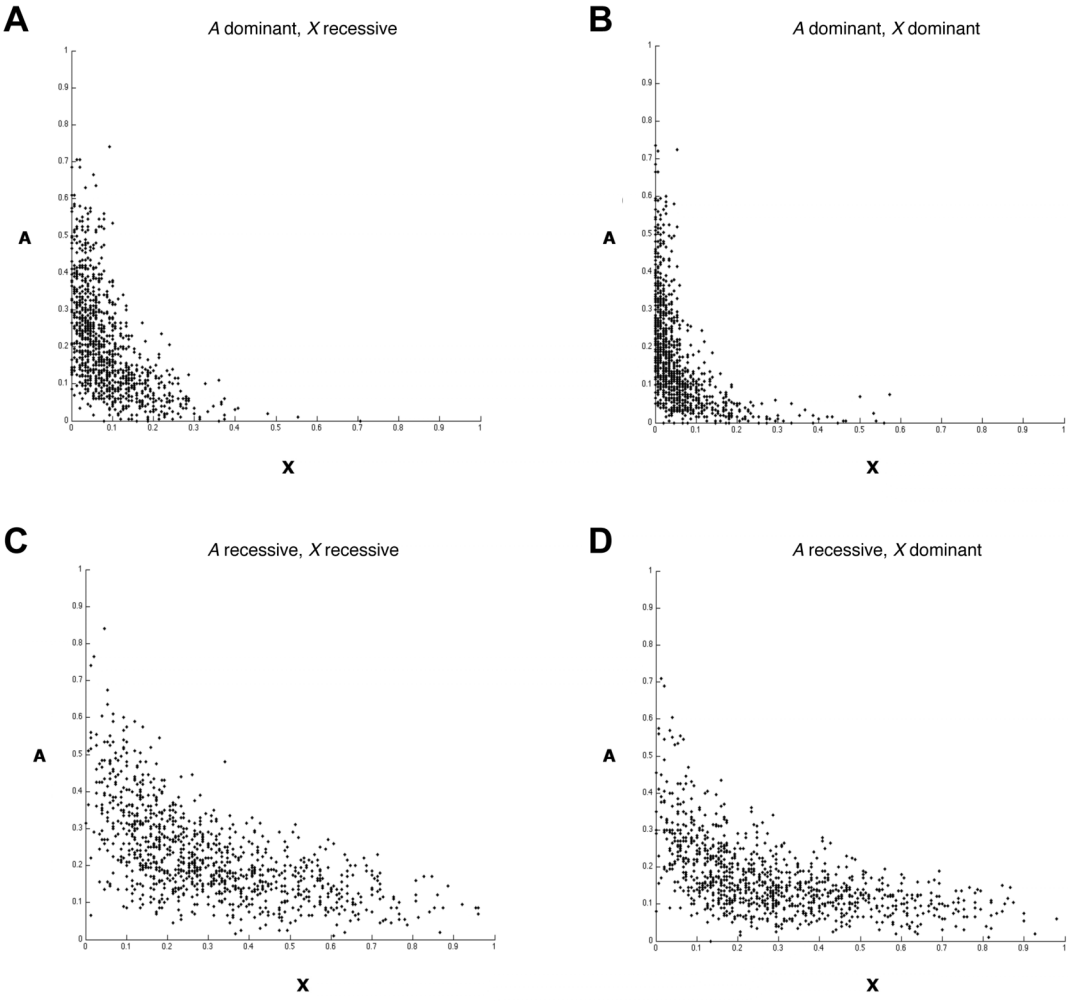


Figure 8. Simulations reveal population size effects

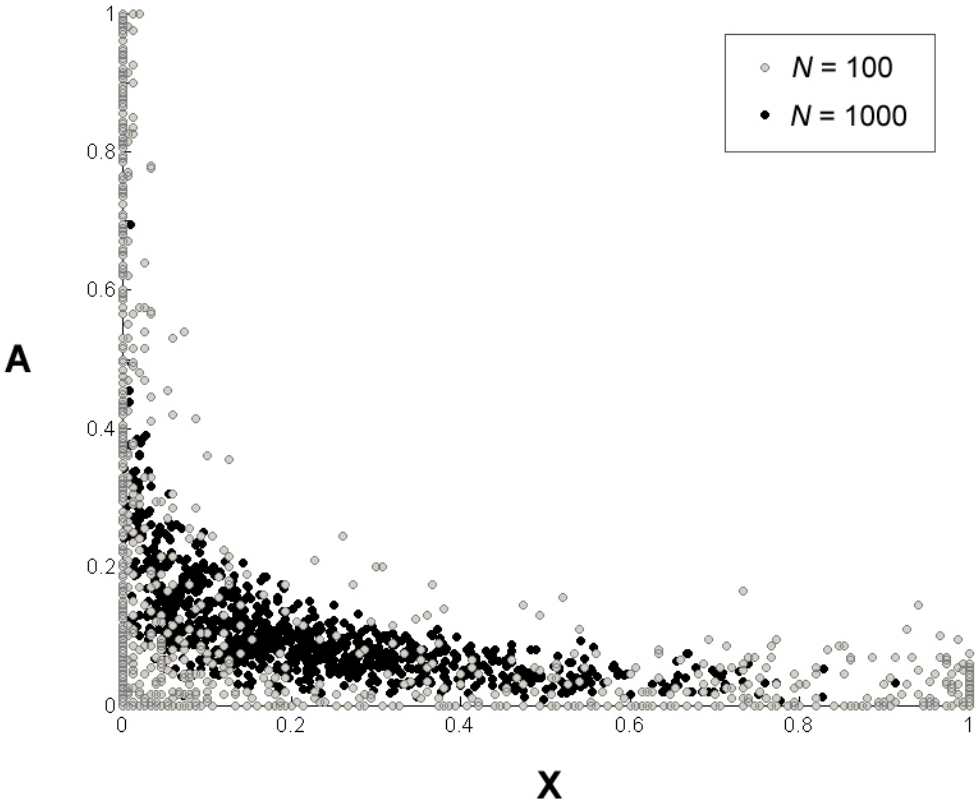


Figure 9. Secondary contact on a fitness landscape

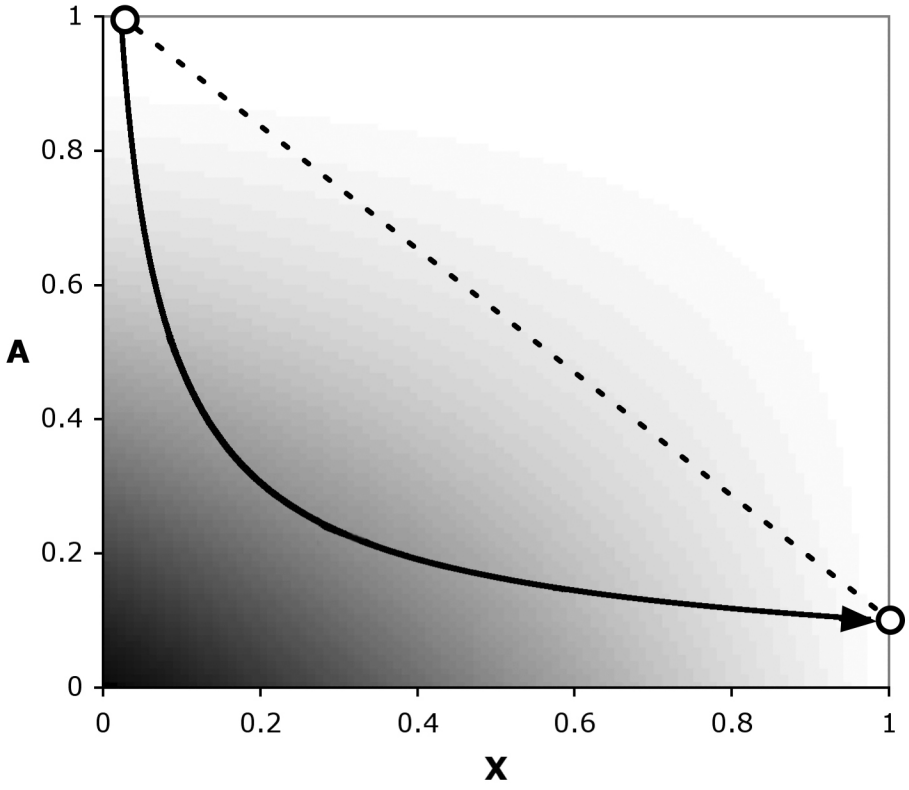
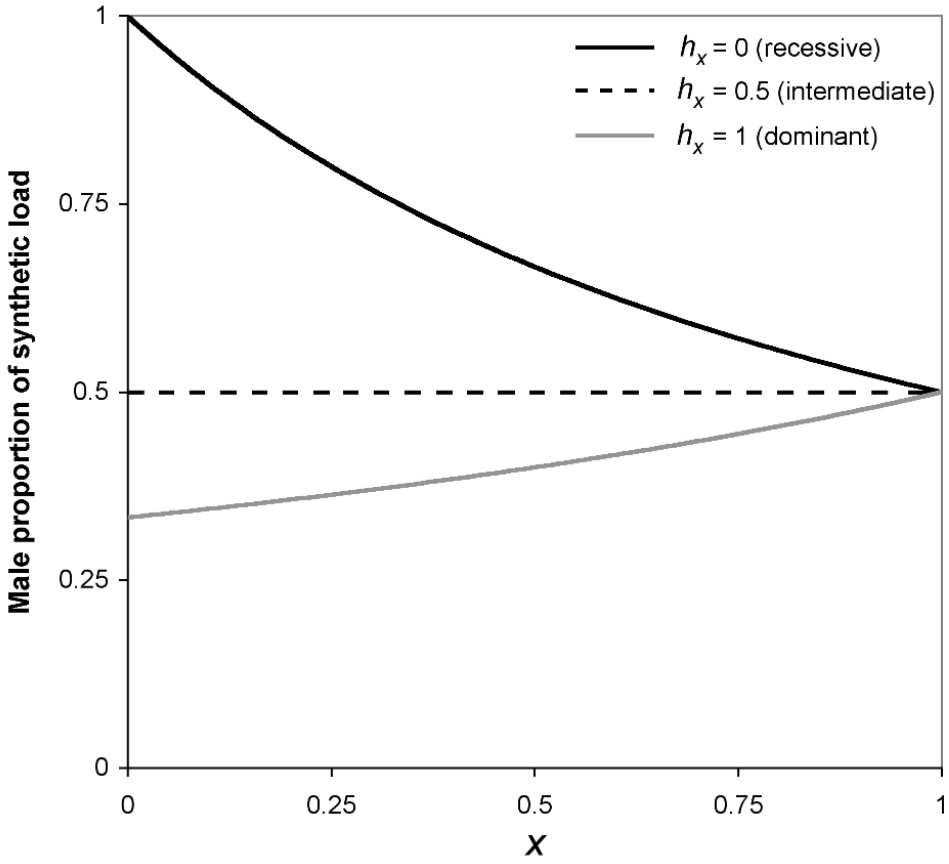


Figure 10. Male proportion of synthetic load for different levels of X-linked dominance



Chapter 4

Complex interactions underlie incompletely penetrant natural variation of the *vesiculated* locus in *Drosophila melanogaster*

Joseph Lachance, Lawrence Jung, and John R. True

Abstract

Many genes involved in producing complex traits are incompletely penetrant. One such example is *vesiculated*, an X-linked gene in *Drosophila melanogaster* that results in flies with wing defects. A naturally occurring X-linked variant was placed into multiple genetic backgrounds and grown at a range of developmental temperatures. Penetrance and expressivity of a wing vesicle phenotype varied across treatments, and the natural segregating variant was found to be allelic with *vesiculated*. Penetrance of this allele ranged from 0% to 79%, and wing phenotypes ranged from wild-type to vesicles encompassing an entire wing. We examined the genetic architecture of this complex trait (vesiculated wings), finding significant epistasis, genotype-by-environment interactions, and maternal effects. Sex and temperature effects were modulated by genetic background. *vesiculated* mutants acted as unbuffered genotypes, allowing the effects of cryptic genetic variation to be seen. The severity of wing phenotypes also varied across different genetic backgrounds and expressivity was positively correlated with penetrance. We found evidence of naturally segregating suppressors of *vesiculated*. These suppressors were present on both the 2nd and 3rd chromosomes, and complex interactions were observed. Penetrance was not an intrinsic property of *vesiculated* alleles. Instead, it depended on genetic background and other factors.

Introduction

The mapping of genotype to phenotype is central to developmental genetics and has important evolutionary consequences (Benfey and Mitchell-Olds 2008; Lunzer et al. 2005; Rockman 2008). Observed phenotypes depend on many factors, including environmental effects, genotype-by-environment (GxE) interactions, and epistatic interactions (Lewontin 2000). For example, cancer susceptibility depends upon genotype-by-environment interactions (Shields and Harris 2000), and epistatic interactions are known to occur between QTL for wing shape in *Drosophila melanogaster* (Mezey et al. 2005). One example of epistasis involves the appearance of suppressors, whereby the effects of an allele at one locus are masked by genetic variation at a second locus. In addition, the effects of genes are often modulated by genetic background, such as *Egfr* and *scalloped* in *D. melanogaster* (Dworkin et al. 2009; Polaczyk et al. 1998). Genotype-phenotype maps are also influenced by the positions of genes in developmental pathways (Stern 2010).

When genotypes are incompletely penetrant, one genotype is associated with two or more different phenotypes. The terms penetrance and expressivity were originally coined by the German neurologist Oskar Vogt after visiting Timofeeff –Ressovsky's fly lab in the mid-1920's (Laubichler and Sarkar 2002). A modern definition of penetrance is "the proportion of individuals of a specified genotype that show the expected phenotype under a defined set of environmental conditions" ((King et al. 2006), p. 327). Similarly, expressivity is defined as "the degree to which a novel phenotype is aberrant" ((Gibson and Dworkin 2004),p. 683). Incomplete penetrance can be viewed as a lack of developmental canalization. To avoid confusion with the concept of recessivity, incomplete penetrance in this paper is a property of genotypes rather than alleles. Many traits are associated with incompletely penetrant genotypes,

such as sterility due to the *Hybrid male rescue* gene in *Drosophila* (Aruna et al. 2009) and heart disease in humans (Pierpont et al. 2007). Penetrance can act as a nuisance parameter in human genetics, making it harder to detect associations in genome-wide association studies (Hirschhorn and Daly 2005). Multiple environmental and genetic causes underlie incomplete penetrance, such as thresholds in gene expression (Raj et al. 2010) and the presence of molecular chaperones (Carey et al. 2006). Penetrance can also reflect levels of genetic buffering (Gibson and Dworkin 2004). Although it is known that penetrance can be modified by environment and/or genetic background (Schmalhausen 1949), the relative importance of each of these factors and whether they interact is largely unknown. In addition, one can ask whether penetrance and expressivity are correlated. Do conditions that favor high penetrance also result in more severe phenotypes? In recent years there has been increased emphasis on the role of epigenetics (Bjornsson et al. 2004; Javierre et al. 2010; Youngson and Whitelaw 2008), and an open question is whether maternal or paternal effects influence the penetrance of alleles.

In a previous study, we placed a number of X chromosomes from natural *Drosophila melanogaster* populations into different autosomal backgrounds (Lachance and True 2010). One of these X chromosomes, *2214*, resulted in incompletely penetrant wing defects. These abnormal wing phenotypes only occurred when the *2214* X chromosome was placed into non-*2214* genetic backgrounds. Lines containing a *2214* X chromosome were otherwise healthy, and no other abnormal phenotypes were observed. After eclosion, the wings of flies unfold due to an increase in hemolymph pressure (Johnson and Milner 1987). Wing hearts (lateral muscular pumps located in the thorax) then function as suction pumps that remove hemolymph from newly unfolding wings (Togel et al. 2008). During wing maturation the wing cuticle delaminates, and components of the extracellular matrix are produced by wing epithelial cells

(Kiger et al. 2007). These components include position-specific integrins and other molecules that allow the dorsal and ventral surfaces of the wing to bond (Brabant et al. 1996; Brown et al. 2000). We observed clear bubbles of fluid in newly unfolded wings of flies containing *2214* X chromosomes. As these flies aged, these bubbles either flattened to become wrinkled wing-blades or they remained as vesicles.

vesiculated (*vs*) is an intriguing X-linked mutant that results in wing defects in *D. melanogaster*. It was originally discovered by O. L. Mohr 86 years ago (Evang 1925). *vs* mutants have wrinkled wings that contain fluid-filled blisters or vesicles and an early drawing of this mutant matches the phenotype observed in our *2214* X chromosome lines (Evang 1925). Also, the original description of *vesiculated* mentions that autosomal chromosomes are able to suppress the wing phenotype (Evang 1925). Waddington suggested that *vesiculated* probably "act(s) by causing breakage of some of the fibres [sic] which normally hold the two surfaces of the wing together during unfolding" (Waddington 1939). Existing *vs* mutants are known to display incomplete penetrance and variable phenotypes (Lindsley and Zimm 1992). Although the recombination and cytogenetic map positions of *vs* are known (1-16.3 and 6B2-6B3, respectively), it has yet to be mapped to the DNA sequence level (Judd et al. 1972; Tweedie et al. 2009). Thus, further studies of *vesiculated* mutants require a classical genetics approach.

In this study, we used complementation tests to determine the genetic basis of naturally segregating wing variants in *Drosophila melanogaster*. The penetrance of wing defects was assessed for flies with different genetic backgrounds at multiple temperatures. We also tested whether maternal and/or paternal effects modify penetrance. In addition, the severity of wing defects (expressivity) was quantified, as was the chromosomal basis of naturally segregating

suppressors. Together, these findings give insight into the genetic architecture of a complex trait.

Materials and methods

Stocks and construction of lines

X chromosomes were derived from wild caught and laboratory stocks of *D. melanogaster*. A wild caught X chromosome from Tuscaloosa, Alabama (vs^{2214}) resulted in abnormal wing phenotypes when in other genetic backgrounds, but not when it was in its natural genetic background, *2214*. This line was collected by R. Yukilevich in 2004. Two X-linked candidate loci were used in complementation tests: *vesiculated* (*vs*) and *inflated* (*if*). vs^{1} *and* if^{3} mutant lines were obtained from the Bloomington Stock Center (stocks 144 and 3960, respectively). BLAST searches of *CG34417* identified an actin-binding domain and sequence similarity to the human smooth muscle protein Smoothelin (Goldstein and Gunawardena 2000). This information and the cytogenetic map position of *CG34417* (6B3:6C1) suggests that this is a plausible candidate locus for *vesiculated*. $CG34417^{P\{GT1\}}$ is a P-element gene-disruption stock (as per (Lukacsovich et al. 2001)) that was obtained from the Bloomington Stock Center (stock 12670). An X chromosome deficiency line for the cytogenetic region 6B2:6C4, labeled *Df(1)Exel6240*, was also used in this paper. This region spans 125 kb and includes ten genes. *Df(1)Exel6240* was generated by Exelisis, Inc. and it also was obtained from the Bloomington Stock Center (stock 7714).

X-linked alleles were tested on multiple autosomal backgrounds: *6326*, *2214*, Rum Cay and Sudbury. *6326* was derived from a mapping line from the Bloomington Stock Center (stock

6326). *2214* and Rum Cay were collected by R. Yukilevich in 2004, and Sudbury was collected by T. Merritt in 2005. The *6326* and Sudbury lines were isogenized with balancers, whereas the Rum Cay line was produced by ten generations of sib-mating. Because of this, flies sharing the same autosomal background effectively had identical or nearly identical genomes. The balancer stock w^{1118};*T(2;3)ap*^{Xa}/*CyO:TM3* was used in the construction of lines and was obtained from the Bloomington Stock Center (stock 2475).

The *2114* X chromosome was placed into *6326*, Sudbury and Rum Cay genetic backgrounds using balancer chromosomes (Lachance and True 2010). Lines used in tests of chromosomal suppressors were constructed, resulting in autosomal backgrounds that were mixtures of *2214*, *6326*, and Sudbury chromosomes. Fourth chromosome effects were not tested. Flies were cultured on standard corn meal/molasses/agar medium supplemented with antibiotics (either penicillin at 40 μg/ml or a mix of tetracycline and streptomycin at 63 μg/ml and 19 μg/ml, respectively).

Complementation tests

The X-linked mutations used in this paper were recessive with respect to wing phenotype. This allowed us to perform complementation tests by generating heterozygous individuals and looking for the presence of wing vesicles. Wing vesicles present in heterozygous F1 flies indicated that mutations involved the same gene (allelism). Conversely, the presence of only wild-type wings indicated that each mutation involved a different gene. The *2214* X chromosome was tested for complementation with: *inflated* (*if*^{3}), *vesiculated* (*vs*^{1}), an X-linked deletion (*Df(1)Exel6240*), and a P-element disruption stock (*CG34417*$^{P\{GT1\}}$). The latter two lines were used to determine the approximate genomic region of the *vesiculated* gene. These

tests were done in multiple autosomal backgrounds (constructed via balancer chromosomes). Complementation tests of vs^{1} with *Df(1)Exel6240* and $CG34417^{P\{GT1\}}$ were also done. For each test, at least 50 F1 females were phenotyped.

Phenotypic assays

Wing phenotypes ranged from wild-type to wings with large vesicles encompassing an entire wing. Intermediate phenotypes involved a characteristic wing vesicle or blister (Figure 1). Because wings do not immediately unfold post-eclosion and wing vesicles fray after one to two weeks, flies were aged three to five days before being phenotyped. All flies tested in this section contained vs^{2214} X chromosomes. Both sexes were scored for multiple combinations of genetic background (*6326*, Sudbury, and Rum Cay) and developmental temperature (17.5ºC, 20ºC, 21.5ºC, and 25ºC). Flies were mass mated and there were at least six vials per combination of treatments. For each wing, phenotypes were scored on a zero to three scale: 0) wild-type, 1) small vesicle or wrinkled, 2) large vesicle spanning half the length of a wing, 3) vesicle encompassing entire wing giving a balloon-like appearance (Figure 4). The number of flies scored for each combination of treatments (sex, background, and temperature) ranged from 107 to 367. Statistical tests for single factors and pairwise interactions involved a three-factor ANOVA (fixed effects model). Calculations were done using MATLAB (Mathworks 2005). Because flies with Rum Cay autosomes could not be maintained at 17.5ºC, data for this temperature were omitted from ANOVA calculations.

It was previously shown that fluorescent lights can modify the penetrance and expressivity of *Curly* mutants (Pavelka et al. 1996). We examined whether this environmental factor had an effect on *vesiculated* mutants. All flies tested contained vs^{2214} X chromosomes.

Two autosomal backgrounds were tested: Sudbury and Rum Cay. Flies were grown at 25°C in an incubator and placed 15cm from a fluorescent light source. Light cycles were 12 h light:dark. Half of the vials were wrapped in index cards, resulting in dark conditions. After waiting four to nine days post-eclosion, wing phenotypes were assayed.

Maternal and paternal effects were tested by crossing parents with different wing phenotypes. Four different types of crosses were done: vesiculated mothers and vesiculated fathers, vesiculated mothers and wild-type fathers, wild-type mothers and vesiculated fathers, and wild-type mothers and wild-type fathers. Offspring of each cross were genetically identical, differing only in the penetrance of an abnormal wing phenotype. For each cross, there were six replicate vials (with three females and three males per vial). Genotypes tested were vs^{2214}; *6326*; *6326* and vs^{2214}; Sudbury; Sudbury. At least 370 F1 flies were phenotyped for each of these crosses. Developmental temperature for tests of maternal and paternal effects was 25ºC.

Interestingly, the *2214* X chromosome does not exhibit abnormal wings in its natural genetic background. To test whether naturally occurring suppressors were on the second or third chromosome we generated lines with mixed autosomal backgrounds. These genetic backgrounds contained a mix of *2214*, *6326*, and Sudbury autosomes. We also checked whether suppressors acted in a dominant or recessive fashion. Heterozygous suppressor genotypes were generated by crossing homozygous lines. A total of 18 different genotypes were tested (all possible autosomal combinations), and at least 55 flies of each sex were scored for each genotype. The 95% confidence intervals of proportions were calculated using the Agresti-Coull method (Agresti and Coull 1998).

Results

Natural segregating genetic variation is allelic with vesiculated and a 125kb region on the X chromosome

Complementation tests were performed for the *2214* X chromosome and the candidate genes *inflated* and *vesiculated* (Table 1). Both of these X-linked genes are known to result in incompletely penetrant wing defects (Lindsley and Zimm 1992). All F1 females heterozygous for the *2214* X chromosome and if^3 had wild-type wings. Conversely, a fraction of F1 females heterozygous for the *2214* X chromosome and vs^1 had wing vesicles. This failure to complement was observed in Sudbury and *6326* autosomal backgrounds. Because the mutation on the *2214* X chromosome was allelic with *vesiculated*, but not *inflated*, it was designated vs^{2214}.

Further complementation tests reiterated that vs^{2214} was allelic to vs^1 and indicated the approximate genomic region of the *vesiculated* gene. A previous study suggests that the *vesiculated* locus lies in the 6B2-6B3 cytogenetic region (Judd et al. 1972; Tweedie et al. 2009). Complementation tests of vs^1 and a deletion spanning this region (*Df(1)Exel6240*) resulted in flies with wing vesicles. Similarly, vs^{2214} showed a failure to complement the deletion construct in three different genetic backgrounds. Because vs^1 and vs^{2214} had similar results, this reinforces the idea that both of these X chromosomes involve mutations in the same gene. The $CG34417^{P\{GT1\}}$ X chromosome was able to complement both vs^{2214} and vs^1. This suggests that *CG34417* is not *vesiculated.* However, it is still possible that the P-element insertion in $CG34417^{P\{GT1\}}$ did not completely knock out this gene (especially since the insertion point is within an intron near the 3′ end of *CG34417*).

Penetrance varies by genetic background, sex, and temperature

We found that penetrance was not an intrinsic characteristic of alleles, as indicated in the reactions norms of Figure 2. Recall that all of the flies tested in this section have vs^{2214} X chromosomes, but they differ in their autosomal background, sex, and developmental temperature. Penetrance of vs^{2214} alleles ranged from 0% to 79%. A three-way ANOVA indicated significant effects of genetic background and temperature (Table 2). Penetrance was highest for flies with a Rum Cay autosomal background (43% to 79%), and lowest for flies with a *2214* autosomal background (0%). A general trend was that penetrance was greater for flies grown at higher temperatures. For example, the penetrance in *6326* and Sudbury backgrounds was approximately 2% higher for each additional degree Celsius. However, there were clear interactions between temperature and genetic background. While there was no general sex effect, there were interactions between background and sex. In particular, female flies with Rum Cay autosomes had higher penetrance than male flies, while male flies with *6326* or Sudbury autosomes had higher penetrance than female flies. These sex-specific differences were on the order of 4% and 12%, respectively. Therefore, accurate estimates of the penetrance of *vesiculated* mutants require knowledge of genetic background, sex, and temperature. When different phenotypic cutoffs were used (i.e. wings were required to have major defects), differences between genetic backgrounds were less noticeable (Figure 3).

Fluorescent lights moderately increased the penetrance of *vesiculated* mutants. Data from both sexes were pooled and the mean number of flies assayed per treatment was 158. Sudbury autosome-containing flies exposed to light had a mean penetrance of 8.1%, and flies kept in the dark had a mean penetrance of 4.5% (p-value = 0.270, two-tailed Fisher's exact test). Rum Cay autosome-containing flies exposed to light had a mean penetrance of 96.2%, and flies kept in the dark had a penetrance of 60.0% (p-value = 0.0391, two-tailed Fisher's exact test).

The mechanism causing these patterns is unknown, and the possibility of temperature differences between light and dark treatments cannot be ruled out.

Expressivity data

The severity of wing phenotypes was also measured and genetic background was found to have a large effect. Out of a grand total of 9,369 flies, the mean phenotypic score was 0.225 on a zero to three scale. Given the presence of a vesiculated wing, the mean phenotypic score was 1.792. Overall, 12.66% of flies had left wing defects and 12.40% had right wing defects. The mean phenotypic score for left wings was 0.228 and the mean score for right wings was 0.221. Given the presence of a vesiculated wing, the mean phenotypic score for left wing vesicles was 1.800 and the mean score for right wing vesicles was 1.785. Overall, there was no significant left-right asymmetry in the presence and magnitude of vesiculated wings (p-value > 0.5, two sample Z-test). Also, the probability that one wing was defective was not independent of the probability that the other wing was defective (p-value < 0.00001, χ^2 test of independence with 1 d.f.). We observed an overabundance of flies with both wings affected (4.49% compared to 1.57%, the product of left and right wing penetrance). This suggests that factors influencing the wing phenotypes of individual flies acted globally rather than locally.

Expressivity patterns are shown in Figure 4. For each treatment, the proportion of flies with a particular combination of left and right wing scores is indicated via shading. Mean phenotypic scores varied by genetic background: 0.000 for *2214*, 0.343 for *6326*, 0.216 for Sudbury, and 0.647 for Rum Cay. Most flies with Sudbury or *6326* autosomes were wild-type. By contrast, many flies with Rum Cay autosomes had wing vesicles, often in both wings. Penetrance and expressivity varied greatly across genetic backgrounds. To a lesser extent,

penetrance and expressivity also varied by temperature (contrast rows and columns in Figure 4). Sex differences in the severity of wing phenotypes were minimal. Figure 5 shows the proportion of wings with a particular non-zero phenotypic score for each set of treatments.

We also found that penetrance and expressivity were related. There was a positive correlation between the penetrance for a set of conditions (background, sex, and temperature) and the average phenotypic score of wings containing defects ($r = 0.2884$, p-value < 0.05, 2-tailed t-test). This pattern was stronger when the natural logarithm of penetrance was used ($r = 0.4873$, p-value < 0.001, 2-tailed t-test). Conditions that increased the probability of observing vesiculated wings also increased the severity of wing defects when they occurred. For example, flies raised at higher temperatures had both higher penetrance and more severe wing phenotypes.

Maternal and paternal effects

Flies were more likely to have wing vesicles if their parents also had wing vesicles. Pooling *6326* and Sudbury backgrounds, the penetrance of vs^{2214} flies with vesiculated mothers was 33.1%, and the penetrance of flies with wild-type mothers was 26.8% (p-value = 0.003, two sample Z-test). However, when individual autosomal backgrounds were considered, significant maternal effects were only observed for the *6326* genetic background (p-value = 0.026 for *6326*, p-value = 0.067 for Sudbury, two-sample Z-tests). Although slight differences were observed, there were no significant paternal effects. Pooling *6326* and Sudbury backgrounds, the penetrance of vs^{2214} flies with vesiculated fathers was 30.3%, and the penetrance of flies with wild-type fathers was 28.8% (p-value > 0.25, two sample Z-test). Thus, the overall trend was that maternal effects modified penetrance, and this pattern varied by genetic background.

Suppression of vesiculated involves complex epistasis

Suppressors of vs^{2214} were found on both autosomes. As indicated in Figure 6, flies with a *2214* autosomal background had wild-type wings. However, the presence of either *6326* or Sudbury autosomes resulted in flies with wing vesicles. While some sex differences were observed, penetrance was largely determined by autosomal background. Flies homozygous for the *2214* third chromosome had higher penetrance than flies homozygous for the *2214* second chromosome (Figure 6). This indicates that the *2214* second chromosome had a greater suppressive effect than the third chromosome. However, suppression of vs^{2214} by *2214* autosomes was not additive and complex patterns were observed (i.e. penetrance of vs^{2214} was not simply determined by the number of *2214* autosomes). In particular, flies with *2214* second chromosomes and Sudbury third chromosomes were more likely to have wing vesicles than flies with only Sudbury autosomes.

Suppression of vs^{2214} by *2214* autosomes was partially dominant (Figure 6). In most cases, heterozygous genotypes containing *2214* autosomes had low penetrance. For example, the penetrance of males heterozygous for *2214* and *6326* autosomes was closer to that of males homozygous for *2214* autosomes than males homozygous for 6326 autosomes (8.0% vs. 0.0% and 28.0%). This dominance pattern was observed for both sexes and two genetic backgrounds (6326 and Sudbury).

Discussion

A major finding of this study is that penetrance is not an intrinsic property of the *vesiculated* gene. This agrees with Schmalhausen's view that penetrance and expressivity are the

result of many environmental and genetic factors (Schmalhausen 1949). The presence of vesiculated wings was influenced by genetic background (GxG interactions), temperature (GxE interactions), and maternal effects. Importantly, temperature, sex, and maternal effects did not occur in isolation: they were modulated by genetic background. This underscores the importance of epistatic interactions, corroborating previous studies (Atallah et al. 2004; Carlborg and Haley 2004; Donehower et al. 1995; Dworkin et al. 2009; Phillips 2008). Our study suggests that the reductionist statements like "gene X does Y" need to be amended with descriptions of the genetic context of alleles.

The complex patterns observed in this paper allow some initial inferences about mechanism to be made. For example, we observed an overabundance of flies with vesicles on both wings, suggesting developmental stochasticity acted on an organismal scale. Autosomal suppression of *vesiculated* involved multiple autosomes and complex epistasis. This indicates the presence of multiple modifier genes. Although the identities of these modifiers are unknown, previous findings suggest that transcriptionally similar modules of genes can be associated with traits (Ayroles et al. 2009). This suggests that genes with modify the penetrance of *vesiculated* will have similar expression profiles. Despite evidence of epistatic interactions, *vesiculated* mutations did not appear to reduce egg-to-adult viability of stocks in a laboratory setting. However, fitnesses of these stocks were not explicitly assayed. It was also somewhat surprising that maternal effects influenced penetrance. Although levels of DNA methylation in *Drosophila melanogaster* are minimal, chromatin remodeling may be able to explain transgenerational epigenetic inheritance in this species (Ruden and Lu 2008).

Our findings are also relevant to the topics of robustness and genetic buffering. Robustness refers to the ability of organic systems to function in the face of perturbations, and

genotypes vary in their ability to buffer perturbations (de Visser et al. 2003; Kitano 2004; Masel and Siegal 2009; Wagner 2005). These perturbations are particularly important when thresholds exist, because continuous traits like gene expression can be converted into discrete phenotypes via development (Stern 2010). A study of intestinal cell fate in *Caenorhabditis elegans* revealed that incomplete penetrance can result from stochastic fluctuations in gene expression of unbuffered genotypes (Raj et al. 2010). Similarly, vs^{1} and vs^{2214} alleles can be viewed as less buffered than wild-type alleles at the *vesiculated* locus. The effects of different autosomes on the penetrance of wing vesicles were only observed in flies that had *vesiculated* mutations, indicating the presence of cryptic genetic variation (Gibson and Dworkin 2004; Hansen 2006).

What causes mutations of some genes (such as *vesiculated*) to be incompletely penetrant, while mutations of other genes are completely penetrant? In contrast to a wealth of knowledge about genetic dominance and recessivity (Kacser and Burns 1981; Kondrashov and Koonin 2004; Wilkie 1994), relatively little is known about the basis of incomplete penetrance. One consideration is developmental noise. Stochastic effects are important when phenotypes are determined by a small number of molecules or cells (Raser and O'Shea 2005). If developmental noise causes gene expression to span both sides of a threshold, incomplete penetrance can result. A second consideration involves developmental stage and reversibility. When a critical developmental time exists, the effects of genes are more likely to result in discrete phenotypes. These phenotypes can potentially be incompletely penetrant. Because wing unfolding only occurs once, chance events cannot be reversed. The position of a gene in a metabolic or developmental pathway may also affect whether mutations result in incomplete penetrance. There is evidence that genes at the center of hourglass-shaped pathways have large amounts of metabolic and/or developmental control (Kitano 2004; Stern 2010), and it is unknown whether

these genes are more or less likely to be incompletely penetrant. It has also been hypothesized that Mendelian traits are more likely to involve completely penetrant variants, while complex traits are more likely to involve low penetrance variants (Antonarakis et al. 2010). A large number of genes modify wing shape in *D. melanogaster* (Grieder et al. 2007; Zimmerman et al. 2000), and we found that complex interactions underlie the penetrance of *vesiculated*. However, it is unknown whether a general pattern exists in which mutant forms of highly epistatic genes are more likely to be incompletely penetrant than genes with few interactions. Penetrance and expressivity may also be related to pleiotropy. Non-permissive conditions (i.e. high penetrance conditions) for one trait may be conditions in which other associated traits manifest.

Importantly, the vs^{2214} allele and *2214* autosomal suppressors segregate in natural populations. This adds to a growing body of literature that indicates that standing genetic variation can modify the effects of genes (Barrett and Schluter 2008; Dworkin et al. 2003; Polaczyk et al. 1998; Wade et al. 1997). Although fitness effects in the wild are likely to be severe (flight appears to be difficult with wing vesicles), the natural vs^{2214} allele can segregate because of multiple reasons: it is recessive, it does not always manifest (effectively reducing any fitness reduction), it is masked by background, and phenotypic effects are modulated by temperature. An evolutionary implication is that suppressors facilitate the persistence of naturally segregating *vesiculated* alleles. The temperature effects we observed give further clues regarding natural variation. For example, vs^{2214} in a Sudbury background had lower penetrance at colder temperatures. If the vs^{2214} allele is viewed as a developmental perturbation, then the Sudbury (northern) background can be regarded as robust at lower temperatures. Conversely, flies with a Rum Cay (tropical) autosomal background were unable to be maintained at colder temperatures.

The DNA sequence and molecular basis of the *vesiculated* gene and its incomplete penetrance remain to be discovered. However, on a broader scale, general questions about penetrance can be asked in light of the *vesiculated* and other data. What is the penetrance of new mutations and does this vary throughout a genetic network? Is penetrance evolvable? The fact that genotype to phenotype mappings need not be one-to-one can also affect evolutionary trajectories (Weinreich et al. 2005), and theoretical models of incomplete penetrance are needed in future studies.

References

Agresti, A., and B. A. Coull. 1998. Approximate is better than "exact" for interval estimation of binomial proportions. American Statistician 52:119-126.

Antonarakis, S. E., A. Chakravarti, J. C. Cohen, and J. Hardy. 2010. Mendelian disorders and multifactorial traits: the big divide or one for all? Nat Rev Genet 11:380-384.

Aruna, S., H. A. Flores, and D. A. Barbash. 2009. Reduced fertility of Drosophila melanogaster hybrid male rescue (Hmr) mutant females is partially complemented by Hmr orthologs from sibling species. Genetics 181:1437-1450.

Atallah, J., I. Dworkin, U. Cheung, A. Greene, B. Ing, L. Leung, and E. Larsen. 2004. The environmental and genetic regulation of obake expressivity: morphogenetic fields as evolvable systems. Evol Dev 6:114-122.

Ayroles, J. F., M. A. Carbone, E. A. Stone, K. W. Jordan, R. F. Lyman, M. M. Magwire, S. M. Rollmann et al. 2009. Systems genetics of complex traits in Drosophila melanogaster. Nat Genet 41:299-307.

Barrett, R. D., and D. Schluter. 2008. Adaptation from standing genetic variation. Trends Ecol Evol 23:38-44.

Benfey, P. N., and T. Mitchell-Olds. 2008. From genotype to phenotype: systems biology meets natural variation. Science 320:495-497.

Bjornsson, H. T., M. D. Fallin, and A. P. Feinberg. 2004. An integrated epigenetic and genetic approach to common human disease. Trends Genet 20:350-358.

Brabant, M. C., D. Fristrom, T. A. Bunch, and D. L. Brower. 1996. Distinct spatial and temporal functions for PS integrins during Drosophila wing morphogenesis. Development 122:3307-3317.

Brown, N. H., S. L. Gregory, and M. D. Martin-Bermudo. 2000. Integrins as mediators of morphogenesis in Drosophila. Dev Biol 223:1-16.

Carey, C. C., K. F. Gorman, and S. Rutherford. 2006. Modularity and intrinsic evolvability of Hsp90-buffered change. PLoS One 1:e76.

Carlborg, O., and C. S. Haley. 2004. Epistasis: too often neglected in complex trait studies? Nat Rev Genet 5:618-625.

de Visser, J. A., J. Hermisson, G. P. Wagner, L. Ancel Meyers, H. Bagheri-Chaichian, J. L. Blanchard, L. Chao et al. 2003. Perspective: Evolution and detection of genetic robustness. Evolution 57:1959-1972.

Donehower, L. A., M. Harvey, H. Vogel, M. J. McArthur, C. A. Montgomery, Jr., S. H. Park, T. Thompson et al. 1995. Effects of genetic background on tumorigenesis in p53-deficient mice. Mol Carcinog 14:16-22.

Dworkin, I., E. Kennerly, D. Tack, J. Hutchinson, J. Brown, J. Mahaffey, and G. Gibson. 2009. Genomic consequences of background effects on scalloped mutant expressivity in the wing of Drosophila melanogaster. Genetics 181:1065-1076.

Dworkin, I., A. Palsson, K. Birdsall, and G. Gibson. 2003. Evidence that Egfr contributes to cryptic genetic variation for photoreceptor determination in natural populations of Drosophila melanogaster. Curr Biol 13:1888-1893.

Evang. 1925. The sex-linked mutants vesiculated and semi-lethal in Drosophila melanogaster. Z Indukt Abstamm Vererbungsl 39:165-183.

Gibson, G., and I. Dworkin. 2004. Uncovering cryptic genetic variation. Nat Rev Genet 5:681-690.

Goldstein, L. S., and S. Gunawardena. 2000. Flying through the drosophila cytoskeletal genome. J Cell Biol 150:F63-68.

Grieder, N. C., I. Charlafti, U. Kloter, H. Jackle, U. Schafer, and W. J. Gehring. 2007. Misexpression screen in Drosophila melanogaster aiming to reveal novel factors involved in formation of body parts. Genetics 175:1707-1718.

Hansen, T. 2006. The Evolution of Genetic Architecture. Annu Rev Ecol Syst 37:123-157.

Hirschhorn, J. N., and M. J. Daly. 2005. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 6:95-108.

Javierre, B. M., A. F. Fernandez, J. Richter, F. Al-Shahrour, J. I. Martin-Subero, J. Rodriguez-Ubreva, M. Berdasco et al. 2010. Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. Genome Res 20:170-179.

Johnson, S. A., and M. J. Milner. 1987. The final stages of wing development in Drosophila melanogaster. Tissue Cell 19:505-513.

Judd, B. H., M. W. Shen, and T. C. Kaufman. 1972. The anatomy and function of a segment of the X chromosome of Drosophila melanogaster. Genetics 71:139-156.

Kacser, H., and J. A. Burns. 1981. The Molecular Basis of Dominance. Genetics:639-666.

Kiger, J. A., Jr., J. E. Natzle, D. A. Kimbrell, M. R. Paddy, K. Kleinhesselink, and M. M. Green. 2007. Tissue remodeling during maturation of the Drosophila wing. Dev Biol 301:178-191.

King, R. C., W. D. Stansfield, and P. K. Mulligan. 2006, A dictionary of genetics. Oxford, Oxford University Press.

Kitano, H. 2004. Biological robustness. Nat Rev Genet 5:826-837.

Kondrashov, F. A., and E. V. Koonin. 2004. A common framework for understanding the origin of genetic dominance and evolutionary fates of gene duplications. Trends Genet 20:287-290.

Lachance, J., and J. R. True. 2010. X-autosome incompatibilities in *Drosophila melanogaster*: tests of Haldane's rule and geographic patterns within species. Evolution 64:3035-3046.

Laubichler, M. D., and S. Sarkar. 2002. Flies, genes, and brains: Oskar Vogt, Nikolai Timofeeff-Ressovsky, and the origin of the concepts of penetrance and expressivity *in* L. S. Parker, and R. A. Ankeny, eds. Mutating concepts, evolving disciplines: genes, medicine, and society. Philosophy and Medicine. Dordrecht, Kluvier.

Lewontin, R. C. 2000, The Triple Helix: Gene, Organism and Environment. Cambridge, Mass., Harvard University Press.

Lindsley, D. L., and G. G. Zimm. 1992, The Genome of *Drosophila melanogaster*. San Diego, CA, Academic Press.

Lukacsovich, T., Z. Asztalos, W. Awano, K. Baba, S. Kondo, S. Niwa, and D. Yamamoto. 2001. Dual-tagging gene trap of novel genes in Drosophila melanogaster. Genetics 157:727-742.

Lunzer, M., S. P. Miller, R. Felsheim, and A. M. Dean. 2005. The biochemical architecture of an ancient adaptive landscape. Science 310:499-501.

Masel, J., and M. L. Siegal. 2009. Robustness: mechanisms and consequences. Trends Genet 25:395-403.

Mathworks. 2005.MATLAB 7.Mathworks, Natick, MA.

Mezey, J. G., D. Houle, and S. V. Nuzhdin. 2005. Naturally segregating quantitative trait loci affecting wing shape of Drosophila melanogaster. Genetics 169:2101-2113.

Pavelka, J., A. M. Kulikov, and F. Marec. 1996. Suppression of the Drosophila Curly mutation by fluorescent light. Hereditas 124:191-197.

Phillips, P. C. 2008. Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet 9:855-867.

Pierpont, M. E., C. T. Basson, D. W. Benson, Jr., B. D. Gelb, T. M. Giglia, E. Goldmuntz, G. McGee et al. 2007. Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. Circulation 115:3015-3038.

Polaczyk, P. J., R. Gasperini, and G. Gibson. 1998. Naturally occurring genetic variation affects Drosophila photoreceptor determination. Dev Genes Evol 207:462-470.

Raj, A., S. A. Rifkin, E. Andersen, and A. van Oudenaarden. 2010. Variability in gene expression underlies incomplete penetrance. Nature 463:913-918.

Raser, J. M., and E. K. O'Shea. 2005. Noise in gene expression: origins, consequences, and control. Science 309:2010-2013.

Rockman, M. V. 2008. Reverse engineering the genotype-phenotype map with natural genetic variation. Nature 456:738-744.

Ruden, D. M., and X. Lu. 2008. Hsp90 affecting chromatin remodeling might explain transgenerational epigenetic inheritance in Drosophila. Curr Genomics 9:500-508.

Schmalhausen, I. I. 1949, Factors of Evolution. The theory of Stabilizing selection. Chicago, University of Chicago Press.

Shields, P. G., and C. C. Harris. 2000. Cancer risk and low-penetrance susceptibility genes in gene-environment interactions. J Clin Oncol 18:2309-2315.

Stern, D. L. 2010, Evolution, development, & the predictable genome. Greenwood Village, Colo., Roberts And Company Publishers.

Togel, M., G. Pass, and A. Paululat. 2008. The Drosophila wing hearts originate from pericardial cells and are essential for wing maturation. Dev Biol 318:29-37.

Tweedie, S., M. Ashburner, K. Falls, P. Leyland, P. McQuilton, S. Marygold, G. Millburn et al. 2009. FlyBase: enhancing Drosophila Gene Ontology annotations. Nucleic Acids Res 37:D555-559.

Waddington, C. H. 1939. Preliminary Notes on the Development of the Wings in Normal and Mutant Strains of Drosophila. Proc Natl Acad Sci U S A 25:299-307.

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.

Wagner, A. 2005, Robustness and Evolvability in Living Systems. Princeton, NJ, Princeton University Press.

Weinreich, D. M., R. A. Watson, and L. Chao. 2005. Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. Evolution 59:1165-1174.

Wilkie, A. O. 1994. The molecular basis of genetic dominance. J Med Genet 31:89-98.

Youngson, N. A., and E. Whitelaw. 2008. Transgenerational epigenetic effects. Annu Rev Genomics Hum Genet 9:233-257.

Zimmerman, E., A. Palsson, and G. Gibson. 2000. Quantitative trait loci affecting components of wing shape in Drosophila melanogaster. Genetics 155:671-683.

Table 1. Complementation tests

X chromosome genotype	Autosomal background	Wing phenotype
vs^{2214} / if^{3}	$+^{Sudbury}$	wild-type
vs^{2214} / if^{3}	$+^{6326}$	wild-type
vs^{2214} / vs^{1}	$+^{Sudbury}$	vesiculated
vs^{2214} / vs^{1}	$+^{6326}$	vesiculated
vs^{1} / *Df(1)Exel6240*	$+^{144}$ / $+^{7714}$	vesiculated
vs^{2214} / *Df(1)Exel6240*	$+^{Sudbury}$ / $+^{7714}$	vesiculated
vs^{2214} / *Df(1)Exel6240*	$+^{6326}$ / $+^{7714}$	vesiculated
vs^{2214} / *Df(1)Exel6240*	*CyO* / $+^{7714}$; *TM3* / $+^{7714}$	vesiculated and curly
vs^{1} / $CG34417^{P\{GT1\}}$	$+^{144}$ / $+^{12670}$	wild-type
vs^{2214} / $CG34417^{P\{GT1\}}$	$+^{Sudbury}$ / $+^{12670}$	wild-type
vs^{2214} / $CG34417^{P\{GT1\}}$	$+^{6326}$ / $+^{12670}$	wild-type
vs^{2214} / $CG34417^{P\{GT1\}}$	*CyO* / $+^{12670}$; *TM3* / $+^{12670}$	curly

Table 1. Complementation tests. Heterozygous female flies were assayed for wing defects. Superscripts of autosomal backgrounds indicate the Bloomington stock number or wild-caught origin of autosomes. These tests indicate that vs^{2214} is allelic with vs^{1} and *Df(1)Exel6240*.

Table 2. Three-factor ANOVA

Source	Sum Sq.	d.f.	Mean Sq.	F	p-value
Background	10681.2	3	3560.39	197.26	< 0.0001
Sex	63.8	1	63.77	3.53	0.1092
Temperature	617.4	2	308.72	17.1	0.0033
Background*Sex	537.3	3	179.09	9.92	0.0097
Background*Temperature	723.9	6	120.66	6.68	0.018
Sex*Temperature	8.7	2	4.36	0.24	0.7926
Error	108.3	6	18.05		
Total	12740.6	23			

Table 2. Three-factor ANOVA. The effects of genetic background, sex, and developmental temperature were tested, as were pairwise interactions. Data tested are shown in Figure 2. Note that 17.5ºC data was omitted from this test. Significant (p-value < 0.05) effects were observed for background, temperature, background*sex, and background*temperature.

Figure legends

Figure 1. Example of a wing vesicle phenotype. The phenotypic score for this wing was one on a zero to three scale. The full phenotypic range of *vesiculated* mutations includes wild-type wings, wrinkled wings, small vesicles, and balloon-shaped wings.

Figure 2. Penetrance reaction norms. The Y-axis indicates the proportion of flies containing vesiculated wings. All flies tested contain X chromosomes with a vs^{2214} genotype. Shading indicates the autosomal background (Rum Cay: black, *6326*: dark gray, Sudbury light gray, *2214*: white). Males are represented with squares and females with circles. Penetrance varies by autosomal background, sex, and developmental temperature. Note that Rum Cay flies were unable to be maintained at 17.5ºC.

Figure 3. Penetrance reaction norms with different phenotypic cutoffs. Phenotypic cutoffs differed for each panel. A) Wings defects were considered to be penetrant if a vesicle spanned at least half a wing (i.e. the wing score was two or three). B) Wing defects were considered to be penetrant only if a vesicle encompassed an entire wing (i.e. the wing score was three).

Figure 4. Expressivity of vesiculated wings in multiple genetic backgrounds. All flies tested harbored vs^{2214} X chromosomes. The severity of wing phenotypes varied by treatment (autosomal backgrounds, temperature, and sex). Scores ranged from 0 (wild-type) to 3 (balloon-

like wings). For each treatment, shading in each cell of a 4x4 grid indicates the proportion of flies with a particular combination of left and right wing scores. Shading ranges from white (0%) to black (>10%). For scores >10%, the percentage is also listed. Panel A shows expressivity data for males and panel B shows expressivity data for females. Median number of flies per combination of treatments is 213.

Figure 5. Expressivity histograms. For each set of treatments (temperature and autosomal background), the proportion of wings with a particular non-zero phenotypic score is quantified. Expressivity data for males is represented in blue and expressivity data for females is represented in red.

Figure 6. Suppression of vs^{2214} by *2214* autosomes. Proportions of flies that had wing defects at 20ºC are depicted. Different chromosomes are represented by shading: *2214* in white, *6326* in gray, and Sudbury in black. Error bars indicate 95% confidence intervals.

Figures

Figure 1. Example of a wing vesicle phenotype

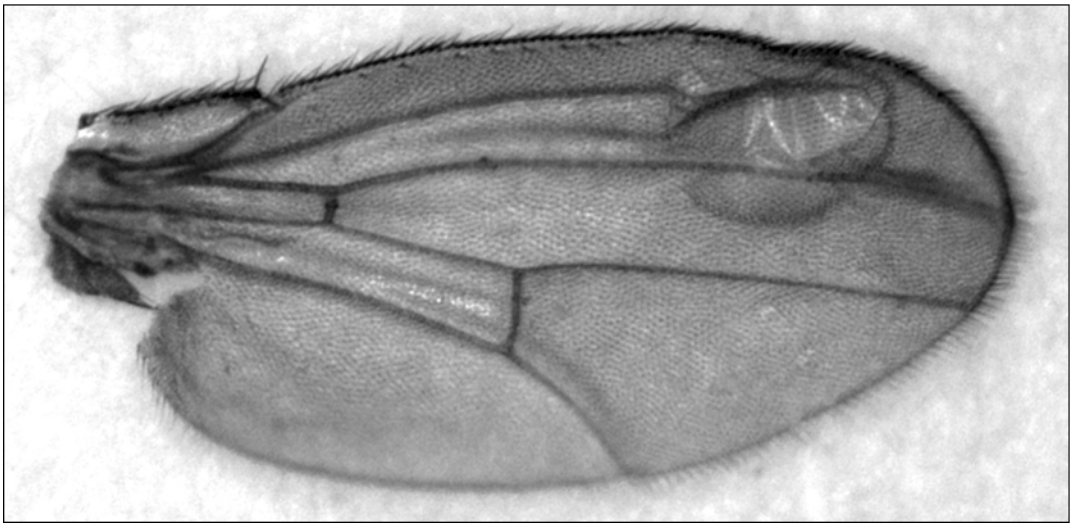


Figure 2. Penetrance reaction norms

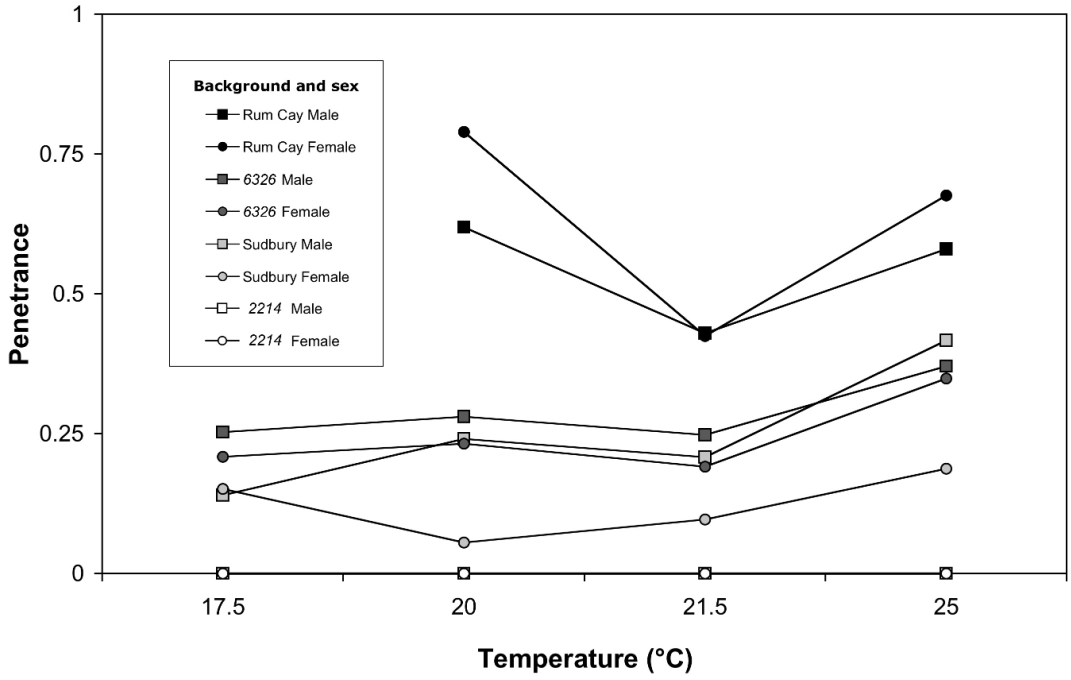


Figure 3. Penetrance reaction norms with different phenotypic cutoffs

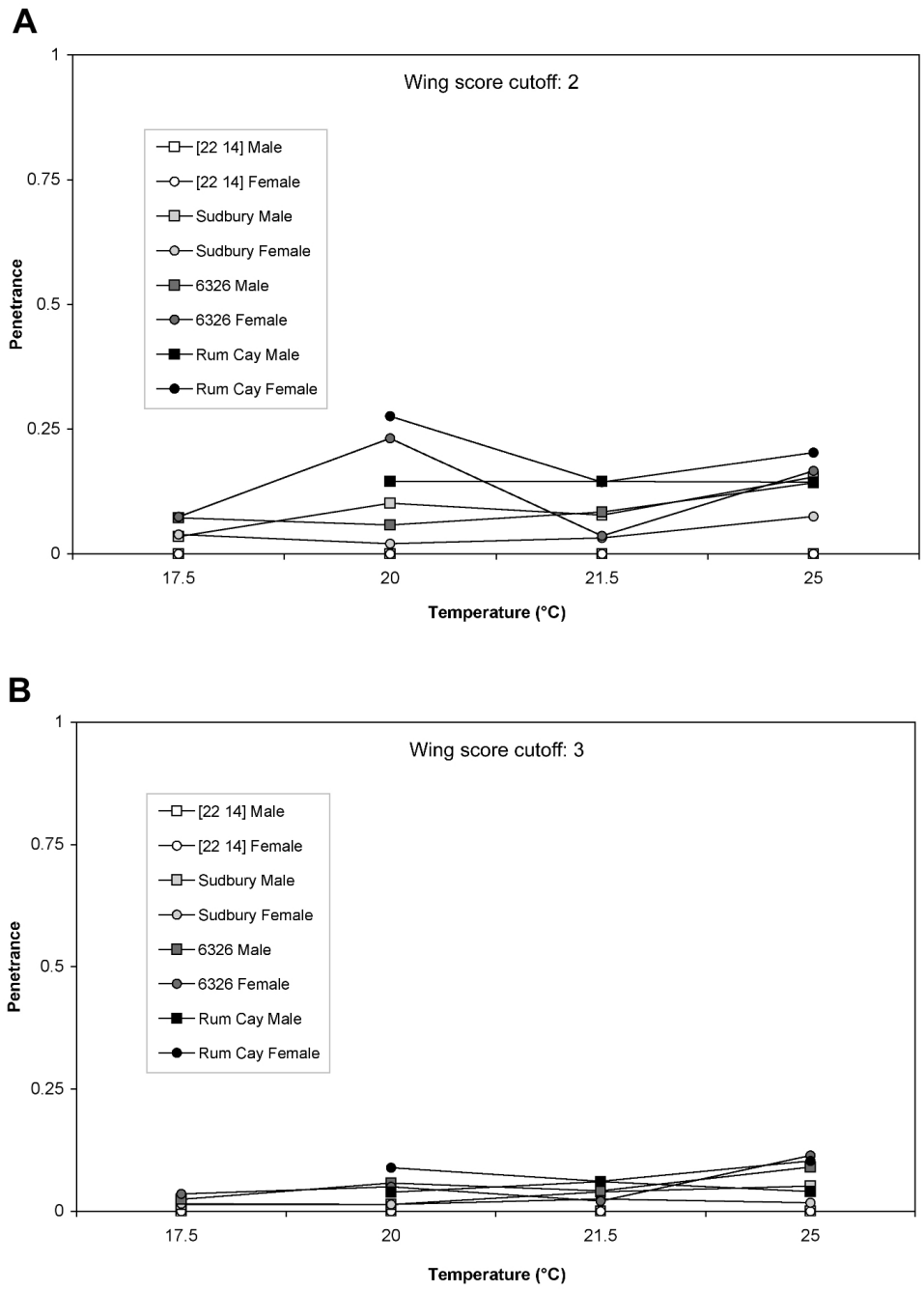


Figure 4. Expressivity of vesiculated wings in multiple genetic backgrounds

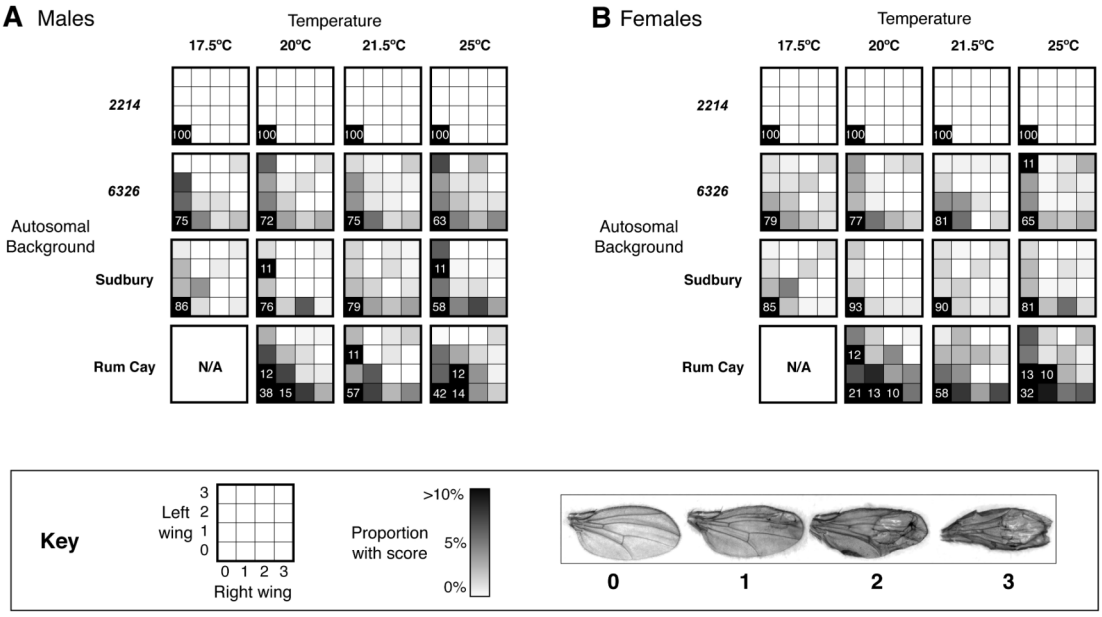


Figure 5. Expressivity histograms

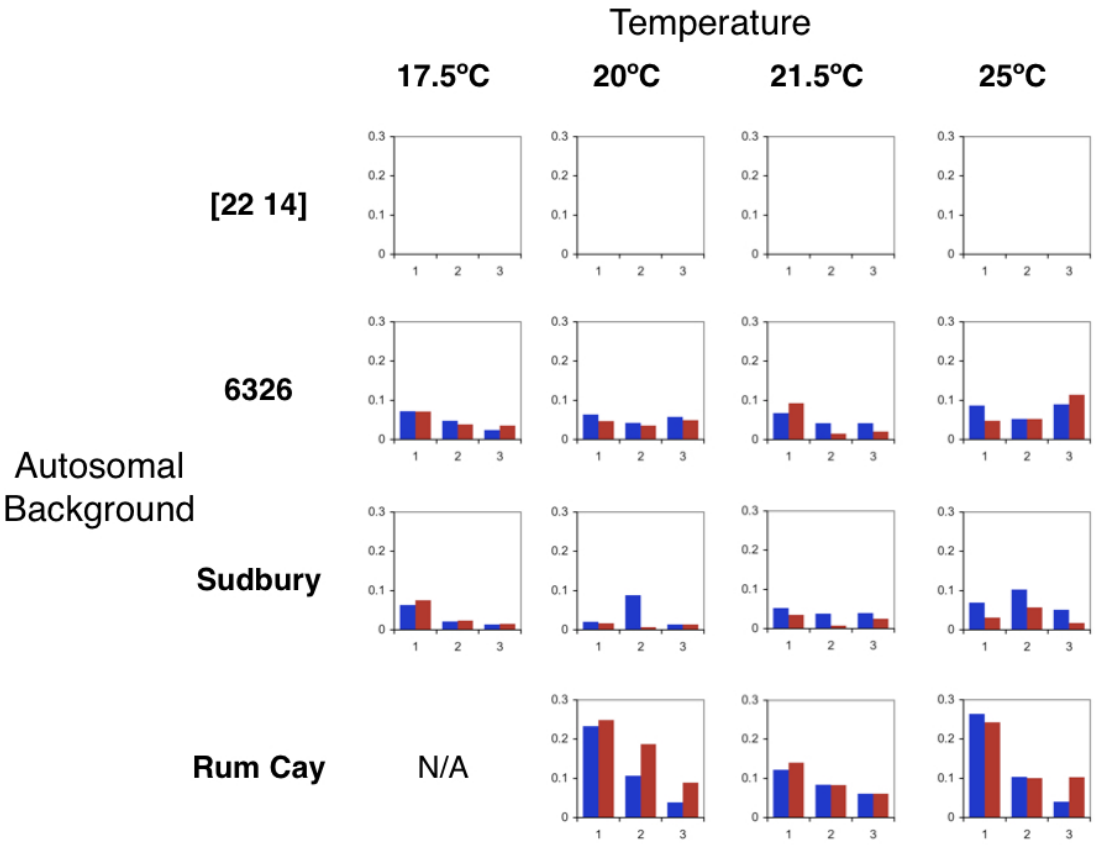
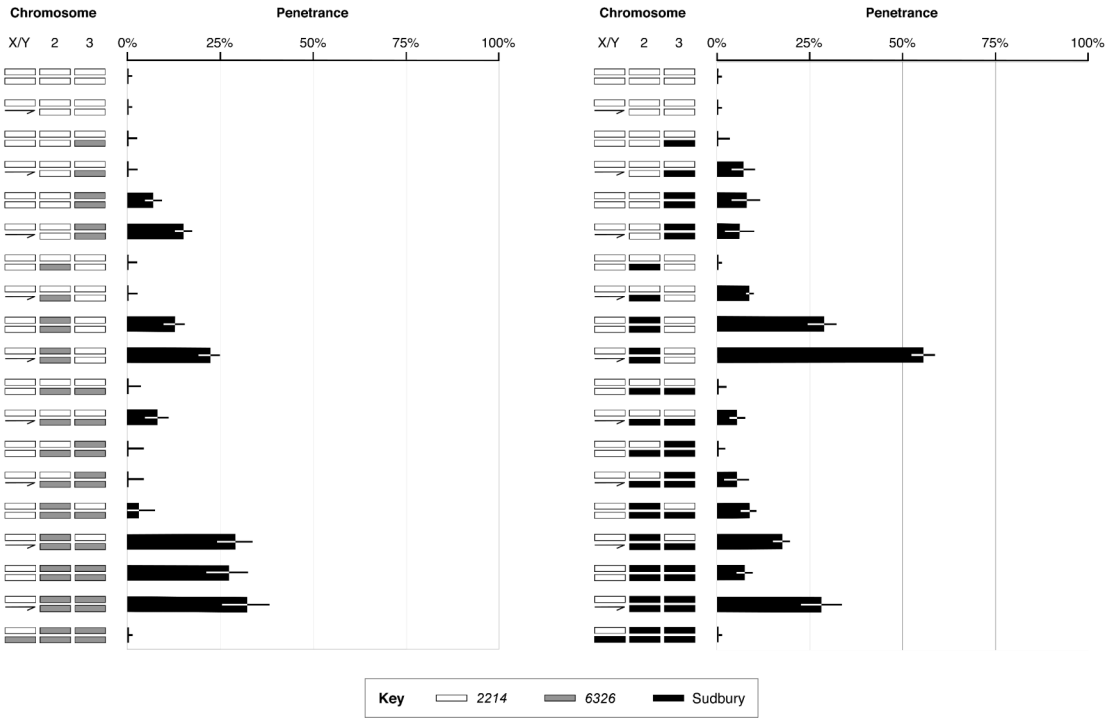


Figure 6. Suppression of *vs*^{2214} by *2214* autosomes



Chapter 5

Theoretical population genetics of incomplete penetrance: stochastic fitness and population epigenetics

Joseph Lachance

Abstract

Individuals with the same genotype do not always have the same phenotype. This phenomenon (incompletely penetrance) is widespread in natural populations and occurs for many traits. Using a mix of analytic theory and computer simulations, we examined how incomplete penetrance affects the evolutionary fate of alleles. Fitness was treated as a random variable, and maternal effects (epigenetic inheritance) was incorporated into our model. We found that incomplete penetrance reduced the sojourn times of new mutations. To a first approximation, mean fitness accurately predicted the probability of fixation of advantageous alleles. However, incomplete penetrance caused alleles that were neutral on average to behave as if they were slightly deleterious. Finally, we found that maternal effects lowered the probability of fixation of beneficial alleles and increased the probability of fixation of alleles that were neutral on average. However, maternal effects in natural populations are unlikely to be of sufficient magnitude to significantly impact the long-term evolutionary dynamics of incompletely penetrant alleles.

Introduction

Genotype-phenotype maps

In population genetics, genotypes are assigned fitness values. However, genotypes map to phenotypes and natural selection acts on organismal phenotypes rather than genotypes (Lewontin 2000). The study of genetic architecture of phenotypes is an important modern research program (Mackay 2001; Rockman 2008; Wu and Lin 2006), and understanding the structure of genotypes has been viewed as the most challenging question remaining in evolutionary biology (Mayr 2001). Multiple factors influence the genotype-to-phenotype map, including epistatic interactions, genotype-by-environment interactions, and developmental noise (Benfey and Mitchell-Olds 2008; Dowell et al. 2010). The genotype-phenotype map is context dependent, with Dobzhansky going so far to describe genes as norms of reaction (Dobzhansky 1937). Norms of reaction are the result of development, a process that has been envisioned as a ball rolling down an epigenetic landscape, with genotypes differing in the extent to which they are canalized (Waddington 1942). And while recent advances in evolutionary developmental biology have increased our understanding of how genotypes lead to phenotypes, much of this work does not bridge the gap to population genetics (Johnson and Porter 2001; Stern 2010). When genotypes map to more than one phenotype, simply knowing the average effect of a gene may not be enough to explain evolutionary dynamics. One such situation involves incomplete penetrance.

Incomplete penetrance

Dating back to the early twentieth century (Laubichler and Sarkar 2002), penetrance refers to "the proportion of individuals of a specified genotype that show the expected phenotype

under a defined set of environmental conditions" ((King et al. 2006), p. 327). Genotypes that always result in the same phenotype are completely penetrant while genotypes that only sometimes result in the expected phenotype are said to be incompletely penetrant. This phenomenon is ubiquitous and it has been observed in a wide range of natural populations and experimental systems (Bejjani et al. 2000; Gergen and Wieschaus 1986; Lalucque and Silar 2004; Legeai-Mallet et al. 1997). Multiple causes of incomplete penetrance exist, including threshold traits and variability in expression levels (Raj et al. 2010). Because the effects of genes can sometimes be masked, similarities exist between incomplete penetrance and recessivity. In addition, incompletely penetrant alleles can be viewed as uncanalized genotypes. If fitnesses differ for each phenotype, an important implication of incomplete penetrance is that that alleles can have multiple fitnesses. From a theoretical standpoint this means that fitness can be treated as a random variable. The fact that individual alleles can map to multiple phenotypes defines a form of stochasticity acting in addition to genetic drift (Gillespie 2006). How much does this affect the probability of fixation and sojourn times of alleles? Also, what conditions allow incompletely penetrant alleles to out-compete completely penetrant alleles?

Stochastic fitness-

Incomplete penetrance is one of many situations where fitness can be treated as a random variable instead of a constant. In effect, stochastic fitness models function as more accurate black boxes than classical constant-fitness models of population genetics (Felsenstein 1976; Gillespie 2006). The concept of stochastic fitness is also related to that of demographic stochasticity, where there is random variation in the mortality and reproduction of individuals. Classic studies of stochastic fitness revealed that arithmetic mean fitness fails to capture the full

evolutionary dynamics for a wide variety of situations (Felsenstein 1976; Gillespie 1974; Haldane and Jayakar 1963; Hartl and Cook 1974; Karlin and Levikson 1974; Levene 1953; Takahata et al. 1975; Templeton 1977). The overall effects of stochastic fitness can be non-intuitive and complex. This is because stochastic fitness reduces both the effective population size (Kimura and Crow 1963) and the effective fitnesses of alleles (Gillespie 1977). Both of these pressures act in the same direction if alleles are beneficial, reducing the probability of fixation. However, if alleles are deleterious these two pressures can act in opposite directions.

Aided by advances in computing power, the study of stochastic fitness and genetic context has been rekindled in recent years (Table 1). Such topics as genetic draft, stochastic gene expression, and variance in offspring number are now receiving attention (Gillespie, 2000; Kærn *et al*, 2005; Shpak and Proulx, 2007). Despite these advances, existing theory lacks a treatment of the population genetics of incompletely penetrant alleles. A thorough understanding of stochastic fitness and incomplete penetrance will allow one to distinguish between scenarios in which it is safe to use the simple models of classical population genetics as opposed to more complex models.

Multiple types of neutrality

Organisms with the same genotype can have different fitnesses, forcing us to re-examine what it means for alleles to be neutral. Just because an allele is neutral "on average" does not mean that every individual with that allele has the same fitness. With this in mind, Proulx and Adler developed a taxonomy of neutrality. They considered three types of neutrality: eu-neutrality, circum-neutrality, and iso-neutrality (Proulx and Adler 2010). Eu-neutrality encompasses the classic view of neutrality, where genotypes are always uncorrelated with

fitness. Circum-neutrality includes alleles that do not differ in terms of function, but which occur in different contexts (such as linkage to other non-neutral alleles). Iso-neutral alleles have different functional characteristics, yet under some conditions are neutral with respect to fitness. For example, frequency-dependent selection may cause alleles to have the same fitness at some frequencies, but not others (Heino et al. 1998; Trotter and Spencer 2008). In this paper, we are primarily concerned with iso-neutrality (where an incompletely penetrant allele has a mean fitness of one).

Maternal effects

Maternal effects can also influence the penetrance of alleles. These effects occur whenever there is "a non-lasting influence of the genotype or phenotype of the mother upon the phenotype of the immediate offspring" ((King et al. 2006), p. 235). Evidence of maternal effects exists in a number of systems, and potential causes include epistatic interactions, DNA methylation, and environmental correlations (Rasanen and Kruuk 2007; Wolf 2000; Youngson and Whitelaw 2008). Constant fitness models implicitly ignore maternal effects, and stochastic fitness is required to understand the evolution of genes with maternal effects. The simple genotype-phenotype map of incomplete penetrance (either an allele is penetrant or non-penetrant) facilitates incorporating maternal effects into population genetics models.

Epigenetics has alternatively been used to describe the process of development (Waddington 1942) and non-DNA inheritance (Berger et al. 2009). From a theoretical perspective, population genetic models that explicitly incorporate maternal effects can be viewed as a form of population epigenetics. Previous theory of maternal effects has indicated that mutation-selection balance frequencies are modified and selection may act within and between

families (Wade 1998). We are not concerned with the exact mechanism causing epigenetic inheritance; instead, what matters is that there is a positive correlation between the penetrance states of parents and offspring.

Outline

In this paper we extend population genetics theory to encompass incompletely penetrant alleles. Analytic theory was developed, as were stochastic computer simulations. These simulations were used to explore the evolutionary trajectories of incompletely penetrant alleles for a range of fitness parameters (including advantageous alleles and iso-neutral alleles). In addition, simulations allowed us to explore epigenetics on a population level by incorporating correlations between the penetrance states of parental and offspring alleles.

Methods

Model

A single locus Wright-Fisher model is assumed. Here, populations are haploid and generations are discrete. Population sizes are finite and denoted by N. Two alleles segregate (A and B), and mutations are ignored. Alleles contribute to next generation's gene pool via a process of sampling with replacement. In any given individual, allele A is either penetrant or non-penetrant, while allele B always maps to a single phenotype. The total number of copies of A is set equal to i, and the number of penetrant copies of A is set equal to j. This yields a simple expression for the allele frequency of A:

$$p = \frac{i}{N} \tag{1}$$

A full list of parameters is given in Table 2. After weighting by fitness, alleles are sampled with replacement to generate the population state of the next generation. However, because the *A* allele is incompletely penetrant, the number of penetrant copies of *A* needs to be calculated. Alleles inherit the penetrance state of their parent with a probability of *r* (with values of *r* above zero indicating the presence of maternal effects). In the absence of selection, *r* conveniently equals the correlation between parent and offspring penetrance. Maternal effects are initially assumed to be absent ($r = 0$). However, this assumption is relaxed later in the paper.). In the event that penetrance state is not directly inherited, the parameter π indicates the probability that a particular copy of the *A* allele is penetrant. $\pi = 1$ corresponds to complete penetrance.

The fitness of penetrant *A* alleles equals $1 + s$, and the fitness of non-penetrant alleles equals $1 + t$. For mathematical simplicity, fitness of *B* alleles is set at equal to one. Thus, negative values of *s* indicate that penetrant forms of the *A* allele are deleterious, and negative values of *t* indicate that non-penetrant forms of the *A* allele are deleterious. If *s* and *t* have opposite signs it is possible for the overall fitness of *A* alleles to be neutral. Conditions for iso-neutrality (assuming $r = 0$):

$$\pi s + (1 - \pi)t = 0 \tag{2}$$

Computer simulations

To test simplifying assumptions used in analytical theory and explore mathematically difficult regions of parameter space, stochastic simulations were also used. These Monte Carlo simulations were programmed in MATLAB (Mathworks 2005). Each simulation run began with a single copy of the *A* allele. Using the model described above, populations were allowed to evolve until either the *A* allele was fixed or it was lost. Sojourn times were recorded for each run, and simulations were run 10,000 to 250,000 times for each set of parameter values, with larger number of runs for scenarios in which fixation was rare. Table 2 lists parameters used in simulations. Computer simulations were particularly useful for situations in which $r > 0$ (i.e. maternal effects were present), as this condition was analytically intractable.

Results

Transition probabilities

Exact equations can be derived for the transition probabilities of the number of penetrant and non-penetrant copies of *A* in subsequent generations. These equations allow the effects of incompletely penetrant alleles to be directly compared to completely penetrant alleles that have the same expected fitness. During any generation, there are i copies of the *A* allele, j of which are penetrant. Assuming $r = 0$, the number of penetrant copies of the *A* allele is equal to j, where:

$$j \sim Binomial(i,\pi) \tag{3}$$

Equation 3 indicates that chance effects will be averaged out if many copies of *A* are present. If only a few copies of *A* exist, it is possible that they all are penetrant or non-penetrant. When maternal effects are present (i.e. $r > 0$), the number of penetrant copies of *A* depends on the relative proportion of *A* alleles that were penetrant in the previous generation.

$$j \sim Binomial\big(i, r(\text{realized penetrance of } A \text{ last generation}) + (1-r)\pi\big), \qquad (4)$$

The realized penetrance of *A* in Equation 4 is equal to the proportion of alleles that are actually penetrant in a given generation. Maternal effects are akin to a form of inertia, whereby the effects of early generations persist over time, only to be gradually averaged out. The marginal fitness of *A* alleles is weighted by the proportion of penetrant and non-penetrant alleles.

$$\dot{w} = 1 + \frac{js + (i-j)t}{i} \qquad (5)$$

Similarly, the mean fitness of the entire population is equal to:

$$\overline{w} = 1 + \frac{js + (i-j)t}{N} \qquad (6)$$

Assuming $r = 0$, the number of copies of *A* the next generation is equal to:

$$i' \sim Binomial\left(N, \frac{j(1+s) + (i-j)(1+t)}{N + js + (i-j)t}\right) \qquad (7)$$

Note that the parameter *j* in the equation above is itself binomially distributed. As per classical population genetics, genetic drift causes there to be a spread in allele frequencies next generation (Figure 1). The expected number of copies of an incompletely penetrant allele next generation is less than the expected number of copies of a completely penetrant allele with the same mean fitness. In addition, there is greater spread around the mean for incompletely penetrant alleles. This relates to Gillespie's concept of *variance in offspring number* (Gillespie 1974; Gillespie 1977). He stressed that variance in offspring number carries a cost, reducing the effective fitness of alleles. The idea is that the cost of having low fitness some of the time can overwhelm the benefits of having high fitness some of the time. In this framework, constant fitness can be viewed as a form of evolutionary bet hedging (Seger and Brockman 1987). As given by (Gillespie 1974):

$$w_{eff} = \mu - \frac{\sigma^2}{N}, \tag{8}$$

where μ is the mean number of offspring and σ^2 is the variance in offspring number. The effect of variance in offspring number is greatest in small populations. In the nomenclature of this paper (assuming $r = 0$):

$$w_{eff} = \left(1 + \pi s + (1-\pi)t\right) - \frac{\pi\left((1-\pi)(s-t)\right)^2 + (1-\pi)\left(\pi(t-s)\right)^2}{N} \tag{9}$$

Equation 9 can be used to describe the constant fitness allele that yields the same mean change in allele frequency as an incompletely penetrant allele. However, this equation only describes a single fitness moment. To fully understand the evolutionary trajectories of incompletely penetrant alleles one needs to know more than the effective fitness of an allele.

Jensen's Inequality states that the convex transformation of a mean is less than or equal to the mean after a convex transformation (Ruel and Ayres 1999). This inequality is reversed if functions are concave. If f is concave: $E[f(X)] \leq f(E[X])$. Given allele A's contribution to the current gene pool, the function for allele frequency next generation ($p' = p\dot{w}/\overline{w}$) is concave. This occurs because alleles under selection also affect the mean fitness of a population. Allele-specific fitness variance, such as that arising from incomplete penetrance, reduces expected allele frequencies next generation.

This concept can be illustrated by considering a single copy of allele A in a population of size N ($i = 1$). If allele A always has a fitness of one: $E(i') = 1$. However, if individual copies of allele A have a 50% chance of having a fitness of two and a 50% chance of having a fitness of zero: $E(i') = N/(N+1)$. It is more difficult for incompletely penetrant alleles to invade populations.

Diffusion approximation

There is a long history of approximating the behavior of Markov chains in population genetics via diffusion theory (Ewens 2004; Kimura 1962; Kimura 1964). In the mid-1970s Gillespie applied diffusion approximations to stochastic fitness models. Using Equations 4 and 5 from (Gillespie 1974), we can incorporate details from our incomplete penetrance model into a diffusion approximation framework. However, some caveats to Gillespie's derivations exist,

including the possibility that allele-specific variance in offspring violates constant population size assumptions and that fitness variance terms should be weighted differently (Shpak 2005). With respect to our model, it is important that the equations below assume that allele *B* always has a fitness of one. The incompletely penetrant allele *A* competes against the completely penetrant allele *B*. The mean change in allele frequency each in a single generation follows from (Gillespie 1974).

$$M(p) \approx p(1-p)(\pi s + (1-\pi)t) \tag{10}$$

The variance in allele frequency in a single generation also follows from (Gillespie 1974). Variance terms for genetic drift and allele-specific variance in fitness were added together (i.e. the covariance between different type of stochasticity was assumed to be zero).

$$V(p) \approx \frac{p(1-p)}{N}\left(1 + p(1-p)(s-t)^2\pi(1-\pi)\right) \tag{11}$$

The rightmost term in Equation 11 indicates the variance in allele frequency that is due to stochasticity in the number of penetrant copies of *A*. This term is larger at intermediate allele frequencies, when large differences in fitness exist between penetrant and non-penetrant forms, and when penetrance is intermediate.

Given the mean and variance terms for allele frequency change (Equations 10 and 11) the Kolmogorov backwards equation allows the probability of fixation to be obtained. The backward equation stems from the decomposition of the probability of fixation. For additional

details, see (Ewens 2004) and the Appendix B in (Gillespie 2004). After satisfying the boundary conditions $P(fix \mid p = 0) = 0$ and $P(fix \mid p = 1) = 1$, we have:

$$P(fix) \approx \frac{\int_0^{p_{initial}} e^{-B(y)} dy}{\int_0^1 e^{-B(y)} dy}, \tag{12}$$

where $B(y) = 2\int_0^y \frac{M(p)}{V(p)} dp$. Using the parameterization of Table 2:

$$B(y) = 2\int_0^y \frac{N(\pi s + (1-\pi)t)}{1 + p(1-p)(s-t)^2 \pi(1-\pi)} dp \tag{13}$$

If selection coefficients are small, the expression in Equation 13 can be simplified.

$$P(fix)_{|s|<<1,\ |t|<<1} \approx \frac{1 - e^{-2N(\pi s + (1-\pi)t)p_{initial}}}{1 - e^{-2N(\pi s + (1-\pi)t)}} \tag{14}$$

This probability of fixation is simply the probability of fixation given the mean fitnesses of an allele (Kimura 1962). Diffusion approximations can also be used to estimate fixation and loss times of mutant alleles (Kimura and Ohta 1969a; Kimura and Ohta 1969b).

Despite the apparent utility of Equation 14, a few caveats exist. First, diffusion approximations are only truly valid for moderately large *N* (Wakeley 2005), yet Gillespie's work

suggests that allele-specific variance in fitness only has a large effect in small populations. Second, the standard diffusion model of population genetics assumes that population size is constant (Feller 1951). Third, selection coefficients were assumed to be small. In contrast to constant fitness theory, the diffusion approximation for incompletely penetrant alleles has a term that contains the squared difference in selection coefficients. If selection coefficients have different signs (as is the case for iso-neutral alleles), individual selection coefficients can be small, but the difference can be large enough to render Equation 14 invalid. Also, population genetics models with maternal effects are not Markov chains. Because of these caveats, computer simulations were used to test expectations from the diffusion approximation and explore a wider range of the parameter space of incompletely penetrant alleles.

Probability of fixation

If an allele is advantageous and incompletely penetrant, the probability of fixation is largely determined by the mean fitness of that allele. This statement is true only when maternal effects are absent. Assuming that non-penetrant forms are neutral, the effective fitness of advantageous alleles is reduced when penetrance is incomplete. Conversely, the effective fitness of deleterious alleles is increased when penetrance is incomplete. As shown in Figure 2, there is strong agreement between simulation results and the diffusion approximation (Equation 14). This occurs for both moderate ($N = 1000$) and small ($N = 100$) population sizes. Incomplete penetrance causes there to be allele-specific variance in fitness, but mean fitness effects of advantageous alleles dwarf the cost of fitness variance. Substituting $t = 0$ into Equation 9 yields

$$w_{eff|(s>0,t=0)} = 1 + \pi s - \frac{s^2 \pi (1-\pi)}{N} \quad (15)$$

For reasonable selection coefficients and population sizes ($s << 1$, $N >> 10$), the effective fitness of advantageous alleles is greater than one.

If alleles are neutral "on average" (i.e. they are iso-neutral), they behave as if they are slightly deleterious. Recall that conditions for iso-neutrality are given by Equation 2. In Figure 3, the probabilities of fixation of iso-neutral alleles are compared to neutral expectations. For values of $s > 0.01$, incomplete penetrance reduces the probability of fixation of iso-neutral alleles. The relative probability of fixation of iso-neutral alleles appears to follow

$$P(fix)_{iso-neutral} \approx \frac{1}{(1+s^2)N} \quad (16)$$

Equation 16 is similar to the neutral probability of fixation, but with a scaled population size. Interestingly, the expression $N/(1+s^2)$ approximately equals the effective size of a haploid population with increased levels of fitness variance (Wright 1938). For the parameter values used in Figure 3, s^2 is the variance in fitness of a single copy of the A allele. Other than the fact that alleles must initially appear as a single copy, it is unclear why the probability of fixation of iso-neutrals scales with the fitness variance of a single copy of A rather than the fitness variance of the ensemble of all A alleles (which is lower when there are more copies of A present).

Sojourn times

Incomplete penetrance reduces the sojourn times of mutant alleles (Figure 4 and Table 3). One reason is because incomplete penetrance lowers the probability of fixation of new mutations, and fixation times tend to be longer than loss times. However, even when one conditions upon alleles ultimately fixing or becoming lost, sojourn times are lower. For example, simulations for neutral alleles yielded mean fixation times of 2000.9 generations, with a mean loss time of 13.29 generations ($r = 0$, $s = 0$, $t = 0$, $N = 1000$). This is consistent with expectations from coalescent theory (Wakeley 2009). Incompletely penetrant alleles that were iso-neutral had a mean fixation time of 1386.6 generations, and a mean loss time of 7.00 generations ($\pi = 0.5$, $r = 0$, $s = 1$, $t = -1$, $N = 1000$). Fixation and loss times for a range of parameter values are shown in Table 3. Fixation times scaled linearly with population size, and loss times were largely independent of population size. Highly advantageous alleles had shorter sojourn times than slightly advantageous alleles. However, given the same mean fitness, incomplete penetrance did not appreciably change the fixation or loss times of advantageous alleles. By contrast, incomplete penetrance affected the sojourn times of iso-neutral alleles, with higher magnitudes of fitness variance yielding shorter fixation and loss times. Note that simulations assumed that alleles were initially present as a single copy; the effect of incomplete penetrance on sojourn times may differ for alleles initially found at intermediate frequencies.

Modified sojourn times occur because of increased variance in transition probabilities. When penetrance is incomplete, newly arising alleles can be quickly lost. On the other hand, if the initial fitnesses of alleles are high they can quickly end up with a large frequency after only a few generations. Plots of allele frequency vs. time for advantageous alleles reaching fixation tend to be sigmoidal (Crow and Kimura 1970). High initial fitness values allow populations to bypass much of the "lag phase." Conversely, low initial fitness values lead to alleles being lost.

Once alleles are relatively common, stochastic effects are averaged out and chances are low that the ensemble of *A* alleles will have a particularly high or low marginal fitness.

Maternal effects

Maternal effects magnify the initial contexts of alleles. Consider a situation in which the penetrant form of an allele has higher fitness than the non-penetrant form ($s > t$). If the initial copy of *A* is penetrant, the ultimate probability of fixation will be higher than if the initial copy is non-penetrant. Positive correlations between parent and offspring penetrance states allow initial fitness effects to affect subsequent generations. In addition, higher values of *r* increase the time it takes to erase the initial fitness context of an allele. The model in this paper does not make any assumptions as to the cause of maternal effects. Non-zero values of *r* can occur via epigenetic modification of allele *A* (including methylation), genotype-by-environment interactions in which the environments of parents and offspring are correlated, or epistatic interactions with alleles at other loci. Ultimately, the impact of maternal effects depends on the magnitude and sign of selection coefficients.

Maternal effects reduce the probability of fixation of advantageous alleles. Consider a situation in which the penetrant form of an allele is high fitness and the non-penetrant form is neutral ($s > 0$, $t = 0$). For most values of *r*, simulations indicate that maternal effects have a minimal effect (Figure 5). However, when *r* is close to one, the probability of fixation is reduced relative to constant fitness expectations. Increased understanding can be gleaned by examining extreme parameter values. If maternal effects are complete (i.e. $r = 1$) all subsequent copies of *A* will have the same fitness as the initial mutant. From Kimura's diffusion equations

(Kimura 1962) we know that the probability of fixation of a new mutation with constant fitness is

$$P(fix)_{\text{constant fintess}} \approx \frac{1-e^{-2s}}{1-e^{-2Ns}} \tag{17}$$

This equation indicates that there are diminishing returns for higher values of s (e.g. $N = 1000$, $s = 0.05$ gives P(fix) = 0.095 and $s = 0.1$ gives P(fix)= 0.181). Doubling the value of a selection coefficient does not double the probability of fixation. When $r = 1$, the overall probability of fitness of incompletely penetrant alleles is calculated by weighting by the probability that the initial mutation is penetrant.

$$P(fix)_{r=1} \approx \pi \frac{1-e^{-2s}}{1-e^{-4Ns}} + (1-\pi)\frac{1-e^{-2t}}{1-e^{-4Nt}}. \tag{18}$$

Equation 18 allows the relative probabilities of fixation of constant fitness and incompletely penetrant alleles to be compared (holding mean fitness constant). Figure 6A illustrates that selection coefficients need to be moderately large (on the order of $s > 0.05$) for incomplete penetrance to reduce the probability of fixation of advantageous alleles.

Maternal effects increase the probability of fixation of iso-neutral alleles. As shown in Figure 7, simulations reveal that moderate ($r > 0.5$) levels of maternal effects are required for probability of fixation to be modified. If $r = 1$, even small selection coefficients can modify the probability of fixation of iso-neutral alleles (Figure 6B). Iso-neutral alleles that initially have high fitness are able to rise to moderate frequencies due to selection and maternal effects. So

long as $r < 1$, the initial context is eventually averaged out. This effectively increases the initial allele frequencies of iso-neutral alleles, increasing the probability of fixation. On the other hand, iso-neutral alleles that initially have low fitness are likely to be quickly lost from a population. However, a classic finding in population genetics is that most new mutations are lost by drift (Crow and Kimura 1970). This means that loss of alleles because of low initial fitness "comes for free" (i.e. the overall probability of fixation is not greatly reduced). What is important here is that the fitness of non-penetrant forms of allele *A* is less than the fitness of the allele *B*.

Discussion

Summary

In population genetics it is not sufficient to know the fitness of penetrant and non-penetrant forms of an allele. A full understanding of evolutionary dynamics also requires one to know the probability that an allele is penetrant. To a first approximation, the mean fitness of an allele determines its evolutionary trajectory. This agrees with theory that examines the population genetics of conditionally expressed alleles (Van Dyken and Wade 2010). Incomplete penetrance reduced the effective magnitudes of selection coefficients for advantageous and deleterious alleles. Diffusion approximations, with assumptions of weak selection coefficients, suggested that mean fitness determines the probability of fixation. Computer simulations agreed with this prediction for advantageous alleles, but not for iso-neutral alleles. This is because fitness variance associated with incomplete penetrance modifies transition probabilities. Finally, maternal effects decreased the probability of fixation of advantageous alleles and increased the probability of fixation of iso-neutral alleles.

Mean fitnesses of alleles

When is it safe to simply consider the mean effects of an allele? Transition probabilities reveal that if alleles are common, individual-level fitness differences are averaged out. Incomplete penetrance matters most when new mutations enter populations. Variance in the proportion of alleles that are penetrant is maximized at $\pi = 0.5$ (see Equation 3). This suggests that only considering mean fitness is most applicable when alleles are almost always penetrant or non-penetrant (i.e. π is close to 0 or 1). So long as the mean effect of alleles is at least nearly neutral (i.e. $|N(\pi s + (1-\pi)t)| > 1$) incomplete penetrance has only a slight effect on the probability of fixation (Figure 2). Maternal effects can shrink the region of parameter space where it is safe to only consider the mean fitness of an allele. However, high levels of maternal effects ($r > 0.9$ for advantageous alleles and $r > 0.5$ for iso-neutral alleles) are required for this to be the case. Maternal effects in most natural populations are unlikely to be this strong. For example, incomplete penetrance of the *vesiculated* locus in *D. melanogaster* yielded parameter values on the order of $r = 0.038$ and $\pi = 0.290$ (see Chapter 4).

Incomplete penetrance vs. recessivity

While incomplete penetrance and recessivity both involve situations where the effects of genes can be masked, there are important differences between each of these phenomena. First, dominance is only relevant for diploid individuals, while incomplete penetrance applies regardless of ploidy. Ample empirical evidence of incomplete penetrance exists for diploids (Hollocher et al. 1992; Ogas et al. 1997; Shields and Harris 2000) and haploids (Veening et al. 2008). Second, variance in fitness due to incomplete penetrance reduces the effective fitness of

alleles when population size is small. No such cost exists for recessive alleles. Third, marginal fitnesses of recessive alleles are frequency dependent (as the proportion of alleles that are found in heterozygous or homozygous form varies by frequency). By contrast, marginal fitnesses of incompletely penetrant alleles do not vary by frequency. This difference is likely to affect allele frequencies at mutation-selection balance.

Haldane's sieve refers to bias against the establishment of recessive beneficial mutations (Turner 1981). This is because the fitness effects of beneficial recessive alleles are masked while the dominant mutations are subject to selection. While new mutations are subject to Haldane's sieve, recent theory indicates that this pattern does not hold for standing variation that has been subject to a change in environment (Orr and Betancourt 2001). Similarly, a filter exists whereby completely penetrant beneficial alleles are more likely to be fixed than incompletely penetrant beneficial alleles. Because they have been filtered by selection, wild-type alleles are more likely to be more highly penetrant than new mutations. At present, it is unknown whether this filter also applies to incompletely penetrant alleles fixing from standing genetic variation after a change in environment. This question is testable in laboratory populations of *D. melanogaster*.

Conclusion

This study indicates that evolutionary effects cannot always be reduced to the mean fitness of alleles. Genetic context matters, especially when an allele first enters a population. Future theoretical work can benefit by either explicitly modeling the genetic and environmental contexts of alleles or by treating fitness as a random variable. Whenever possible, empirical studies should aim to estimate multiple fitness moments.

References

Bejjani, B. A., D. W. Stockton, R. A. Lewis, K. F. Tomey, D. K. Dueker, M. Jabak, W. F. Astle et al. 2000. Multiple CYP1B1 mutations and incomplete penetrance in an inbred population segregating primary congenital glaucoma suggest frequent de novo events and a dominant modifier locus. Hum Mol Genet 9:367-374.

Benfey, P. N., and T. Mitchell-Olds. 2008. From genotype to phenotype: systems biology meets natural variation. Science 320:495-497.

Berger, S. L., T. Kouzarides, R. Shiekhattar, and A. Shilatifard. 2009. An operational definition of epigenetics. Genes Dev 23:781-783.

Crow, J. F., and M. Kimura. 1970, An introduction to populaiton genetics theory. New York, Harper and Row.

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

Dowell, R. D., O. Ryan, A. Jansen, D. Cheung, S. Agarwala, T. Danford, D. A. Bernstein et al. 2010. Genotype to phenotype: a complex problem. Science 328:469.

Ewens, W. J. 2004, Mathematical population genetics: Interdisciplinary applied mathematics v. 27. New York, Springer.

Feller, W. 1951, Diffusion processes in genetics. S. Neyman, ed. Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability:227-246.

Felsenstein, J. 1976. The theoretical population genetics of variable selection and migration. Annu Rev Genet 10:253-280.

Gergen, J. P., and E. F. Wieschaus. 1986. Localized requirements for gene activity in segmentation of *Drosophila* embryos: analysis of *armadillo*, *fused*, *giant* and *unpaired* mutations in mosaic embryos. Development Genes and Evolution 195:49-62.

Gillespie, J. H. 1974. Nautural selection for within-generation variance in offspring number. Genetics 76:601-606.

—. 1977. Natural selection for variance in offpspring number: a new evolutionary principle. American Naturalist 111:1010-1014.

—. 2001. Is the population size of a species relevant to its evolution? Evolution Int J Org Evolution 55:2161-2169.

—. 2004, Population Genetics: A Concise Guide. Baltimore, The Johns Hopkins University Press.

—. 2006. Stochastic Processes in Evolution *in* C. W. Fox, and J. B. Wolf, eds. Evolutionary Genetics: Concepts and Case Studies. Oxford, UK, Oxford University Press.

Haldane, J. B. S., and S. D. Jayakar. 1963. Polymorphism due to selection of varying direction. J. Genet. 58:237-242.

Hartl, D. L., and R. D. Cook. 1974. Autocorrelated Random Environments and Their Effects on Gene Frequency. Evolution 28:275-280.

Heino, M., J. A. J. Metz, and V. Kaitala. 1998. The enigma of frequency-dependent selection. Trends Ecol. Evol. 13:367-370.

Hollocher, H., A. R. Templeton, R. DeSalle, and J. S. Johnston. 1992. The molecular through ecological genetics of abnormal abdomen. IV. Components of genetic variation in a natural population of Drosophila mercatorum. Genetics 130:355-366.

Johnson, N. A., and A. H. Porter. 2001. Toward a new synthesis: population genetics and evolutionary developmental biology. Genetica 112/113:45-58.

Kærn, M., T. Elston, W. Blake, and J. Collins. 2005. Stochasticity in Gene Expression: from Theories to Phenotypes. Nature Reviews Genetics 6:451-464.

Karlin, S., and B. Levikson. 1974. Temporal fluctuations in selection intensities: case of small selection size. Theor Popul Biol 63:383-412.

Kim, Y. 2007. Rate of Adaptive Peak Shifts with Partial Genetic Robustness. Evolution 61:1847-1856.

Kimura, M. 1962. On the probability of fixation of mutant genes in a population. Genetics 47:713-719.

—. 1964. Diffusion models in population genetics. J. Appl. Prob. 1:177-232.

Kimura, M., and J. F. Crow. 1963. The measurement of effective population number. Evolution 17:279-288.

Kimura, M., and T. Ohta. 1969a. The Average Number of Generations until Extinction of an Individual Mutant Gene in a Finite Population. Genetics 61:701-709.

—. 1969b. The Average Number of Generations until Fixation of a Mutant Gene in a Finite Population. Genetics 61:763-771.

King, R. C., W. D. Stansfield, and P. K. Mulligan. 2006, A dictionary of genetics. Oxford, Oxford University Press.

Lalucque, H., and P. Silar. 2004. Incomplete penetrance and variable expressivity of a growth defect as a consequence of knocking out two K(+) transporters in the euascomycete fungus Podospora anserina. Genetics 166:125-133.

Lande, R. 2007. Expected Relative Fitness and the Adaptive Topography of Fluctuating Selection. Evolution 61:1835-1846.

Laubichler, M. D., and S. Sarkar. 2002. Flies, genes, and brains: Oskar Vogt, Nikolai Timofeeff-Ressovsky, and the origin of the concepts of penetrance and expressivity *in* L. S. Parker, and R. A. Ankeny, eds. Mutating concepts, evolving disciplines: genes, medicine, and society. Philosophy and Medicine. Dordrecht, Kluvier.

Legeai-Mallet, L., A. Munnich, P. Maroteaux, and M. Le Merrer. 1997. Incomplete penetrance and expressivity skewing in hereditary multiple exostoses. Clin Genet 52:12-16.

Levene, H. 1953. Genetic Equilibrium When More Than One Ecological Niche is Available. American Naturalist 87:331-333.

Lewontin, R. C. 2000, The Triple Helix: Gene, Organism and Environment. Cambridge, Mass., Harvard University Press.

Mackay, T. F. 2001. The genetic architecture of quantitative traits. Annu Rev Genet 35:303-339.

Mathworks. 2005.MATLAB 7.The Mathworks, Natick, MA.

Mayr, E. 2001, What evolution is. New York, Basic Books.

Ogas, J., J. C. Cheng, Z. R. Sung, and C. Somerville. 1997. Cellular differentiation regulated by gibberellin in the Arabidopsis thaliana pickle mutant. Science 277:91-94.

Okasha, S. 2004. The "averaging fallacy" and the levels of selection. Biology and Philosophy 19:167-184.

Orr, H. A. 2007. Absolute Fitness, Relative Fitness, and Utility. Evolution 61:2997-2300.

Orr, H. A., and A. J. Betancourt. 2001. Haldane's sieve and adaptation from the standing genetic variation. Genetics 157:875-884.

Proulx, S. R. 2000. The ESS under Spatial Variation with Applications to Sex Allocation. Theor Popul Biol 58:33-47.

Proulx, S. R., and F. R. Adler. 2010. The standard of neutrality: still flapping in the breeze? J Evol Biol 23:1339-1350.

Proulx, T., and T. Day. 2001. What can invasion analyses tell us about evolution under stochasticity in finite populations? Selection 2:1-15.

Raj, A., S. A. Rifkin, E. Andersen, and A. van Oudenaarden. 2010. Variability in gene expression underlies incomplete penetrance. Nature 463:913-918.

Rasanen, K., and E. B. Kruuk. 2007. Maternal effects and evolution at ecological time-scales. Functional Ecology 21:408-421.

Rice, S. H. 2008. A stochastic version of the Price equation reveals the interplay of deterministic and stochastic processes in evolution. BMC Evol Biol 8:262.

Rockman, M. V. 2008. Reverse engineering the genotype-phenotype map with natural genetic variation. Nature 456:738-744.

Ruel, J. J., and M. P. Ayres. 1999. Jensen's inequality predicts effects of environmental variation. Trends Ecol Evol 14:361-366.

Salathe, M., J. Van Cleve, and M. W. Feldman. 2009. Evolution of stochastic switching rates in asymmetric fitness landscapes. Genetics 182:1159-1164.

Seger, J., and H. J. Brockman. 1987. What is bet-hedging?, Pages 182-211 *in* P. H. Harvey, and L. Partridge, eds. Oxford surveys in evolutionary biology. New York, Oxford University Press.

Shields, P. G., and C. C. Harris. 2000. Cancer risk and low-penetrance susceptibility genes in gene-environment interactions. J Clin Oncol 18:2309-2315.

Shpak, M. 2005. Evolution of Variance in Offspring Number: the Effects of Population Size and Migration. Theory in Bioscience 124:65-85.

—. 2007. Selection Against Demographic Stochasticity in Age-Structured Populations. Genetics 177:2181-2194.

Stern, D. L. 2010, Evolution, development, & the predictable genome. Greenwood Village, Colo., Roberts And Company Publishers.

Takahata, N., K. Ishii, and H. Matsuda. 1975. Effect of temporal fluctuation of selection coefficient on gene frequency in a population. PNAS 72:4541-4145.

Taylor, J. E. 2009. The genealogical consequences of fecundity variance polymorphism. Genetics 182:813-837.

Templeton, A. R. 1977. Survival Probabilities of Mutant Alleles in Fine-Grained Environments. American Naturalist 111:951-966.

Trotter, M. V., and H. G. Spencer. 2008. The generation and maintenance of genetic variation by frequency-dependent selection: constructing polymorphisms under the pairwise interaction model. Genetics 180:1547-1557.

Turner, J. R. G. 1981. Adaptation and evolution in *Heliconius*: a defense of NeoDarwinism. Annu Rev Ecol Syst 12:99-121.

Van Dyken, J. D., and M. J. Wade. 2010. The genetic signature of conditional expression. Genetics 184:557-570.

Veening, J. W., W. K. Smits, and O. P. Kuipers. 2008. Bistability, epigenetics, and bet-hedging in bacteria. Annu Rev Microbiol 62:193-210.

Waddington, C. H. 1942. The canalization of development and the inheritance of acquired characters. Nature 150:563-565.

Wade, J. D. 1998. The evolutionary genetics of maternal effects, Pages 5 *in* T. A. Mousseau, and C. W. Fox, eds. Maternal effects as adaptations. New York, Oxford Univ. Press.

Wakeley, J. 2005. The limits of theoretical population genetics. Genetics 169:1-7.

—. 2009, Coalescent theory : an introduction. Greenwood Village, Colo., Roberts & Co. Publishers.

Whitlock, M. C., and R. Gomulkiewicz. 2005. Probability of Fixation in a Heterogeneous Environment. Genetics 171:1407-1417.

Wolf, J. B. 2000. Gene interactions from maternal effects. Evolution 54:1882-1898.

Wright, S. 1938. Size of population and breeding structure in relation to evolution. Science 87:430-431.

Wu, R., and M. Lin. 2006. Functional mapping - how to map and study the genetic architecture of dynamic complex traits. Nat Rev Genet 7:229-237.

Youngson, N. A., and E. Whitelaw. 2008. Transgenerational epigenetic effects. Annu Rev Genomics Hum Genet 9:233-257.

Tables

Table 1. Recent stochastic fitness papers

Process	Reference
Game theory and variance in offspring number	(Proulx 2000)
Environmental stochasticity with finite *N*	(Proulx and Day 2001)
Genetic draft	(Gillespie 2001)
Averaging fitness across contexts	(Okasha 2004)
Stochasticity and gene expression	(Kærn et al. 2005)
Variance in offspring number with migration	(Shpak 2005)
Environmental heterogeneity	(Whitlock and Gomulkiewicz 2005)
Multiple forms of stochasticity	(Gillespie 2006)
Fluctuating selection	(Lande 2007)
Perturbation of genetic robustness	(Kim 2007)
Genetic bet-hedging	(Orr 2007)
Demographic stochasticity and age-structure	(Shpak 2007)
Stochastic version of the Price equation	(Rice 2008)
Stochastic switching	(Salathe et al. 2009)
Fecundity variance and coalescent theory	(Taylor 2009)
Conditional expression	(Van Dyken and Wade 2010)

Table 2. List of parameters for the incomplete penetrance model

Parameter	Meaning
p	Frequency of the A allele
$1 - p$	Frequency of the B allele
N	Population size
i	Number of A alleles
j	Number of penetrant A alleles
$i - j$	Number of non-penetrant A alleles
r	Probability of inheriting penetrance state
π	Penetrance of A allele (if state is not inherited)
s	Selection coefficient of penetrant A alleles
t	Selection coefficient of non-penetrant A alleles
$\dot{w}$	Marginal fitness of A alleles
$\overline{w}$	Mean fitness of the population

Table 3. Probability of fixation, fixation times, and loss times

Comparison	Parameter values	P(Fix)	Mean T_{fix}	Mean T_{loss}
Population size	$\pi = 0.5$, $r = 0$, $s = 1$, $t = -1$			
	... $N = 100$	5.06×10^{-3}	133.5	4.71
	... $N = 200$	2.38×10^{-3}	279.6	5.44
	... $N = 400$	1.38×10^{-3}	567.0	6.05
	... $N = 800$	6.56×10^{-4}	1070	6.90
Strength of selection	$r = 0$, $t = 0$, $N = 1000$			
	... $\pi = 1$, $s = 0.01$	1.94×10^{-2}	695.0	8.28
	... $\pi = 1$, $s = 0.1$	1.77×10^{-1}	119.7	4.11
	... $\pi = 0.5$, $s = 0.02$	9.73×10^{-3}	1110.0	9.78
	... $\pi = 0.5$, $s = 0.2$	1.74×10^{-1}	119.6	4.09
Fitness variance	$\pi = 0.5$, $r = 0$, $N = 1000$			
	... $s = 0.2$, $t = -0.2$	1.09×10^{-3}	2073.3	12.66
	... $s = 0.4$, $t = -0.4$	8.8×10^{-4}	2000.3	10.60
	... $s = 0.6$, $t = -0.6$	6.0×10^{-4}	1676.7	6.05
	... $s = 0.8$, $t = -0.8$	6.6×10^{-4}	1535.0	8.42
	... $s = 1$, $t = -1$	5.0×10^{-4}	1339.8	6.87

Table 3. Probability of fixation, fixation times, and loss times. Monte Carlo simulations were programmed in MATLAB. Simulations were run the following number of times for each comparison: 250000 for population size comparisons, 125000 for strength of selection comparisons, and 100000 for fitness variance comparisons.

Figure legends

Figure 1. Transition probabilities. The allele frequency distribution of the number of copies of an allele next generation (i') is shown for an allele that is present as a single copy in the current generation ($i = 1$). The probability distribution of an incompletely penetrant allele is show in gray, and the probability distribution of a completely penetrant allele is show in black. Both distributions refer to alleles with a mean fitness of 1.5. In the incomplete penetrance case, individual alleles have 50% chance of having a fitness of two and a 50% chance of having a fitness of one. However, in the constant fitness case fitness is always equal to 1.5.

Figure 2. Probability of fixation of advantageous alleles. Expectations from diffusion approximations and computer simulations were compared over a range of selection coefficients. Populations initially contain a single copy of the *A* allele. Multiple population sizes were considered. Overall, there is a strong match between probabilities of fixations generated by each method.

Figure 3. Probability of fixation of iso-neutral alleles. Expectations from the neutral theory and computer simulations were compared over a range of selection coefficients. Populations initially contain a single copy of the *A* allele. Penetrance was assumed to be 0.50 and fitnesses of iso-neutral alleles were constrained so that $s = -t$. High levels of allele-specific fitness variance reduced probabilities of fixation relative to neutral expectations. For the set of parameter values tested, the probability of fixation of iso-neutral alleles appears to follow Equation 16.

Figure 4. Fixation times. Data were obtained using computer simulations. In panel A, allele *A* is eu-neutral (i.e. fitness is always equal to one). In panel B, allele *A* is iso-neutral (each copy of *A* has a 50% chance of having a fitness of two and a 50% chance of having a fitness of one). Iso-neutral alleles have shorter sojourn times. Populations initially contain a single copy of the *A* allele.

Figure 5. Maternal effects and advantageous alleles. Computer simulations reveal that the probability of fixation of advantageous alleles is largely unaffected by maternal affects. However, values of *r* close to one yield reduced fixation probabilities compared to a constant fitness scenario with the same mean fitness. Populations initially contain a single copy of the *A* allele.

Figure 6. Relative fixation probabilities with fully inherited maternal effects. Populations initially contain a single copy of the *A* allele, and Equation 18 was used to obtain fixation probabilities. Panel A: Large selection coefficients are required for fully inherited maternal effects ($r = 1$) to modify the fixation probabilities of advantageous alleles. Constant fitness expectations are show in gray ($\pi=1$, s = the mean fitness of *A*, $t = 0$), and incompletely penetrant expectations are show in black ($\pi=0.5$, s = twice the mean fitness of *A,* $t = 0$). Panel B: Allele-specific fitness variance does not have to be large for fully inherited maternal effects to modify the probability of fixation of iso-neutral alleles. Neutral expectations are show in gray ($s = 0$, $t = 0$), and iso-neutral expectations are show in black ($\pi=0.5$, $t = -s$).

Figure 7. Maternal effects and iso-neutral alleles. Computer simulations reveal that maternal effects can increase the probability of fixation of iso-neutral alleles if the parameter r is moderately large ($r > 0.5$ for the tested parameter values). Populations initially contain a single copy of the A allele.

Figures

Figure 1. Transition probabilities

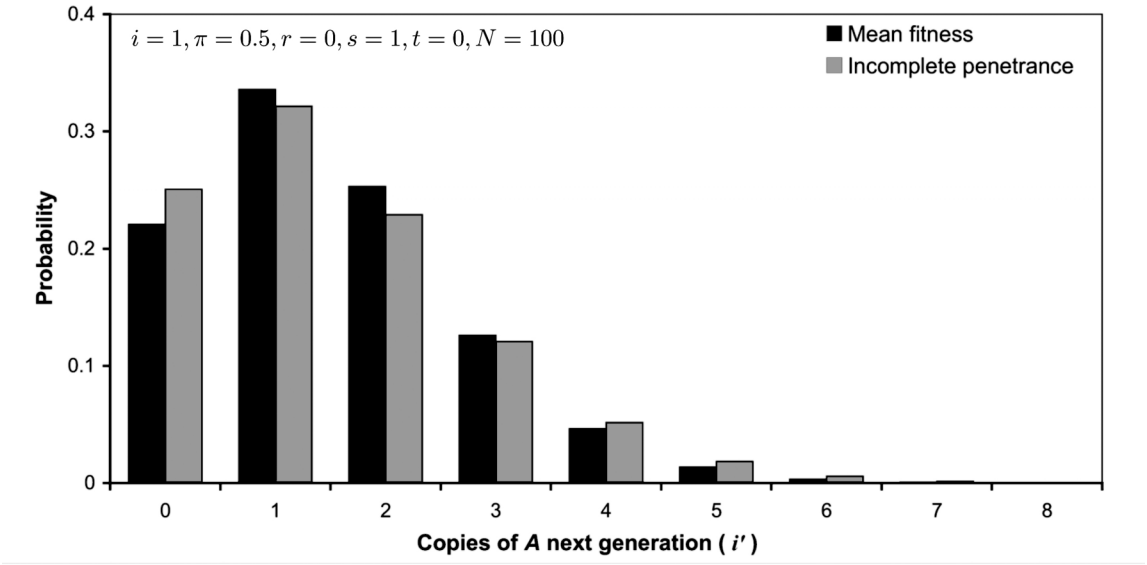


Figure 2. Probability of fixation of advantageous alleles

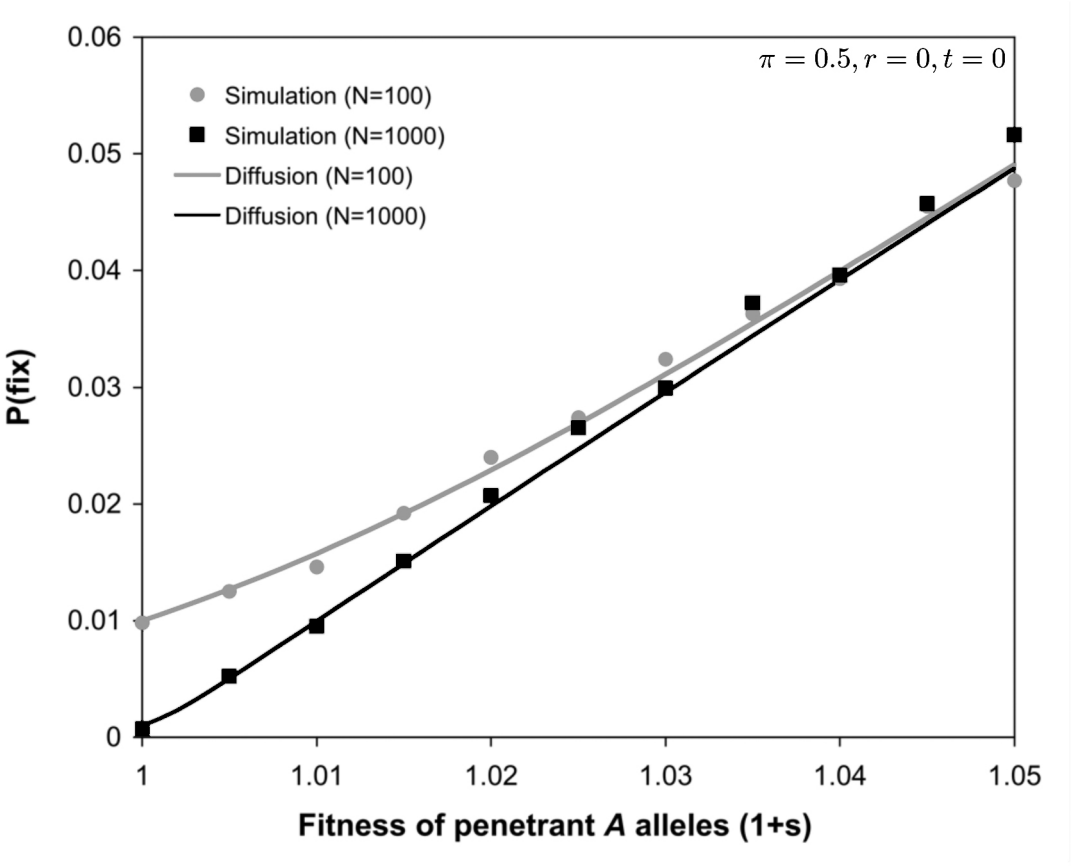


Figure 3. Probability of fixation of iso-neutral alleles

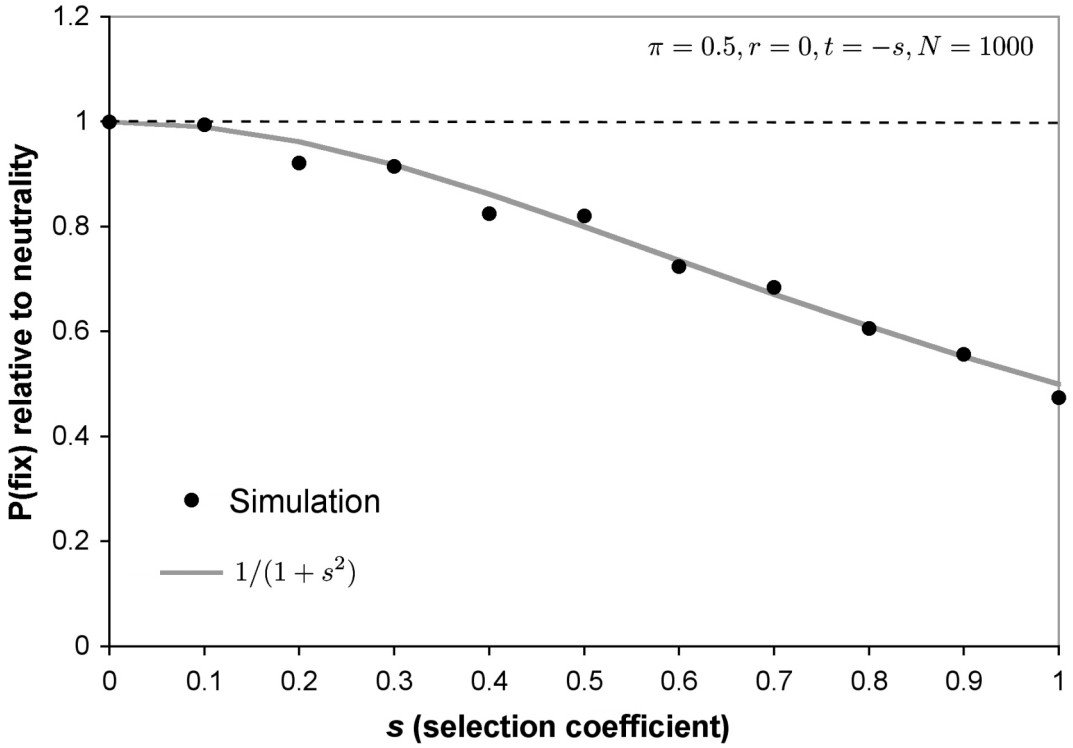


Figure 4. Fixation times

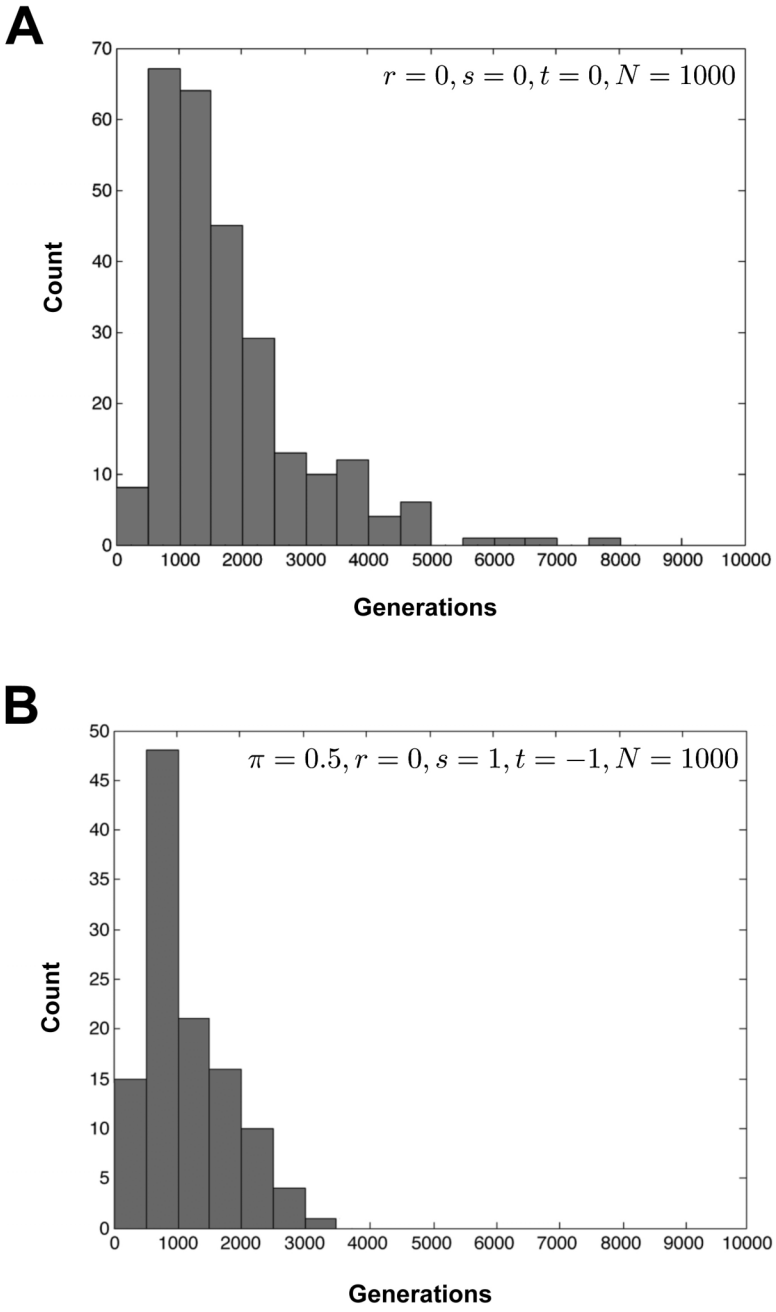


Figure 5. Maternal effects and advantageous alleles

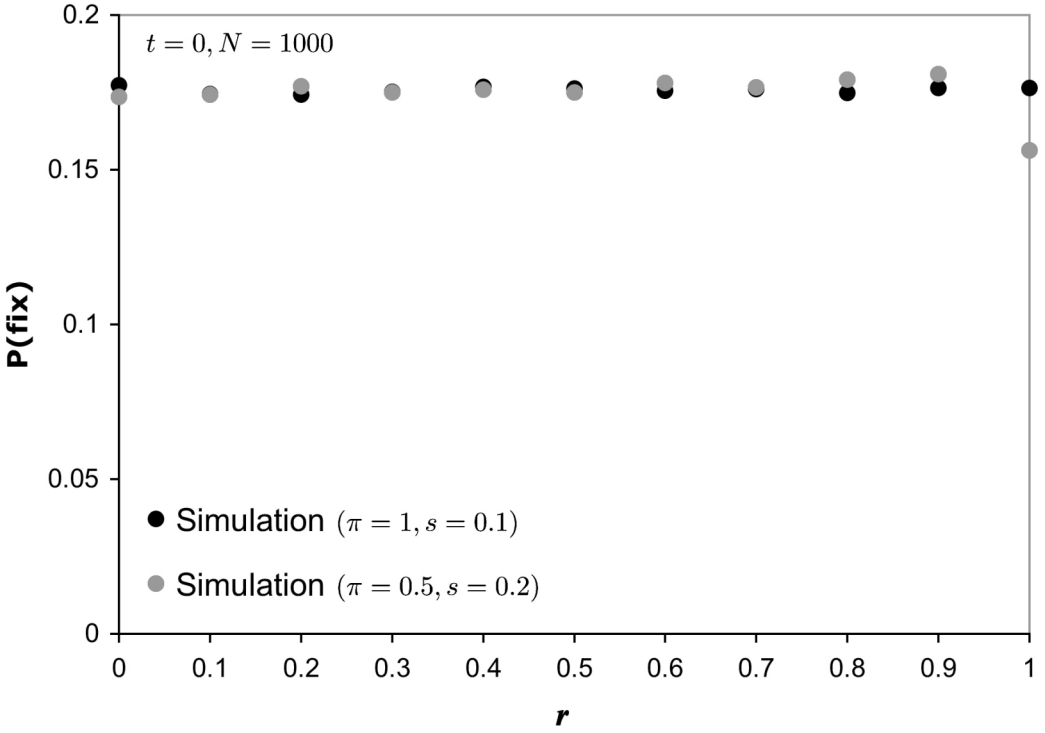


Figure 6. Relative fixation probabilities with fully inherited maternal effects

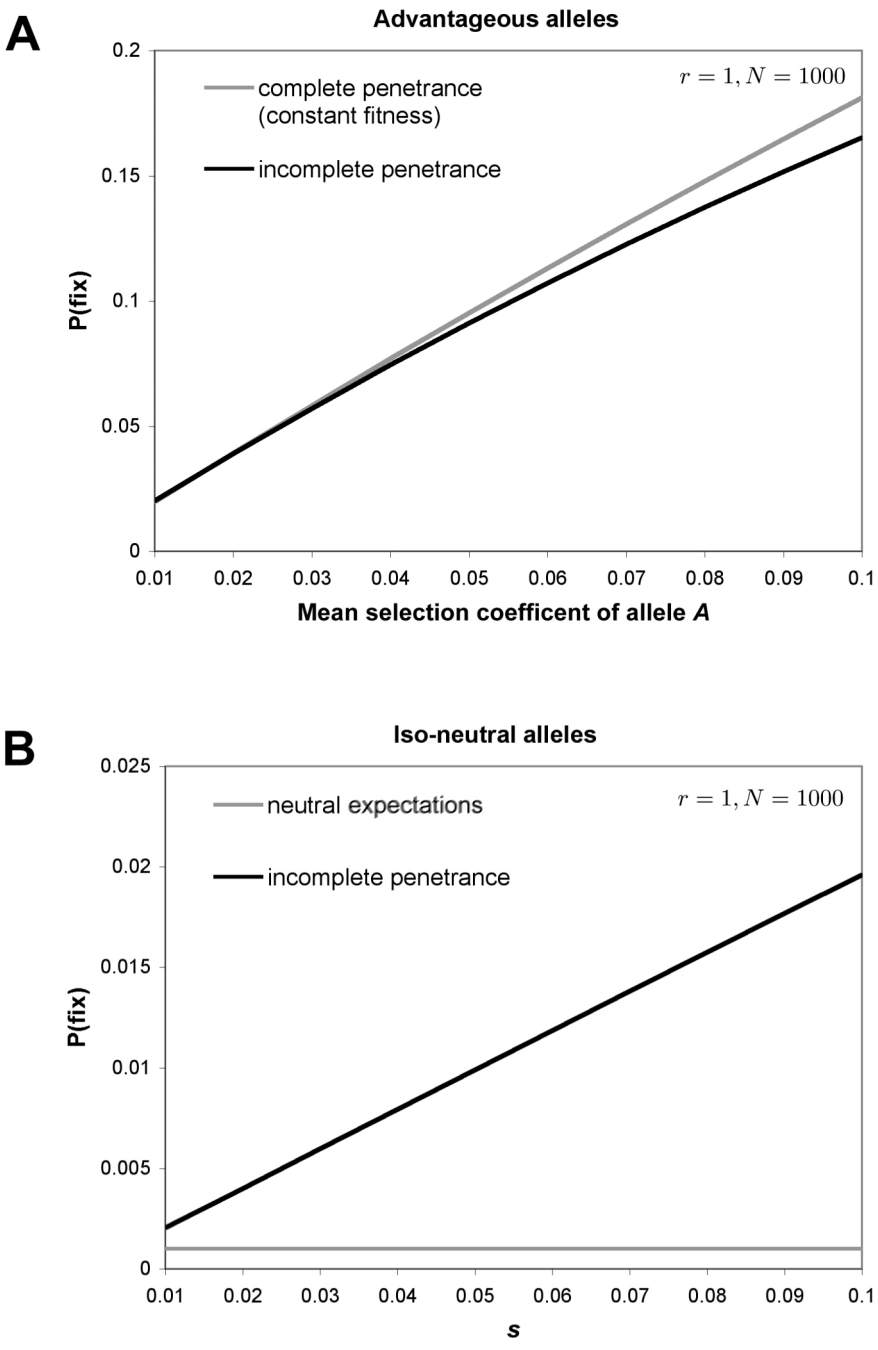
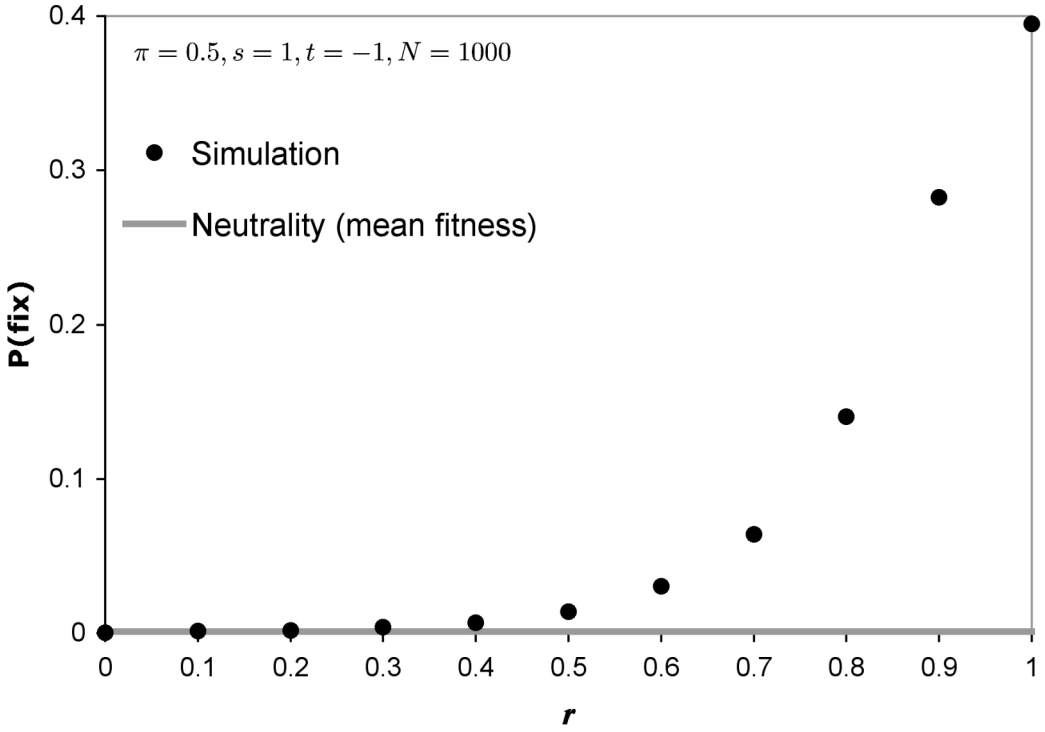


Figure 7. Maternal effects and iso-neutral alleles



References

Chapter 1

Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne, P. G. Amanatides, S. E. Scherer et al. 2000. The genome sequence of Drosophila melanogaster. Science 287:2185-2195.

Begun, D. J., A. K. Holloway, K. Stevens, L. W. Hillier, Y. P. Poh, M. W. Hahn, P. M. Nista et al. 2007. Population genomics: whole-genome analysis of polymorphism and divergence in Drosophila simulans. PLoS Biol 5:e310.

Carlson, E. A. 2004, Mendel's legacy: the origin of classical genetics. Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press.

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

Crow, J. F. 2001. The beanbag lives on. Nature 409:771.

Demuth, J. P., and M. J. Wade. 2007. Population differentiation in the beetle Tribolium castaneum. II. Haldane'S rule and incipient speciation. Evolution Int J Org Evolution 61:694-699.

Eldar, A., and M. B. Elowitz. 2010. Functional roles for noise in genetic circuits. Nature 467:167-173.

Greenspan, R. J. 2004, Fly Pushing: The Theory and Practice of *Drosophila* Genetics. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.

Haldane, J. B. 1964. A Defense of Beanbag Genetics. Perspect Biol Med 7:343-359.

Harrison, J. S., and S. Edmands. 2006. Chromosomal basis of viability differences in *Tigropus californicus* interpopulation hybrids. Journal of Evolutionary Biology 19:2040-2051.

Hollocher, H., A. R. Templeton, R. DeSalle, and J. S. Johnston. 1992. The molecular through ecological genetics of abnormal abdomen. IV. Components of genetic variation in a natural population of Drosophila mercatorum. Genetics 130:355-366.

Kimura, M. 1964. Diffusion models in population genetics. J. Appl. Prob. 1:177-232.

Kohler, R. E. 1994, Lords of the Fly: Drosophila Genetics and the Experimental Life. Chicago, University of Chicago Press.

Lachance, J., and J. R. True. 2010. X-autosome incompatibilities in *Drosophila melanogaster*: tests of Haldane's rule and geographic patterns within species. Evolution 64:3035-3046.

Laubichler, M. D., and S. Sarkar. 2002. Flies, genes, and brains: Oskar Vogt, Nikolai Timofeeff-Ressovsky, and the origin of the concepts of penetrance and expressivity *in* L. S. Parker, and R. A. Ankeny, eds. Mutating concepts, evolving disciplines: genes, medicine, and society. Philosophy and Medicine. Dordrecht, Kluvier.

Lunzer, M., S. P. Miller, R. Felsheim, and A. M. Dean. 2005. The biochemical architecture of an ancient adaptive landscape. Science 310:499-501.

Mayr, E. 1959. Where Are We? Cold Spring Harbor Symposium on Quantitative Biology 24:409-440.

Ogas, J., J. C. Cheng, Z. R. Sung, and C. Somerville. 1997. Cellular differentiation regulated by gibberellin in the Arabidopsis thaliana pickle mutant. Science 277:91-94.

Phillips, P. C. 1998. The language of gene interaction. Genetics 149:1167-1171.

—. 2008. Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet 9:855-867.

Powell, J. R. 1997, Progress and Prospects in Evolutionary Biology: The Drosophila Model. New York, Oxford University Press.

Provine, W. B. 1971, The Origins of Theoretical Population Genetics. Chicago, University of Chicago Press.

Raj, A., S. A. Rifkin, E. Andersen, and A. van Oudenaarden. 2010. Variability in gene expression underlies incomplete penetrance. Nature 463:913-918.

Rao, V., and V. Nanjundiah. 2010. J. B. S. Haldane, Ernst Mayr and the Beanbag Genetics Dispute. J Hist Biol.

Raser, J. M., and E. K. O'Shea. 2005. Noise in gene expression: origins, consequences, and control. Science 309:2010-2013.

Seidel, H. S., M. V. Rockman, and L. Kruglyak. 2008. Widespread Genetic Incompatibility in *C. elegans* Maintained by Balancing Selection. Science 319:589-594.

Shields, P. G., and C. C. Harris. 2000. Cancer risk and low-penetrance susceptibility genes in gene-environment interactions. J Clin Oncol 18:2309-2315.

Stark, A., M. F. Lin, P. Kheradpour, J. S. Pedersen, L. Parts, J. W. Carlson, M. A. Crosby et al. 2007. Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures. Nature 450:219-232.

Tweedie, S., M. Ashburner, K. Falls, P. Leyland, P. McQuilton, S. Marygold, G. Millburn et al. 2009. FlyBase: enhancing Drosophila Gene Ontology annotations. Nucleic Acids Res 37:D555-559.

Waddington, C. H. 1957, The Strategy of Genes. London, Allen & Unwin.

Wakeley, J. 2009, Coalescent theory : an introduction. Greenwood Village, Colo., Roberts & Co. Publishers.

Weinreich, D. M., N. F. Delaney, M. A. Depristo, and D. L. Hartl. 2006. Darwinian evolution can follow only very few mutational paths to fitter proteins. Science 312:111-114.

Weinreich, D. M., R. A. Watson, and L. Chao. 2005. Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. Evolution Int J Org Evolution 59:1165-1174.

Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proc. 6th International Congress of Genetics 1:356-366.

Wu, C. I., H. Hollocher, D. J. Begun, C. F. Aquadro, Y. Xu, and M. L. Wu. 1995. Sexual isolation in Drosophila melanogaster: a possible case of incipient speciation. Proc Natl Acad Sci U S A 92:2519-2523.

Yukilevich, R., J. Lachance, F. Aoki, and J. R. True. 2008. Long-term adaptation of epistatic genetic networks. Evolution 62:2215-2235.

Yukilevich, R., and J. R. True. 2008. Incipient sexual isolation among cosmopolitan Drosophila melanogaster populations. Evolution 62:2112-2121.

Chapter 2

Alcazar, R., A. V. Garcia, J. E. Parker, and M. Reymond. 2009. Incremental steps toward incompatibility revealed by Arabidopsis epistatic interactions modulating salicylic acid pathway activation. Proc Natl Acad Sci U S A 106:334-339.

Anholt, R. R., C. L. Dilda, S. Chang, J. J. Fanara, N. H. Kulkarni, I. Ganguly, S. M. Rollmann et al. 2003. The genetic architecture of odor-guided behavior in Drosophila: epistasis and the transcriptome. Nat Genet 35:180-184.

Bierne, N., F. Bonhomme, P. Boudry, M. Szulkin, and P. David. 2006. Fitness landscapes support the dominance theory of post-zygotic isolation in the mussels Mytilus edulis and M. galloprovincialis. Proc Biol Sci 273:1253-1260.

Bomblies, K., J. Lempe, P. Epple, N. Warthmann, C. Lanz, J. L. Dangl, and D. Weigel. 2007. Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. PLoS Biol 5:e236.

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

Caracristi, G., and C. Schlotterer. 2003. Genetic differentiation between American and European Drosophila melanogaster populations could be attributed to admixture of African alleles. Mol Biol Evol 20:792-799.

Carling, M. D., and R. T. Brumfield. 2008. Haldane's rule in an avian system: using cline theory and divergence population genetics to test for differential introgression of mitochondrial, autosomal, and sex-linked loci across the Passerina bunting hybrid zone. Evolution 62:2600-2615.

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

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

Dixon, S. J., Y. Fedyshyn, J. L. Koh, T. S. Prasad, C. Chahwan, G. Chua, K. Toufighi et al. 2008. Significant conservation of synthetic lethal genetic interaction networks between distantly related eukaryotes. Proc Natl Acad Sci U S A 105:16653-16658.

Dobzhansky, T. 1970, Genetics of the Evolutionary Process. New York, Columbia University Press.

Edmands, S. 2007. Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. Mol Ecol 16:463-475.

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

Good, J. M., M. A. Handel, and M. W. Nachman. 2008. Asymmetry and polymorphism of hybrid male sterility during the early stages of speciation in house mice. Evolution 62:50-65.

Gourbiere, S., and J. Mallet. 2009. Are Species Real? The Shape of the Species Boundary with Exponential Failure, Reinforcement, and the 'Missing Snowball'. Evolution.

Griswold, C. K., and J. Masel. 2009. Complex adaptations can drive the evolution of the capacitor [PSI], even with realistic rates of yeast sex. PLoS Genet 5:e1000517.

Haldane, J. B. S. 1922. Sex ratio and unisexual sterility in hybrid animals. J. Genet. 12:101-109.

Harrison, J. S., and S. Edmands. 2006. Chromosomal basis of viability differences in Tigriopus californicus interpopulation hybrids. J Evol Biol 19:2040-2051.

Henikoff, S., K. Ahmad, and H. S. Malik. 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. Science 293:1098-1102.

Hurst, L. D., and A. Pomiankowski. 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.

Johnson, N. 2000. Gene Interactions and the Origin of Species, Pages 197-212 *in* J. Wolf, E. Brodie, and M. Wade, eds. Epistasis and the Evolutionary Process. New York, Oxford University Press.

Kaplan, W. D., R. L. Seecof, W. E. Trout, and M. E. Pasternack. 1970. Production and Relative Frequency of Maternally Influenced Lethals in Drosophila melanogaster. American Naturalist 104:261-271.

Keller, A. 2007. Drosophila melanogaster's history as a human commensal. Curr Biol 17:R77-81.

Kim, Y. 2007. Rate of adaptive peak shifts with partial genetic robustness. Evolution 61:1847-1856.

Krimbas, C. B. 1960. Synthetic Sterility in Drosophila Willistoni. Proc Natl Acad Sci U S A 46:832-833.

Laurie, C. C. 1997. The weaker sex is heterogametic: 75 years of Haldane's rule. Genetics 147:937-951.

Lucchesi, J. C. 1968. Synthetic lethality and semi-lethality among functionally related mutants of Drosophila melanfgaster. Genetics 59:37-44.

Mank, J. E., B. Vicoso, S. Berlin, and B. Charlesworth. 2010. Effective population size and the faster-X effect: empirical results and their interpretation. Evolution 64:663-674.

Masly, J. P., and D. C. Presgraves. 2007. High-resolution genome-wide dissection of the two rules of speciation in Drosophila. PLoS Biol 5:e243.

Moyle, L. C., and T. Nakazato. 2009. Complex epistasis for Dobzhansky-Muller hybrid incompatibility in solanum. Genetics 181:347-351.

Orr, H. A. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics 139:1805-1813.

Palopoli, M. F., and C. I. Wu. 1994. Genetics of hybrid male sterility between drosophila sibling species: a complex web of epistasis is revealed in interspecific studies. Genetics 138:329-341.

Phillips, P. C. 2008. Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet 9:855-867.

Phillips, P. C., and N. A. Johnson. 1998. The population genetics of synthetic lethals. Genetics 150:449-458.

Pool, J. E., and C. F. Aquadro. 2006. History and structure of sub-Saharan populations of Drosophila melanogaster. Genetics 174:915-929.

Powell, J. 1997, Progress and Prospects in Evolutionary Biology: The Drosophila Model: Oxford Series in Ecology and Evolution. New York, Oxford University Press.

Presgraves, D. C. 2007. Speciation genetics: epistasis, conflict and the origin of species. Curr Biol 17:R125-127.

Proulx, S. R., and P. C. Phillips. 2005. The Opportunity for Canalization and the Evolution of Genetic Networks. American Naturalist 165:147-162.

Ranz, J. M., C. I. Castillo-Davis, C. D. Meiklejohn, and D. L. Hartl. 2003. Sex-dependent gene expression and evolution of the Drosophila transcriptome. Science 300:1742-1745.

Sanjuan, R., J. M. Cuevas, A. Moya, and S. F. Elena. 2005. Epistasis and the adaptability of an RNA virus. Genetics 170:1001-1008.

Schluter, D., and G. L. Conte. 2009. Genetics and ecological speciation. Proc Natl Acad Sci U S A 106 Suppl 1:9955-9962.

Sezgin, E., D. D. Duvernell, L. M. Matzkin, Y. Duan, C. T. Zhu, B. C. Verrelli, and W. F. Eanes. 2004. Single-locus latitudinal clines and their relationship to temperate adaptation in metabolic genes and derived alleles in Drosophila melanogaster. Genetics 168:923-931.

Shao, H., L. C. Burrage, D. S. Sinasac, A. E. Hill, S. R. Ernest, W. O'Brien, H. W. Courtland et al. 2008. Genetic architecture of complex traits: large phenotypic effects and pervasive epistasis. Proc Natl Acad Sci U S A 105:19910-19914.

Tao, Y., S. Chen, D. L. Hartl, and C. C. Laurie. 2003. Genetic dissection of hybrid incompatibilities between Drosophila simulans and D. mauritiana. I. Differential accumulation of hybrid male sterility effects on the X and autosomes. Genetics 164:1383-1397.

Tao, Y., D. L. Hartl, and C. C. Laurie. 2001. Sex-ratio segregation distortion associated with reproductive isolation in Drosophila. Proc Natl Acad Sci U S A 98:13183-13188.

Temin, R. G., H. U. Meyer, P. S. Dawson, and J. F. Crow. 1969. The influence of epistasis on homozygous viability depression in Drosophila melanogaster. Genetics 61:497-519.

Thompson, V. 1986. Synthetic lethals: a critical review. Evolutionary Theory 8:1-13.

Thonneau, P., S. Marchand, A. Tallec, M. L. Ferial, B. Ducot, J. Lansac, P. Lopes et al. 1991. Incidence and main causes of infertility in a resident population (1,850,000) of three French regions (1988-1989). Hum Reprod 6:811-816.

True, J. R., B. S. Weir, and C. C. Laurie. 1996. A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of Drosophila mauritiana chromosomes into Drosophila simulans. Genetics 142:819-837.

Turelli, M., and H. A. Orr. 1995. The dominance theory of Haldane's rule. Genetics 140:389-402.

—. 2000. Dominance, epistasis and the genetics of postzygotic isolation. Genetics 154:1663-1679.

Umina, P. A., A. R. Weeks, M. R. Kearney, S. W. McKechnie, and A. A. Hoffmann. 2005. A rapid shift in a classic clinal pattern in Drosophila reflecting climate change. Science 308:691-693.

Vicoso, B., and B. Charlesworth. 2006. Evolution on the X chromosome: unusual patterns and processes. Nat Rev Genet 7:645-653.

Watanabe, T. K., and W. H. Lee. 1977. Sterile mutation in *Drosophila melanogaster*. Genetical Research 30:107-113.

Wolf, J., E. Brodie, and M. J. Wade. 2000. Epistasis and the Evolutionary Process. New York, Oxford University Press.

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

Wu, C. I., H. Hollocher, D. J. Begun, C. F. Aquadro, Y. Xu, and M. L. Wu. 1995. Sexual isolation in Drosophila melanogaster: a possible case of incipient speciation. Proc Natl Acad Sci U S A 92:2519-2523.

Wu, C. I., N. A. Johnson, and M. F. Palopoli. 1996. Haldane's rule and its legacy: Why are there so many sterile males? TREE 11:281-284.

Wu, C. I., and C. T. Ting. 2004. Genes and speciation. Nat Rev Genet 5:114-122.

Yukilevich, R., J. Lachance, F. Aoki, and J. R. True. 2008. Long-term adaptation of epistatic genetic networks. Evolution 62:2215-2235.

Yukilevich, R., and J. R. True. 2008a. African morphology, behavior and phermones underlie incipient sexual isolation between us and Caribbean Drosophila melanogaster. Evolution 62:2807-2828.

—. 2008b. Incipient sexual isolation among cosmopolitan Drosophila melanogaster populations. Evolution 62:2112-2121.

Chapter 3

Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne, P. G. Amanatides, S. E. Scherer et al. 2000. The genome sequence of Drosophila melanogaster. Science 287:2185-2195.

Bagheri, H. C. 2006. Unresolved boundaries of evolutionary theory and the question of how inheritance systems evolve: 75 years of debate on the evolution of dominance. J Exp Zool B Mol Dev Evol 306:329-359.

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

Bengtsson, B. O., and F. B. Christiansen. 1983. A Two-Locus Mutation-Selection Model and Some of Its Evolutionary Implications. Theoretical Population Biology 24:59-77.

Cabot, E. L., A. W. Davis, N. A. Johnson, and C. I. Wu. 1994. Genetics of reproductive isolation in the Drosophila simulans clade: complex epistasis underlying hybrid male sterility. Genetics 137:175-189.

Charlesworth, B., and D. Charlesworth. 2010, Elements of Evolutionary Genetics. Greenwood Village, Colorado, Roberts and Company Publishers.

Charlesworth, B., J. Coyne, and N. H. Barton. 1987. The Relative Rates of Evolution of Sex Chromosomes and Autosomes. Am Nat 130:113-146.

Christiansen, F. B., and O. Frydenberg. 1977. Selection-mutation balance for two nonallelic recessives producing an inferior double homozygote. Am J Hum Genet 29:195-207.

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

Crow, J. F., and M. Kimura. 1970, An introduction to populaiton genetics theory. New York, Harper and Row.

Davierwala, A. P., J. Haynes, Z. Li, R. L. Brost, M. D. Robinson, L. Yu, S. Mnaimneh et al. 2005. The synthetic genetic interaction spectrum of essential genes. Nat Genet 37:1147-1152.

Davis, A. W., E. G. Noonburg, and C. I. Wu. 1994. Evidence for complex genic interactions between conspecific chromosomes underlying hybrid female sterility in the Drosophila simulans clade. Genetics 137:191-199.

de Visser, J. A., and S. F. Elena. 2007. The evolution of sex: empirical insights into the roles of epistasis and drift. Nat Rev Genet 8:139-149.

Dobzhansky, T. 1936. Studies on Hybrid Sterility. II. Localization of Sterility Factors in Drosophila Pseudoobscura Hybrids. Genetics 21:113-135.

—. 1937, Genetics and the Origin of Species. New York, Columbia University Press.

—. 1946. Genetics of Natural Populations. Xiii. Recombination and Variability in Populations of Drosophila Pseudoobscura. Genetics 31:269-290.

Freeman, J. L., G. H. Perry, L. Feuk, R. Redon, S. A. McCarroll, D. M. Altshuler, H. Aburatani et al. 2006. Copy number variation: new insights in genome diversity. Genome Res 16:949-961.

Gavrilets, S. 2004, Fitness Landscapes and the Origin of Species. Princeton, NJ, Princeton University Press.

Haldane, J. B. S. 1922. Sex ratio and unisexual sterility in hybrid animals. J. Genet. 12:101-109.

Hartman, J. L., B. Garvick, and L. Hartwell. 2001. Principles for the Buffering of Genetic Variation. Science 291:1001-1004.

Johnson, N. A. 2000. Speciation: Dobzhansky-Muller incompatibilities, dominance and gene interactions. Trends Ecol Evol 15:480-482.

—. 2010. Hybrid incompatibility genes: remnants of a genomic battlefield? Trends Genet 26:317-325.

Lachance, J., and J. R. True. 2010. X-autosome incompatibilities in *Drosophila melanogaster*: tests of Haldane's rule and geographic patterns within species. Evolution 64:3035-3046.

Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon et al. 2001. Initial sequencing and analysis of the human genome. Nature 409:860-921.

Laurie, C. C. 1997. The weaker sex is heterogametic: 75 years of Haldane's rule. Genetics 147:937-951.

Lehner, B., C. Crombie, J. Tischler, A. Fortunato, and A. G. Fraser. 2006. Systematic mapping of genetic interactions in Caenorhabditis elegans identifies common modifiers of diverse signaling pathways. Nat Genet 38:896-903.

Liao, C. Y., P. Wu, B. Hu, and K. K. Yi. 2001. Effects of genetic background and environment on QTLs and epistasis for rice (*Oryza sativa* L.) panicle number. Theor Appl Genet 103:104-111.

Lucchesi, J. C. 1968. Synthetic lethality and semi-lethality among functionally related mutants of Drosophila melanfgaster. Genetics 59:37-44.

Mackay, T. F. 2004. The genetic architecture of quantitative traits: lessons from Drosophila. Curr Opin Genet Dev 14:253-257.

Mank, J. E., B. Vicoso, S. Berlin, and B. Charlesworth. 2010. Effective population size and the faster-X effect: empirical results and their interpretation. Evolution 64:663-674.

Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff, D. J. Hunter, M. I. McCarthy et al. 2009. Finding the missing heritability of complex diseases. Nature 461:747-753.

Mathworks. 2005.MATLAB 7.The Mathworks, Natick, MA.

Moore, J. H. 2003. The ubiquitous nature of epistasis in determining susceptibility to common human diseases. Hum Hered 56:73-82.

Moore, J. H., F. W. Asselbergs, and S. M. Williams. 2010. Bioinformatics challenges for genome-wide association studies. Bioinformatics 26:445-455.

Muller, H. J. 1942. Isolating mechanisms, evolution and temperature. Biol. Symp. 6:71-125.

Nadeau, J. H. 2001. Modifier genes in mice and humans. Nat Rev Genet 2:165-174.

Ohta, T. 1973. Slightly deleterious mutant substitutions in evolution. Nature 246:96-98.

Ooi, S. L., X. Pan, B. D. Peyser, P. Ye, P. B. Meluh, D. S. Yuan, R. A. Irizarry et al. 2006. Global synthetic-lethality analysis and yeast functional profiling. Trends Genet 22:56-63.

Orr, H. A. 1993. A Mathematical Model of Haldane's Rule. Evolution 47:1606-1611.

—. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics 139:1805-1813.

Orr, H. A., and S. Irving. 2001. Complex epistasis and the genetic basis of hybrid sterility in the Drosophila pseudoobscura Bogota-USA hybridization. Genetics 158:1089-1100.

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

Otto, S. P., and T. Lenormand. 2002. Resolving the paradox of sex and recombination. Nat Rev Genet 3:252-261.

Phillips, P. C. 2008. Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet 9:855-867.

Phillips, P. C., and N. A. Johnson. 1998. The population genetics of synthetic lethals. Genetics 150:449-458.

Presgraves, D. C. 2003. A fine-scale genetic analysis of hybrid incompatibilities in Drosophila. Genetics 163:955-972.

—. 2010. The molecular evolutionary basis of species formation. Nat Rev Genet 11:175-180.

Sanjuan, R., A. Moya, and S. F. Elena. 2004. The contribution of epistasis to the architecture of fitness in an RNA virus. Proc Natl Acad Sci U S A 101:15376-15379.

Tao, Y., S. Chen, D. L. Hartl, and C. C. Laurie. 2003. Genetic dissection of hybrid incompatibilities between Drosophila simulans and D. mauritiana. I. Differential accumulation of hybrid male sterility effects on the X and autosomes. Genetics 164:1383-1397.

Temin, R. G., H. U. Meyer, P. S. Dawson, and J. F. Crow. 1969. The influence of epistasis on homozygous viability depression in Drosophila melanogaster. Genetics 61:497-519.

Templeton, A. R. 2000. Epistasis and Complex Traits *in* J. B. Wolf, E. D. Brodie, and M. J. Wade, eds. Epistasis and the Evolutionary Process. New York, Oxford University Press.

Thompson, V. 1986. Synthetic lethals: a critical review. Evolutionary Theory 8:1-13.

Tong, A. H., G. Lesage, G. D. Bader, H. Ding, H. Xu, X. Xin, J. Young et al. 2004. Global mapping of the yeast genetic interaction network. Science 303:808-813.

True, J. R., B. S. Weir, and C. C. Laurie. 1996. A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of Drosophila mauritiana chromosomes into Drosophila simulans. Genetics 142:819-837.

Turelli, M., and D. J. Begun. 1997. Haldane's rule and X-chromosome size in Drosophila. Genetics 147:1799-1815.

Turelli, M., and H. A. Orr. 1995. The dominance theory of Haldane's rule. Genetics 140:389-402.

—. 2000. Dominance, epistasis and the genetics of postzygotic isolation. Genetics 154:1663-1679.

Venter, J. C., M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith et al. 2001. The sequence of the human genome. Science 291:1304-1351.

Vicoso, B., and B. Charlesworth. 2006. Evolution on the X chromosome: unusual patterns and processes. Nat Rev Genet 7:645-653.

Wade, M. J. 2001. Epistasis, complex traits, and mapping genes. Genetica 112-113:59-69.

Wang, R. X., and Y. L. Zhao. 2008. Differential barrier strength and allele frequencies in hybrid zones maintained by sex-biased hybrid incompatibilities. Heredity 100:326-336.

Whitlock, M. C., and M. J. Wade. 1995. Speciation: founder events and their effects on X-linked and autosomal genes. Am Nat 145:676-685.

Wolf, J., E. Brodie, and M. J. Wade. 2000. Epistasis and the Evolutionary Process. New York, Oxford University Press.

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

Chapter 4

Agresti, A., and B. A. Coull. 1998. Approximate is better than "exact" for interval estimation of binomial proportions. American Statistician 52:119-126.

Antonarakis, S. E., A. Chakravarti, J. C. Cohen, and J. Hardy. 2010. Mendelian disorders and multifactorial traits: the big divide or one for all? Nat Rev Genet 11:380-384.

Aruna, S., H. A. Flores, and D. A. Barbash. 2009. Reduced fertility of Drosophila melanogaster hybrid male rescue (Hmr) mutant females is partially complemented by Hmr orthologs from sibling species. Genetics 181:1437-1450.

Atallah, J., I. Dworkin, U. Cheung, A. Greene, B. Ing, L. Leung, and E. Larsen. 2004. The environmental and genetic regulation of obake expressivity: morphogenetic fields as evolvable systems. Evol Dev 6:114-122.

Ayroles, J. F., M. A. Carbone, E. A. Stone, K. W. Jordan, R. F. Lyman, M. M. Magwire, S. M. Rollmann et al. 2009. Systems genetics of complex traits in Drosophila melanogaster. Nat Genet 41:299-307.

Barrett, R. D., and D. Schluter. 2008. Adaptation from standing genetic variation. Trends Ecol Evol 23:38-44.

Benfey, P. N., and T. Mitchell-Olds. 2008. From genotype to phenotype: systems biology meets natural variation. Science 320:495-497.

Bjornsson, H. T., M. D. Fallin, and A. P. Feinberg. 2004. An integrated epigenetic and genetic approach to common human disease. Trends Genet 20:350-358.

Brabant, M. C., D. Fristrom, T. A. Bunch, and D. L. Brower. 1996. Distinct spatial and temporal functions for PS integrins during Drosophila wing morphogenesis. Development 122:3307-3317.

Brown, N. H., S. L. Gregory, and M. D. Martin-Bermudo. 2000. Integrins as mediators of morphogenesis in Drosophila. Dev Biol 223:1-16.

Carey, C. C., K. F. Gorman, and S. Rutherford. 2006. Modularity and intrinsic evolvability of Hsp90-buffered change. PLoS One 1:e76.

Carlborg, O., and C. S. Haley. 2004. Epistasis: too often neglected in complex trait studies? Nat Rev Genet 5:618-625.

de Visser, J. A., J. Hermisson, G. P. Wagner, L. Ancel Meyers, H. Bagheri-Chaichian, J. L. Blanchard, L. Chao et al. 2003. Perspective: Evolution and detection of genetic robustness. Evolution 57:1959-1972.

Donehower, L. A., M. Harvey, H. Vogel, M. J. McArthur, C. A. Montgomery, Jr., S. H. Park, T. Thompson et al. 1995. Effects of genetic background on tumorigenesis in p53-deficient mice. Mol Carcinog 14:16-22.

Dworkin, I., E. Kennerly, D. Tack, J. Hutchinson, J. Brown, J. Mahaffey, and G. Gibson. 2009. Genomic consequences of background effects on scalloped mutant expressivity in the wing of Drosophila melanogaster. Genetics 181:1065-1076.

Dworkin, I., A. Palsson, K. Birdsall, and G. Gibson. 2003. Evidence that Egfr contributes to cryptic genetic variation for photoreceptor determination in natural populations of Drosophila melanogaster. Curr Biol 13:1888-1893.

Evang. 1925. The sex-linked mutants vesiculated and semi-lethal in Drosophila melanogaster. Z Indukt Abstamm Vererbungsl 39:165-183.

Gibson, G., and I. Dworkin. 2004. Uncovering cryptic genetic variation. Nat Rev Genet 5:681-690.

Goldstein, L. S., and S. Gunawardena. 2000. Flying through the drosophila cytoskeletal genome. J Cell Biol 150:F63-68.

Grieder, N. C., I. Charlafti, U. Kloter, H. Jackle, U. Schafer, and W. J. Gehring. 2007. Misexpression screen in Drosophila melanogaster aiming to reveal novel factors involved in formation of body parts. Genetics 175:1707-1718.

Hansen, T. 2006. The Evolution of Genetic Architecture. Annu Rev Ecol Syst 37:123-157.

Hirschhorn, J. N., and M. J. Daly. 2005. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 6:95-108.

Javierre, B. M., A. F. Fernandez, J. Richter, F. Al-Shahrour, J. I. Martin-Subero, J. Rodriguez-Ubreva, M. Berdasco et al. 2010. Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. Genome Res 20:170-179.

Johnson, S. A., and M. J. Milner. 1987. The final stages of wing development in Drosophila melanogaster. Tissue Cell 19:505-513.

Judd, B. H., M. W. Shen, and T. C. Kaufman. 1972. The anatomy and function of a segment of the X chromosome of Drosophila melanogaster. Genetics 71:139-156.

Kacser, H., and J. A. Burns. 1981. The Molecular Basis of Dominance. Genetics:639-666.

Kiger, J. A., Jr., J. E. Natzle, D. A. Kimbrell, M. R. Paddy, K. Kleinhesselink, and M. M. Green. 2007. Tissue remodeling during maturation of the Drosophila wing. Dev Biol 301:178-191.

King, R. C., W. D. Stansfield, and P. K. Mulligan. 2006, A dictionary of genetics. Oxford, Oxford University Press.

Kitano, H. 2004. Biological robustness. Nat Rev Genet 5:826-837.

Kondrashov, F. A., and E. V. Koonin. 2004. A common framework for understanding the origin of genetic dominance and evolutionary fates of gene duplications. Trends Genet 20:287-290.

Lachance, J., and J. R. True. 2010. X-autosome incompatibilities in *Drosophila melanogaster*: tests of Haldane's rule and geographic patterns within species. Evolution 64:3035-3046.

Laubichler, M. D., and S. Sarkar. 2002. Flies, genes, and brains: Oskar Vogt, Nikolai Timofeeff-Ressovsky, and the origin of the concepts of penetrance and expressivity *in* L. S. Parker, and R. A. Ankeny, eds. Mutating concepts, evolving disciplines: genes, medicine, and society. Philosophy and Medicine. Dordrecht, Kluvier.

Lewontin, R. C. 2000, The Triple Helix: Gene, Organism and Environment. Cambridge, Mass., Harvard University Press.

Lindsley, D. L., and G. G. Zimm. 1992, The Genome of *Drosophila melanogaster*. San Diego, CA, Academic Press.

Lukacsovich, T., Z. Asztalos, W. Awano, K. Baba, S. Kondo, S. Niwa, and D. Yamamoto. 2001. Dual-tagging gene trap of novel genes in Drosophila melanogaster. Genetics 157:727-742.

Lunzer, M., S. P. Miller, R. Felsheim, and A. M. Dean. 2005. The biochemical architecture of an ancient adaptive landscape. Science 310:499-501.

Masel, J., and M. L. Siegal. 2009. Robustness: mechanisms and consequences. Trends Genet 25:395-403.

Mathworks. 2005.MATLAB 7.Mathworks, Natick, MA.

Mezey, J. G., D. Houle, and S. V. Nuzhdin. 2005. Naturally segregating quantitative trait loci affecting wing shape of Drosophila melanogaster. Genetics 169:2101-2113.

Pavelka, J., A. M. Kulikov, and F. Marec. 1996. Suppression of the Drosophila Curly mutation by fluorescent light. Hereditas 124:191-197.

Phillips, P. C. 2008. Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet 9:855-867.

Pierpont, M. E., C. T. Basson, D. W. Benson, Jr., B. D. Gelb, T. M. Giglia, E. Goldmuntz, G. McGee et al. 2007. Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. Circulation 115:3015-3038.

Polaczyk, P. J., R. Gasperini, and G. Gibson. 1998. Naturally occurring genetic variation affects Drosophila photoreceptor determination. Dev Genes Evol 207:462-470.

Raj, A., S. A. Rifkin, E. Andersen, and A. van Oudenaarden. 2010. Variability in gene expression underlies incomplete penetrance. Nature 463:913-918.

Raser, J. M., and E. K. O'Shea. 2005. Noise in gene expression: origins, consequences, and control. Science 309:2010-2013.

Rockman, M. V. 2008. Reverse engineering the genotype-phenotype map with natural genetic variation. Nature 456:738-744.

Ruden, D. M., and X. Lu. 2008. Hsp90 affecting chromatin remodeling might explain transgenerational epigenetic inheritance in Drosophila. Curr Genomics 9:500-508.

Schmalhausen, I. I. 1949, Factors of Evolution. The theory of Stabilizing selection. Chicago, University of Chicago Press.

Shields, P. G., and C. C. Harris. 2000. Cancer risk and low-penetrance susceptibility genes in gene-environment interactions. J Clin Oncol 18:2309-2315.

Stern, D. L. 2010, Evolution, development, & the predictable genome. Greenwood Village, Colo., Roberts And Company Publishers.

Togel, M., G. Pass, and A. Paululat. 2008. The Drosophila wing hearts originate from pericardial cells and are essential for wing maturation. Dev Biol 318:29-37.

Tweedie, S., M. Ashburner, K. Falls, P. Leyland, P. McQuilton, S. Marygold, G. Millburn et al. 2009. FlyBase: enhancing Drosophila Gene Ontology annotations. Nucleic Acids Res 37:D555-559.

Waddington, C. H. 1939. Preliminary Notes on the Development of the Wings in Normal and Mutant Strains of Drosophila. Proc Natl Acad Sci U S A 25:299-307.

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.

Wagner, A. 2005, Robustness and Evolvability in Living Systems. Princeton, NJ, Princeton University Press.

Weinreich, D. M., R. A. Watson, and L. Chao. 2005. Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. Evolution 59:1165-1174.

Wilkie, A. O. 1994. The molecular basis of genetic dominance. J Med Genet 31:89-98.

Youngson, N. A., and E. Whitelaw. 2008. Transgenerational epigenetic effects. Annu Rev Genomics Hum Genet 9:233-257.

Zimmerman, E., A. Palsson, and G. Gibson. 2000. Quantitative trait loci affecting components of wing shape in Drosophila melanogaster. Genetics 155:671-683.

Chapter 5

Bejjani, B. A., D. W. Stockton, R. A. Lewis, K. F. Tomey, D. K. Dueker, M. Jabak, W. F. Astle et al. 2000. Multiple CYP1B1 mutations and incomplete penetrance in an inbred population segregating primary congenital glaucoma suggest frequent de novo events and a dominant modifier locus. Hum Mol Genet 9:367-374.

Benfey, P. N., and T. Mitchell-Olds. 2008. From genotype to phenotype: systems biology meets natural variation. Science 320:495-497.

Berger, S. L., T. Kouzarides, R. Shiekhattar, and A. Shilatifard. 2009. An operational definition of epigenetics. Genes Dev 23:781-783.

Crow, J. F., and M. Kimura. 1970, An introduction to populaiton genetics theory. New York, Harper and Row.

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

Dowell, R. D., O. Ryan, A. Jansen, D. Cheung, S. Agarwala, T. Danford, D. A. Bernstein et al. 2010. Genotype to phenotype: a complex problem. Science 328:469.

Ewens, W. J. 2004, Mathematical population genetics: Interdisciplinary applied mathematics v. 27. New York, Springer.

Feller, W. 1951, Diffusion processes in genetics. S. Neyman, ed. Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability:227-246.

Felsenstein, J. 1976. The theoretical population genetics of variable selection and migration. Annu Rev Genet 10:253-280.

Gergen, J. P., and E. F. Wieschaus. 1986. Localized requirements for gene activity in segmentation of *Drosophila* embryos: analysis of *armadillo*, *fused*, *giant* and *unpaired* mutations in mosaic embryos. Development Genes and Evolution 195:49-62.

Gillespie, J. H. 1974. Nautural selection for within-generation variance in offspring number. Genetics 76:601-606.

—. 1977. Natural selection for variance in offpspring number: a new evolutionary principle. American Naturalist 111:1010-1014.

—. 2001. Is the population size of a species relevant to its evolution? Evolution Int J Org Evolution 55:2161-2169.

—. 2004, Population Genetics: A Concise Guide. Baltimore, The Johns Hopkins University Press.

—. 2006. Stochastic Processes in Evolution *in* C. W. Fox, and J. B. Wolf, eds. Evolutionary Genetics: Concepts and Case Studies. Oxford, UK, Oxford University Press.

Haldane, J. B. S., and S. D. Jayakar. 1963. Polymorphism due to selection of varying direction. J. Genet. 58:237-242.

Hartl, D. L., and R. D. Cook. 1974. Autocorrelated Random Environments and Their Effects on Gene Frequency. Evolution 28:275-280.

Heino, M., J. A. J. Metz, and V. Kaitala. 1998. The enigma of frequency-dependent selection. Trends Ecol. Evol. 13:367-370.

Hollocher, H., A. R. Templeton, R. DeSalle, and J. S. Johnston. 1992. The molecular through ecological genetics of abnormal abdomen. IV. Components of genetic variation in a natural population of Drosophila mercatorum. Genetics 130:355-366.

Johnson, N. A., and A. H. Porter. 2001. Toward a new synthesis: population genetics and evolutionary developmental biology. Genetica 112/113:45-58.

Kærn, M., T. Elston, W. Blake, and J. Collins. 2005. Stochasticity in Gene Expression: from Theories to Phenotypes. Nature Reviews Genetics 6:451-464.

Karlin, S., and B. Levikson. 1974. Temporal fluctuations in selection intensities: case of small selection size. Theor Popul Biol 63:383-412.

Kim, Y. 2007. Rate of Adaptive Peak Shifts with Partial Genetic Robustness. Evolution 61:1847-1856.

Kimura, M. 1962. On the probability of fixation of mutant genes in a population. Genetics 47:713-719.

—. 1964. Diffusion models in population genetics. J. Appl. Prob. 1:177-232.

Kimura, M., and J. F. Crow. 1963. The measurement of effective population number. Evolution 17:279-288.

Kimura, M., and T. Ohta. 1969a. The Average Number of Generations until Extinction of an Individual Mutant Gene in a Finite Population. Genetics 61:701-709.

—. 1969b. The Average Number of Generations until Fixation of a Mutant Gene in a Finite Population. Genetics 61:763-771.

King, R. C., W. D. Stansfield, and P. K. Mulligan. 2006, A dictionary of genetics. Oxford, Oxford University Press.

Lalucque, H., and P. Silar. 2004. Incomplete penetrance and variable expressivity of a growth defect as a consequence of knocking out two K(+) transporters in the euascomycete fungus Podospora anserina. Genetics 166:125-133.

Lande, R. 2007. Expected Relative Fitness and the Adaptive Topography of Fluctuating Selection. Evolution 61:1835-1846.

Laubichler, M. D., and S. Sarkar. 2002. Flies, genes, and brains: Oskar Vogt, Nikolai Timofeeff-Ressovsky, and the origin of the concepts of penetrance and expressivity *in* L. S. Parker, and R. A. Ankeny, eds. Mutating concepts, evolving disciplines: genes, medicine, and society. Philosophy and Medicine. Dordrecht, Kluvier.

Legeai-Mallet, L., A. Munnich, P. Maroteaux, and M. Le Merrer. 1997. Incomplete penetrance and expressivity skewing in hereditary multiple exostoses. Clin Genet 52:12-16.

Levene, H. 1953. Genetic Equilibrium When More Than One Ecological Niche is Available. American Naturalist 87:331-333.

Lewontin, R. C. 2000, The Triple Helix: Gene, Organism and Environment. Cambridge, Mass., Harvard University Press.

Mackay, T. F. 2001. The genetic architecture of quantitative traits. Annu Rev Genet 35:303-339.

Mathworks. 2005.MATLAB 7.The Mathworks, Natick, MA.

Mayr, E. 2001, What evolution is. New York, Basic Books.

Ogas, J., J. C. Cheng, Z. R. Sung, and C. Somerville. 1997. Cellular differentiation regulated by gibberellin in the Arabidopsis thaliana pickle mutant. Science 277:91-94.

Okasha, S. 2004. The "averaging fallacy" and the levels of selection. Biology and Philosophy 19:167-184.

Orr, H. A. 2007. Absolute Fitness, Relative Fitness, and Utility. Evolution 61:2997-2300.

Orr, H. A., and A. J. Betancourt. 2001. Haldane's sieve and adaptation from the standing genetic variation. Genetics 157:875-884.

Proulx, S. R. 2000. The ESS under Spatial Variation with Applications to Sex Allocation. Theor Popul Biol 58:33-47.

Proulx, S. R., and F. R. Adler. 2010. The standard of neutrality: still flapping in the breeze? J Evol Biol 23:1339-1350.

Proulx, T., and T. Day. 2001. What can invasion analyses tell us about evolution under stochasticity in finite populations? Selection 2:1-15.

Raj, A., S. A. Rifkin, E. Andersen, and A. van Oudenaarden. 2010. Variability in gene expression underlies incomplete penetrance. Nature 463:913-918.

Rasanen, K., and E. B. Kruuk. 2007. Maternal effects and evolution at ecological time-scales. Functional Ecology 21:408-421.

Rice, S. H. 2008. A stochastic version of the Price equation reveals the interplay of deterministic and stochastic processes in evolution. BMC Evol Biol 8:262.

Rockman, M. V. 2008. Reverse engineering the genotype-phenotype map with natural genetic variation. Nature 456:738-744.

Ruel, J. J., and M. P. Ayres. 1999. Jensen's inequality predicts effects of environmental variation. Trends Ecol Evol 14:361-366.

Salathe, M., J. Van Cleve, and M. W. Feldman. 2009. Evolution of stochastic switching rates in asymmetric fitness landscapes. Genetics 182:1159-1164.

Seger, J., and H. J. Brockman. 1987. What is bet-hedging?, Pages 182-211 *in* P. H. Harvey, and L. Partridge, eds. Oxford surveys in evolutionary biology. New York, Oxford University Press.

Shields, P. G., and C. C. Harris. 2000. Cancer risk and low-penetrance susceptibility genes in gene-environment interactions. J Clin Oncol 18:2309-2315.

Shpak, M. 2005. Evolution of Variance in Offspring Number: the Effects of Population Size and Migration. Theory in Bioscience 124:65-85.

—. 2007. Selection Against Demographic Stochasticity in Age-Structured Populations. Genetics 177:2181-2194.

Stern, D. L. 2010, Evolution, development, & the predictable genome. Greenwood Village, Colo., Roberts And Company Publishers.

Takahata, N., K. Ishii, and H. Matsuda. 1975. Effect of temporal fluctuation of selection coefficient on gene frequency in a population. PNAS 72:4541-4145.

Taylor, J. E. 2009. The genealogical consequences of fecundity variance polymorphism. Genetics 182:813-837.

Templeton, A. R. 1977. Survival Probabilities of Mutant Alleles in Fine-Grained Environments. American Naturalist 111:951-966.

Trotter, M. V., and H. G. Spencer. 2008. The generation and maintenance of genetic variation by frequency-dependent selection: constructing polymorphisms under the pairwise interaction model. Genetics 180:1547-1557.

Turner, J. R. G. 1981. Adaptation and evolution in *Heliconius*: a defense of NeoDarwinism. Annu Rev Ecol Syst 12:99-121.

Van Dyken, J. D., and M. J. Wade. 2010. The genetic signature of conditional expression. Genetics 184:557-570.

Veening, J. W., W. K. Smits, and O. P. Kuipers. 2008. Bistability, epigenetics, and bet-hedging in bacteria. Annu Rev Microbiol 62:193-210.

Waddington, C. H. 1942. The canalization of development and the inheritance of acquired characters. Nature 150:563-565.

Wade, J. D. 1998. The evolutionary genetics of maternal effects, Pages 5 *in* T. A. Mousseau, and C. W. Fox, eds. Maternal effects as adaptations. New York, Oxford Univ. Press.

Wakeley, J. 2005. The limits of theoretical population genetics. Genetics 169:1-7.

—. 2009, Coalescent theory : an introduction. Greenwood Village, Colo., Roberts & Co. Publishers.

Whitlock, M. C., and R. Gomulkiewicz. 2005. Probability of Fixation in a Heterogeneous Environment. Genetics 171:1407-1417.

Wolf, J. B. 2000. Gene interactions from maternal effects. Evolution 54:1882-1898.

Wright, S. 1938. Size of population and breeding structure in relation to evolution. Science 87:430-431.

Wu, R., and M. Lin. 2006. Functional mapping - how to map and study the genetic architecture of dynamic complex traits. Nat Rev Genet 7:229-237.

Youngson, N. A., and E. Whitelaw. 2008. Transgenerational epigenetic effects. Annu Rev Genomics Hum Genet 9:233-257.