EVOLUTION OF SEX GENES AND SPECIATION IN DROSOPHILA

THE MOLECULAR EVOLUTION OF SEX AND REPRODUCTION RELATED GENES, HYBRID MALE STERILITY AND SPECIATION IN THE *DROSOPHILA MELANOGASTER* COMPLEX

By

LI XU, B.Sc., M.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

Master of Science (2003) McMaster University

(Biology) Hamilton, Ontario

TITLE: The Molecular Evolution of Sex and Reproduction Related Genes, Hybrid Male Sterility and Speciation in the *Drosophila melanogaster* Complex

AUTHOR: Li Xu, B.Sc., M.Sc. (Xiamen University, Xiamen)

SUPERVISOR: Professor R.S. Singh

NUMBER OF PAGES: xi, 115

ABSTRACT

Haldane's rule, which states that the heterogametic sex is preferentially afflicted if one of hybrid sexes is sterile or inviable, is a general pattern in all animals that possess sex chromosomes. The hybrid sterility component of this rule is especially important because hybrid sterility is involved in the onset of postzygotic isolation. Accumulating evidence on the fast evolution of individual sex genes have stimulated us to hypothesize that the fast evolution of sex genes may be the force underlying the excess of hybrid heterogametic sterility. This study tests the evolutionary patterns of sex genes in comparison to non-sex genes, as a general group. The divergences between a group of 19 sex genes and 20 non-sex genes from X chromosome were compared between *D. melanogaster*, *D. mauritiana*, *D. simulans,* and *D. sechellia* using PCR-RFLP. Within species polymorphism data were also obtained for *D. simulans* and *D. mauritiana*. The results show a significantly higher divergence for sex genes than non-sex genes, while a comparable level of intraspecific polymorphism was revealed in both groups. Among the sex gene group, genes related to male reproduction appear to evolve faster than female-reproductive genes. The evolution of both sex and non-sex genes conforms to the neutral theory under Tajima's test and HKA test. The faster evolution of sex genes supports the fast-sex theory as an explanation for the hybrid sterility component of Haldane's rule.

Localization of some examples of hybrid sterility genes is crucial to ultimately untangle the genetics of hybrid sterility. The *car* region of *D. mauritiana*, which has been

shown to harbor genes that confer full effect of hybrid sterility in the *D. simulans* genetic background, was introgressed into the *D. simulans* genome by continuous backcrosses. Recombination mapping analysis, taking advantage of molecular markers, revealed that at least two regions are capable of causing hybrid sterility in this species group. The phenotypes of hybrid testes were examined during the backcross process.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisor Dr. Rama Singh. From his welcoming me here to the imparting of his immense wealth of knowledge, it has been a wonderful and stimulating two years of learning experience for me. This thesis has largely been guided by his patience, support and above all his warm smile.

Skeptical scrutiny is the means, in both science and religion, by which deep insights can be winnowed from deep emptiness. I am deeply indebted to Dr. Richard Morton a man of few words, for being that source of inspiration. His keen eye and timely advise were like that proverbial beacon of light. A special thanks to Dr. J. P. Xu whose infectious energy and livewire suggestions were a great help to me all the time especially so when I was new to this country.

It was indeed my privilege to have worked with my colleague Dr. Rob Kulathinal. He was largely responsible for tiding me over those early uncertain waters. Hope you are having fun at Harvard Rob. My lab-mates, Aaron Thomson, Santosh Jagadeeshan and Dara Torgerson also supported me in my research work. I would like to thank all of them for all their help, support and valuable suggestions. Heidi Musters and Sanjay Hiremath just joined us. I thank them for creating a pleasant atmosphere for me in the lab.

Last but definitely not the least, I would like to say a very warm thank you to my parents and to my husband Sheng Sun for their patient love and encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES AND APPENDICES

CHAPTER 1

GENERAL INTRODUCTION

Speciation is a fundamental evolutionary process and has resulted in an almost unlimited biological diversity of life on earth. In the process of speciation, reproductive isolation develops as a result of adaptation and genetic differentiation between different groups of organism. According to the Biological Species Concept, new species arise when reproductive barriers prevent effective exchange of genetic matter between potentially interbreeding individuals. Such reproductive barriers have been categorized into premating isolation and postmating isolation according to whether the gene exchange is blocked before or after the formation of hybrid zygotes. In prezygotic isolation, mating or fertilization can not successfully occur between species due to isolating mechanisms such as courtship differences, while postzygotic isolation is manifested through reduced fitness of hybrids, namely, hybrid sterility or inviability. The evolution of prezygotic isolation can be explained by the selection to avoid hybrid matings that produce inferior progeny (Kelly and Noor 1996). By contrast, the evolution of postzygotic isolation is not so clearly explained since it seems to possess no obvious adaptive value.

1.1 Models for the Genetic Basis of Postmating Reproductive Isolation

The evolution of such maladaptive traits as hybrid sterility or inviability makes one wonder that how natural selection could allow the evolution of progeny that are

sterile or inviable? The model proposed by Dobzhansky (1937) and Muller (1942) provided a general picture of the evolution of postzygotic isolation. The "Dobzhansky-Muller model" of speciation forms the basis of all work on the genetics of hybrid sterility and inviability. In this model, they proposed that postzygotic isolation might arise from interactions between two or more loci. Assume two allopatric populations initially have identical two-locus genotype, *AABB*. In one population, mutation *a* replaces allele *A* and goes to fixation. Within this population, *aaBB* or *AaBB*, which is fertile and viable, is equally fit as *AABB* by natural selection. In another population, a *b* mutation goes to fixation and the individual with genotype *AAbb* or *AABb* is also fertile and viable. While allele *a* and *b* both function properly within population, they are potentially incompatible with each other when combined together in a common genetic background, resulting in hybrid sterility or inviability. Thus, two populations become reproductively isolated. Muller (1942) pointed out that mutations do not necessarily occur in both populations. One population could retain ancestral AABB genotype whereas the other population could change to aaBB then aabb. The interaction between A and b may cause the breakdown of hybrid fitness in the hybrid genetic background AaBb. The complete isolation could be a consequence of the cumulative effect of many small incompatibilities between "complementary genes" scattered throughout the genome, which refer to genes showing deleterious interaction in a hybrid genetic background. Thus, this model provides a simple mechanism for reproductive isolation: the reproductive isolation is just a byproduct of normal substitution processes. By recent genetic work on postzygotic

isolation, the Dobzhansky-Muller model has been extensively corroborated (reviewed in Orr 1997).

Orr (1993a, 1995) formalized Dobzhansky-Muller model mathematically and obtained some important conclusions. Firstly, the time of divergence has a "snowball" effect on the number of loci causing postzygotic isolation. Secondly, the severity of postzygotic isolation might increase faster than linearly with time. The epistatic interactions increase at least as fast as the square of the time since the separation between two species, which is not hard to understand because the number of incompatibilities increases faster than linearly with the number of substitutions that have occurred. Thirdly, substitutions at a later stage of divergence will cause more incompatibilities than those that happen earlier. Fourth, all hybrid incompatibilities are asymmetric. For example, the derived allele *a* in population A could be incompatible with the derived allele *b* in population B, but ancestral alleles *A* and *B* must be compatible with each other since they have been tested by natural selection during intermediate ancestral steps. In other words, the derived and ancestral alleles of the first locus (a and A) can not both be incompatible with alleles of the second locus (b and B).

1.2 Genetic Search for Hybrid Sterility Genes

Despite the early emphasis on the study of speciation, empirical data regarding the nature of genetic changes underlying the evolution of reproductive isolation is surprisingly scarce. One of the most important discoveries in this field is Haldane's rule, which states the preferential sterility or inviability of heterogametic sex in species hybrids. The rule was first based on data from Lepidoptera, birds, flies, mammals,

Anoplura, and Cladocera (Haldane, 1922). Consecutive surveys of interspecific crosses revealed that Haldane's rule is well obeyed in all animals known to possess sex chromosomes (reviewed in Orr 1997). Heterogametic hybrid sterility has received special attention for its apparent involvement in the early establishment of reproductive isolation. In *Drosophila*, interspecific hybridization has revealed that hybrid male sterility is far more common than hybrid inviability (Bock 1984). In addition, in heterogametic-male taxa, hybrid male sterility involves spermatogenesis, which is a well-defined system (Lindsley and Tokuyasu 1980).

The first step to elucidate the genetic basis of hybrid sterility is to localize and characterize genes involved in this process. The genetic architecture underlying hybrid sterility is a heated debate in postzygotic isolation studies (Wu and Davis 1993). Whether there are a small number of discrete genes, each capable of causing complete hybrid sterility, or a large number of genes, jointly responsible for the complete hybrid sterility effect, has been intensively investigated in the past decades, yet it still remains unresolved. The basic approach to study hybrid sterility has moved from analyses of F2 backcross hybrids to conventional analyses of introgressions and more recently to high-resolution mapping of introgressions (Wu and Palopoli 1994). The analysis of F2 backcross hybrids was developed by Dobzhansky (1936) to study *D. pseudobscura* / *D. persimilis* hybrids. This method has provided a general assessment of the effect on hybrid sterility of interacting chromosomal regions between closely related species. However, its capacity to resolve the underlying genetic architecture is limited due to the complexity of the genetic compositions of F2 hybrids generated. Marked regions of the genome from

one species of *Drosophila* are moved into otherwise pure genetic background of another species by repeated backcrosses (Wu and Beckenbach 1983; Coyne and Charlesworth 1986). This allows for a cleaner analysis and makes it possible to attribute the hybrid sterility effect to a specific introgression. But a better understanding of the genetic architecture requires a more powerful method to decompose the introgression into smaller ones. A recent genetic analysis with higher resolution was thus developed, providing a direct study of genes involved in hybrid sterility by the involvement of molecular markers. A candidate "speciation gene", *Odysseus* in *Drosophila*, that lowers hybrid male fertility when placed in a foreign genetic background has been identified by this method (Perez et al. 1993) and further characterized (Ting et al. 1998).

1.2.1 F2 Backcross Hybrids

The study of F2 backcross hybrids is an important early step of hybrid sterility analysis and was the most widely used approach. F1 fertile hybrids can be backcrossed to either of the parental species to produce F2 hybrids with different combinations of chromosomes from two species. Morphological markers can be used to identify the origin of chromosomal segments and their effects on hybrid male sterility were examined. All individuals carrying a specific marker from one species will have at least surrounding region of the same species. Therefore, by checking the sterility of hybrids and the corresponding genotype, specific combinations of chromosomes causing male sterility in F2 can be revealed.

The most important conclusion drawn from F2 hybrid analysis is that most chromosome arms have some effect on hybrid male sterility even between closely related

species and the X chromosome has the greatest effect (Dobzhansky 1936; Charlesworth, Coyne, and Barton 1987). But the interacting components from two species can not be revealed in detail due to the substantial heterogeneity in the genetic composition in F2 hybrids. One phenotype might correspond to a collection of several genotypes sharing the same markers. Furthermore, many recombination genotypes remain uncharacterized.

Because of the intrinsic limitations of F2 hybrid analysis, the resolution of this method is far from enough to come up with an accurate explanation to account for hybrid male sterility unless a small number of loci underlying male sterility have been concluded. This method is effective only in pairs that are close to a primary speciation event, in which case, only a few loci may be responsible for hybrid male sterility. Two subspecies of *D. pseudoobscura*, Bogota and USA, have diverged as recently as 155,000 to 230,000 years ago with a Nei's genetic distance of 0.194 (Schaeffer and Miller 1991). Prezygotic isolation between them is very weak and postzygotic isolation is incomplete (Noor 1995), suggesting that they are at the early stage of speciation and their genetic differentiation may not be very extensive. Orr (1989a) detected a strong X chromosome effect on hybrid male sterility between this pair and less than 30% of the X chromosome was responsible, indicative of the involvement of small number of loci. In a recent study on the same species, Orr and Irving (2001) also suggested that hybrid male sterility involved a modest number of loci with a complex pattern of epistatic interactions. Coyne (1984) applied this method to another species pair with a Nei's genetic distance of 0.30 – *D. simulans* and *D. mauritiana*. The study showed that at least one locus on each of the five chromosome arms has a significant effect on male sterility, which represented the

maximum genetic divergence detectable with one recessive marker on each of the five major chromosome arms and indicated that many other hybrid sterility genes were present.

An alternative way to alleviate the problems of F2 hybrid analysis is to take advantage of one directional fertility of F1 males between recently diverged pairs. The resulting F2 hybrids would be less complex in genetic composition because there is no recombination in *Drosophila* males. The origin of chromosomes can be identified by means of markers and genotypes bearing the same set of markers are genetically homogeneous. Hybrid males are fertile from the cross between *D. arizonensis* males and *D. mojavensis* females. Vigneault and Zouros (1986) backcrossed these hybrid males to parental species and revealed that Y chromosome and two autosomes are involved in hybrid male sterility.

F2 hybrid analysis can reveal the presence of at least one region involved in hybrid sterility by using a particular marker. The number of genes causing hybrid male sterility has been estimated by this approach and most of the studies attribute hybrid sterility to genes with major effect (Coyne 1984; Orr 1989b; Orr 1992). However, the evidence obtained from this method was not conclusive. The results were compatible to either major gene or polygene interpretations since the major gene effect could also be explained as a consequence of a large number of clustered loci.

1.2.2 Conventional Introgression Analysis

A small segment of chromosome from one species can be introgressed into the pure genetic background of another species by repeated backcrosses so the male sterility

effect can be more precisely assigned to a specific introgression. In the introgression study by Wu and Beckenbach (1983), two regions that caused male sterility were mapped on the right arm of the X chromosome when introgressed from *D. persimilis* into *D. pseudoobscura*, whereas one sterility region was detected in the reciprocal direction. This is the maximum number of hybrid sterility genes detectable by three markers. Coyne and Charlesworth (1989) also revealed three loci linked to each of the three morphological markers on the *D. mauritiana* X chromosome, each of them capable of causing hybrid male sterility by itself in the *D. simulans* genetic background. Interestingly, Naveira (1992) studied the same chromosomal regions of the same species pair and suggested that at least two of these localized sterility effects could be further subdivided into smaller, linked ones.

1.2.3 Fine-Mapping Recombination Analysis

Although the introgression approach has been successful in detecting hybrid sterility regions, the resolution is still not sufficient enough to detect all possible sterility factors until DNA marker-assisted recombination is employed. This refined method has taken advantage of morphological and molecular markers to precisely define the boundary of introgressions and made it possible to ultimately delineate a sterility factor to a gene-sized interval by manipulating the introgression length by recombination.

A hybrid male sterility gene with major effect has been mapped by this method. Perez et al. (1993) formalized three criteria to infer the existence of major hybrid male sterility genes: (1) the presence of only two distinct phenotypic classes; (2) complementarity from both flanking markers by recombination analysis; and (3) physical

demarcation by molecular markers. They studied a small introgression of *D. mauritiana* in the *D. simulans* background and mapped a major sterility gene, *Odysseus* (*Ods*), at the cytological location of 16D. *Ods* of *D. sechellia* in *D. simulans* background caused no hybrid sterility. Unfortunately, further characterization of this gene revealed that *Ods* alone was not capable of conferring full hybrid sterility effect (Perez and Wu, 1995).

1.3 Genetic Basis of Haldane's Rule

Haldane's rule is obeyed in a wide variety of animals, suggesting shared genetic characteristic underlying the genetics of postzygotic isolation. Ever since the formalization of Haldane's rule, the search for the explanation for this general phenomenon has never stopped. Orr (1993b) suggested that Haldane's rule is most likely to have multiple genetic causes. Its distinct components such as hybrid sterility and hybrid inviability require distinct explanations (Wu and Davis 1993).

A number of theories have been proposed to explain the excess of hybrid sterility in the heterogametic sex. Among them, the dominance theory, the faster-male evolution theory and the faster-X theory are the leading explanations. Each of these mechanisms may explain some cases, but none of them accounts for the ubiquity of Haldane's rule.

1.3.1 The Dominance Theory

The dominance theory is an extension of the Dobzhansky-Muller model, proposed firstly by Muller (1942) then revised by Orr (1993a). To explain this theory, we consider two complementary genes that interact to cause hybrid sterility: Allele A1 from the first species is incompatible with B2 from the second species. If both complementary genes are autosomal, regardless of the dominance or recessivity, males and females have

the same fate as both sexes have the same genotype. In the case that one locus is X-linked and the other is autosomal, it may not be surprising to observe that more heterogametic sex are afflicted than homogametic sex: the former will suffer full effect of the complementary genes, dominance and recessive, while it is partially masked in the latter. Thus, Haldane's rule results as long as some fraction of the alleles lowering hybrid fitness is recessive. From the observation of the large time lag between the evolution of hybrid male and female sterility in *Drosophila*, Turelli and Orr (1995) suggested that the alleles lowering hybrid fitness must be very recessive.

While the dominance theory offers a simple and compelling explanation for most of the hybrid sterility cases, it was challenged in some tests. Coyne (1985) predicted that if hybrid sterility is caused by recessive X-linked genes, then hybrid females carrying homozygous Xs should also be sterile because these females suffer the same effect of the interaction of incompatibilities as hybrid males. He tested this prediction by producing "unbalanced" hybrid females that carry an attached-X of one species on an otherwise hybrid genetic background. The result has contrasted the prediction, showing that these females remained perfectly fit despite being homozygous for their X-chromosomes. Similar results were obtained in several other independent hybridizations (Orr 1989a; Orr and Coyne 1989). A test made in Lepidoptera also refuted a simple prediction of the dominance theory that the hybrid sterility in species having large X-chromosomes should have evolved faster than species with small X-chromosomes since large X chromosomes are expected to harbor more recessive X-linked incompatibilities than small X chromosome (Presgraves 2002). Moreover, the dominance theory encountered difficulties

in the species lacking a degenerate Y chromosome. Mosquitoes of the genus *Aedes* have two functional X-chromosomes in both sexes. Cross experiment showed that only male *Aedes* suffer hybrid sterility (Presgraves and Orr 1998). The dominance theory has relied on the hemizygous X state of the heterogametic sex and hence fails to provide an explanation in *Aedes*. All these findings seem to falsify the dominance theory and have stimulated us to search for a more convincing theory for the hybrid sterility component of Haldane's rule.

1.3.2 Faster-Male Evolution Theory

The second theory, faster-male evolution, posits that pervasive hybrid sterility is due to the faster accumulation of male sterility genes than female sterility genes (Wu and Davis 1993). The evidence for the rapid evolution of male reproductive traits has been extensively provided (reviewed in Singh and Kulathinal 2000; Swanson and Vacquier 2002). Under the Dobzhansky-Muller model, rapidly diverging genes would be reasonable to develop more incompatibilities than slowly evolving genes. Two factors, unique developmental properties of spermatogenesis in heterogametic males and stronger sexual selection for male reproductive characters, were proposed to possibly underlie such faster-male evolution.

The faster-male theory suffers the obvious weakness that it can not explain the excess hybrid female sterility in taxa with heterogametic-female. Both possible reasons for the appearance of hybrid male sterility work against the preference of hybrid female sterility in heterogametic-female species. Other forces must have been involved to overcome the faster-male effect to give rise to Haldane's rule in these taxa.

1.3.3 Faster Evolution of X-linked Genes

Numerous genetic studies on hybrid sterility have revealed a second general pattern in postzygotic isolation, so-called "large X effect". Experimental studies have showed that X chromosome had the disproportionately large effect on hybrid male sterility (Dobzhansky 1936; Wu and Beckenback 1983; Coyne and Kreitman 1986; Orr 1987; Orr 1989a; True et al. 1996). This large X chromosome effect has been suggested to be due to the hemizygous bias, since the studies usually compared the effect of replacing hemizygous X chromosome locus with the effect of replacing heterozygous autosome locus. True, Weir, and Laurie (1996) tested this bias by comparing hemizygous replacements of the X with homozygous replacements of the autosomes. They separately placed 87 marked positions with P elements throughout the *D. mauritiana* genome in an otherwise homozygous *D. simulans* background and revealed that the density of hybrid male sterility factors is much greater on the hemizygous X chromosome than homozygous autosomes.

Charlesworth, Coyne, and Barton (1987) attributed this phenomenon to the recessive or partially recessive nature of favorable mutations on the average. Substitutions of partially recessive mutations will have a larger effect when on the X chromosome than on autosomes. A general prediction that X-linked genes evolve faster thus follows. Although homogametic hybrids suffer from twice as many X-linked incompatibilities as heterogametic ones, the excess of hybrid sterility in the heterogametic sex results due to the full effect of recessive X-linked genes. However, the assumption that favorable mutations are typically partially recessive encounters the

problem for the lack of large X-effects on morphological and behavioral differences: large X-effects should display not only on postzygotic isolation, but also on morphological and behavioral (species) differences if this assumption is correct. Furthermore, this assumption is biochemically contrary to the metabolic theory and genetic data that show the loss-of-function nature of recessive alleles (Orr 1997).

1.4 The Species

The *D. melanogaster* complex includes *D. melanogaster* and the *D. simulans* clade that is comprised of *D. simulans*, *D. mauritiana* and *D. sechellia*. In distribution, *D. melanogaster* and *D. simulans* are cosmopolitan species whereas *D. mauritiana* and *D. sechellia* are insular endemic, restricted in the Mauritius and Seychelles, respectively (Lachaise et al. 1988). *D. melanogaster* is the closest relative to the *D. simulans* clade with the split time of 2.5 – 3.4 mya. Within the *D. simulans* clade, the separations from each other occurred 0.58 – 0.86 mya ago (Hey and Kliman 1993). These four sibling species are morphologically indistinguishable and have identical karyotype. They are homosequential in chromosomes except for a very small inversion on the third chromosome. The DNA divergence is small in the *D. melanogaster* complex (Lemeunier and Ashburner 1976). As a model species for genetics for almost a century, the genome of *D. melanogaster* was completely sequenced (Adams et al. 2000) and is a valuable resource of sequence information for the *D. simulans* clade because of their close relatedness. Interspecific hybrids within the *D. simulans* clade produce fertile females and sterile males (Lachaise et al. 1988), following Haldane's rule, which makes the backcross study possible. Moreover, there are various phenotypic mutations in *D.*

simulans that can facilitate the genetic analysis of the hybrids. All these advantages have made this complex a subject of extensive genetic studies.

1.5 Objectives of the Study

After intensive study for half of a century, a convincing consensus for the explanation of Haldane's rule has not been reached yet. A growing number of studies have revealed a higher divergence of individual sex-related genes than non-sex genes (reviewed in Singh and Kulathinal 2000). In the first part of the thesis, we studied the molecular evolution of sex and reproduction related genes in a comprehensive manner and tested the hypothesis that sex-related genes, as a general class, evolve faster than non-sex genes on the X chromosome of the *D. melanogaster* complex. The fast-sex evolution theory, which explains the hybrid sterility component of Haldane's rule including all the incongruities of the dominance theory, was supported (Singh 2000).

The evolution of reproductive barriers between conspecific populations is the fundamental step to speciation. A complete explanation of the genetics of speciation requires the understanding of the type, number and effect of genes that prevent free exchange of genetic matter. As a first step, a precise mapping of hybrid sterility genes would be necessary. The second part of this study addressed the question of the genetic basis of postmating reproductive isolation, through the investigation of hybrid male sterility involving a specific X chromosome region that has been shown to have a large hybrid sterility effect between *D. mauritiana* and *D. simulans*.

CHAPTER 2

Faster Evolution of Sex and Reproduction Related Genes on the X chromosome of the *Drosophila melanogaster* Complex

ABSTRACT

Many individual genes that are related to sex and reproduction have been shown to evolve at a high rate in a variety of taxa. A sample of 19 sex genes and 20 non-sex genes, in a comprehensive manner, on the X chromosome of *Drosophila melanogaster, D. mauritiana, D. simulans, and D. sechellia* were amplified by PCR and characterized by restriction site analysis (PCR-RFLP). These genes come from an important region that is known to contain hybrid sterility genes. The divergences of these two sets of genes for every pair of these four sibling species were compared. The results revealed a significantly higher divergence for sex genes compared to non-sex genes for every pairwise comparison. Among the sex gene group, genes involved in male reproductive traits appeared to evolve faster than genes for female reproductive traits. The intraspecific variations in *D. simulans* and *D. mauritiana* were also obtained, revealing a comparable polymorphism level between sex and non-sex genes. Tajima's and HKA tests were applied to test the neutral theory of molecular evolution. Neither the evolution of sex genes nor of non-sex genes revealed significant deviation from neutrality. The elevated divergence of sex genes, in contrast to the dominance theory's prediction of the rapid evolution of X-linked genes in general, provides support for the fast-sex theory as an explanation for the hybrid sterility component of Haldane's rule. The phylogeny among the *D. melanogaster* complex were obtained from site presences/absences of both sex and non-sex genes.

Many individual sex and reproduction-related genes (SRR genes, referred to as sex genes hereafter) have exhibited a high divergence between species. The early evidence comes from morphological studies which show that male genitalia differentiate faster so much so that they often act as the only reliable species-diagnostic trait between related species such as those of the *D. melanogaster* complex (Eberhard 1985). More recently, Makalowski and Boguski (1998) compared 2,820 proteins encoded by orthologous genes from humans and rodents and revealed that proteins involved in reproduction are among the most rapidly evolving groups. In line with this view, faster evolution of reproductive tissues in comparison to non-reproductive tissues in *Drosophila* has been revealed in phenotypic (Civetta and Singh 1998a), protein (Coulthart and Singh 1988) and molecular (Ting et al. 1998) traits. More evidence comes from a fertilization protein in marine invertebrate (Lee et al. 1995), sex determination and mate recognition genes in eukaryotic phyla (Ferris et al. 1997), and sperm genes in humans and primates (Wyckoff, Wang, and Wu 2000). Such highly diverged genes are important in evolution because they may have played an essential role in the formation of new species. Nei and Zhang (1998) have suggested that incompatibilities of alleles related to mating and spermatogenesis appear to be responsible for reproductive isolation.

Recent studies have extended the scope of faster evolving characters from mainly male reproductive traits to a wider range (reviewed in Civetta and Singh 1999). Although the evidence is not overwhelming, female reproductive traits have been demonstrated to evolve fast as well (Civetta and Singh 1995; Swanson et al. 2001b). However, how sex

genes evolve, as a general group, remains untested. In this study, we address the question whether sex genes, in a general sense, have a higher rate of nucleotide divergence between closely related species in *Drosophila* than non-sex genes. We focused here on the genes on the X chromosome because the X chromosome has shown a large effect in postzygotic reproductive isolation (Charlesworth, Coyne, and Barton 1987), and from the dominance theory, generally a faster evolution of all X-linked genes is expected. We performed sequence analysis using Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP) data from a random sample of 19 sex and 20 non-sex genes on the X chromosome of the *D. melanogaster* complex. This region is known to contain major genes affecting hybrid sterility (Perez et al. 1993; Zeng and Singh 1995). Tajima's test and HKA test were performed on the data to test the neutral theory of molecular evolution. Together with other studies (Nurminsky et al. 1998; Ting et al. 1998), these data provide a broad, chromosome-wide assessment of the rate of divergence of sex gene group and show how these genes differ in their rate of evolution from non-sex genes. The hypothesis that the fast-sex evolution accounts for the hybrid sterility component of Haldane's rule was tested.

MATERIALS AND METHODS

2.1 *Drosophila* Strains

Four species of the *D. melanogaster* complex were used for the between-species divergence study. Stock center designations are indicated in parentheses. *D. melanogaster* (0231.0) from Hawaii, *D. simulans* (0251.2) from Colombia, *D. sechellia* (0248.3) from Cousin Island, which were obtained from the Drosophila Tucson Stock Center at Arizona University, and *D. mauritiana* (S080) from the Drosophila Stock Center at Umea, Sweden. The population genetic study has revealed a low level of variation for *D. melanogaster* as well as for *D. sechellia* and a relatively high polymorphism level in *D. simulans* and *D. mauritiana* (Kliman and Hey 1993; Kliman et al. 2000). Five additional lines of *D. simulans:* 0251.166 from Florida, 1088 with unknown origin (obtained from the Drosophila Species Stock Center at Bowling Green), S132 from Italy (obtained from the Drosophila Stock Centre at Umea, Sweden), S24 from Madagascar (provided by John Roote in Cambridge University, UK), and three additional lines of *D. mauritiana* (0241.1, 0241.5 and 0251.7 with unknown origin) from the Tucson Stock Center were used to determine the within species polymorphism.

2.2 Selection of Genes

Sex and non-sex genes were chosen based on their known/predicted functions from Flybase (http://flybase.bio.indiana.edu). The region 7B to 19F on the X chromosome was examined in order to develop molecular markers to search for hybrid

male sterility genes. This area represents a region with a high concentration of genes affecting hybrid male sterility (True, Weir, and Laurie 1996). The recombination rates of these genes in this region range from low to intermediate and high. Genes affecting sex determination or reproductive traits in a broad sense were classified as sex genes. Generally speaking, first sex genes were chosen and then randomly non-sex genes were chosen from the nearby sub-regions. In total, 19 sex genes and 20 non-sex genes were examined (Appendix A).

2.3 PCR-RFLP Analysis

Primers for each gene were designed based on the coding sequence of *D. melanogaster* retrieved from GenBank. Most of the target segments defined by PCR are coding regions. Some contained introns, in which case the restriction sites within them were not counted. PCR amplifications were performed under standard conditions. In the total of 25 μl reaction mix, there was 50 mM KCl, 10 mM Tris-HCl (pH 9.0), 1.5 mM $MgCl_2$, 0.01% gelatin (w/v), 0.1% Triton X-100, 0.2 mM for each of the dNTPs, 40 pmol each primer, about 20 μg of genomic DNA, 1.2 units of Taq DNA polymerase (Promega) and ddH_2O. The mix was amplified by thirty-five reaction cycles with 45 sec at 94 °C, 45 sec at annealing temperature and 1 min at 72 °C then finally left for 10 min at 72 °C. The annealing temperatures varied from gene to gene.

PCR products were then digested by several restriction enzymes to detect restriction site differences between pairs of species. The digestion reaction contained 4-5 μl of PCR product, 2 U of restriction enzyme, 2 μl of 10 × reaction buffer, and ddH_2O to a total volumn of 20 μl. The reaction was incubated at 37°C overnight. Restriction digests

were detected in 1.2 % agarose electrophoresis and viewed by UV transillumination. 1Kb ladder was used as a standard to determine the restriction fragment sizes. Enzymes used included AccI (GT/MKAC), AluI (AG/CT), BclI (T/GATCA), BglII (A/GATCT), CfoI (GCGC), ClaI (AT/CGAT), DdeI (C/TNAG), EcoRV (GAT/ATC), HaeII (RGCGC/Y), HincII (GTY/RAC), HinfI (G/ANTC), KpnI (GGTAC/C), MboI (/GATC), MseI (T/TAA), MspI (C/CGG), PstI (CTGCA/G), PvuII (CAG/CTG), RsaI (GT/AC), Sau96I (G/GNCC), TaqI (T/CGA), XbaI (T/CTAGA) and XhoI (C/TCGAG).

2.4 Data Analysis

2.4.1 Divergence Analysis

Several enzymes with different base recognition sites were used for each gene. The probability of a nucleotide difference at a given nucleotide position between pair of species (*P*) was estimated from Nei's (1987, Pp104) maximum likelihood method in equation (1). The number of sites present in each species and the number of sites shared for each pair produced by groups of enzymes with the same number of bases in their recognition sequence were counted.

$$P = \hat{P}_I \frac{\Sigma_i r_i (m_i - m_{xyi}) / \{[1 - (1 - \hat{P}_I)^{ri}][2 - (1 - P_I)^{ri}]\}}{\Sigma_i r_i [m_i / (2 - (1 - \hat{P}_I)^{ri})]} \qquad (1)$$

$$V(\hat{d}) = 1 / \sum(1 / V(\hat{d}_i)) \qquad (2)$$

$$V(\hat{d}_i) = \frac{9 (1 - P)^2 (2 - S_i) (1 - S_i)}{2 r_i^2 m_i (3 - 4P)^2 S_i} \qquad (3)$$

Where r_i is the number of nucleotides of a recognition sequence of a restriction enzyme i; m_{xi} and m_{yi} are the numbers of restriction sites for DNA sequences X and Y, respectively; m_{xyi} is the number of restriction sites shared by the two sequences and $m_i = (m_{xi} + m_{yi}) / 2$; $\hat{P}_1$ is a trial value of P. The first trial value of P is equal to $1 - S_{il}^{1/ri}$ where $S_{il} = m_{xyi} / m_i$. When $P = \hat{P}_1$, P is the maximum likelihood estimateor. Divergence (the number of nucleotide substitutions per site) can be obtained by $d = -3/4 \ln (1 - (4/3)P)$. The variance of divergence was obtained from equation (2) where $V(\hat{d}_i)$ was obtained from equation (3) for the ith type of restriction enzymes when $S_i = (1 - P)^{ri}$. The test statistics, t-tests, were used to measure any deviation of divergences between sex and non-sex genes. Data of site presences/absences were used in the PHYLIY software package (Felsenstein 1995) to obtain phylogenetic relationship among the four species of the *D. melanogaster* complex.

2.4.2 Within-species Polymorphism in *D. simulans* and *D. mauritiana*

Two measures of nucleotide diversity have been employed. The first was to estimate the number of nucleotide differences, π, following the method of Nei and Li (1979). The second measure was to estimate θ from the number of polymorphic sites following Hudson (1982). The parameter θ is equal to $4N\mu$ for automosomal genes and $3N\mu$ for X-linked genes, where N is the population size and μ is the neutral mutation rate, because the effective population size for X-linked genes is smaller than autosomal genes. The θ was calculated by dividing the proportion of polymorphic sites (p) by $\sum 1/i$ (i = 1 to n-1), where n is the number of sequences sampled (Begun and Aquadro 1991). The length of the recognition sequence of restriction enzymes, j, is the weighted average of

the lengths of the recognition sequences $\sum im_i / \sum m_i$, where m_i is the number of cleavage sites in the sample that correspond to recognition sequences of length i. The following equations show the calculations for p and the variances of θ (Hudson 1982).

$$p = k / (2m-k)j \quad (4)$$

$$Var(\theta) = \theta^2 / k \quad (5)$$

Where m be the number of the cleavage sites found, and k be the number of the cleavage sites which are polymorphic.

2.4.3 Test of Neutrality

Two tests of the neutrality of molecular evolution have been used. Assuming all the nucleotide variants neutral and the populations panmixia and equilibrium, $\pi = \theta = 3N\mu$ should be observed. Tajima's test (1989) statistically tests whether π and θ are significantly different. Only intraspecific variations are required for this test. The restriction site data were also used for HKA test (Hudson, Kreitman and Aguade 1987), using the HKA program at http://lifesci.rutgers.edu/~heylab.

RESULTS

Target gene segments with expected lengths were obtained by PCR from all species, and then digested by restriction enzymes. For each gene, the restriction sites for individual species as well as the sites shared by two species were counted according to the digestion patterns (Appendix B and C). Since gel electrophoresis can not clearly show the restriction fragments smaller than about 100bp and can not distinguish doublets with similar lengths, the determination of the restriction sites was assisted by the predicted patterns based on the *D. melanogaster* sequence. Although the resolution of the restriction mapping is relatively coarse, Nei and Tajima (1981) have suggested that the lack of accuracy of small fragment detection does not significantly affect this kind of analysis. The maps of restriction sites indicated that changes in restriction sites between pairs of species could be reasonably attributed to the gain or loss of restriction sites resulting from nucleotide substitutions. The GC contents in the nucleotide sequences of these two sets of genes, as well as those in recognition sites of restriction enzymes used in sex and non-sex genes, are similar and therefore the bias of nucleotide base composition variation is minimized (Galtier and Gouy, 1995).

2.5 Between-Species Divergence

The nucleotide divergence between each pair of species was estimated by the matrix of restriction sites from these two sets of genes. The average divergence of sex and non-sex genes is about 2% between *D. melanogaster* and the *D. simulans* clade, and

0.8% within the *D. simulans* clade, which is consistent with previous studies (Table 3 in Hilton, Kilman, and Hey 1994). Sex genes show a significantly higher rate of divergence than non-sex genes for every pairwise comparison (Table 1.1 and Fig 1.1). More sex genes show variance between species compared to non-sex genes. Moreover, divergences from individual genes reveal a higher evolution rate for male-reproductive related genes than female-reproductive related ones between *D. melanogaster* and the *D. simulans* clade. Male *vs.* female divergence is 0.045 *vs.* 0.024, 0.036 *vs.* 0.021, and 0.034 *vs.* 0.020 between *D. melanogaster* and *D. mauritiana*, *D. melanogaster* and *D. simulans*, and *D. melanogaster* and *D. sechellia*, respectively. The first two fastest evolving sex genes are both male-reproductive related (Fig 2a, b, and c). Within the *D. simulans* clade, a higher divergence is also revealed for male reproductive genes (0.0089) compared to female reproductive genes (0.0055) between *D. simulans* and *D. sechellia*. Although the other two pairs show slightly higher divergence for female reproductive genes than male reproductive genes (male *vs.* female is 0.0090 *vs.* 0.0136 and 0.0088 *vs.* 0.0092 for *D. mauritiana* – *D. sechellia* and *D. mauritiana* – *D. simulans*, respectively), the fastest evolving sex gene (*otu*) belongs to male-reproductive class.

2.6 Within-Species Variation

The generally low level of variation present throughout the genome of *D. melanogaster* and *D. sechellia* makes it difficult to perform any statistic test for intraspecific variation. Therefore, polymorphism data were only obtained from *D. simulans* and *D. mauritiana,* which have been shown to be relatively highly polymorphic. In *D. simulans*, there were four polymorphic restriction-sites out of 209 in sex genes as

Table 1.1. Divergence of the *D. melanogaster* complex inferred from sex and non-sex genes

species pairs[a]	*P*		d ± S.D.		t-test
	sex genes	non-sex genes	sex genes	non-sex genes	P value
mel-mau	0.03035	0.01171	0.03098 ± 0.0044	0.01180 ± 0.0027	<0.0001
mel-sim	0.02553	0.01229	0.02597 ± 0.0040	0.01240 ± 0.0028	<0.0001
mel-sec	0.02508	0.01156	0.02551 ± 0.0039	0.01165 ± 0.0027	<0.0001
mau-sec	0.01336	0.007	0.01348 ± 0.0027	0.00704 ± 0.0021	<0.0001
mau-sim	0.00964	0.00411	0.00971 ± 0.0023	0.00413 ± 0.0016	<0.0001
sim-sec	0.00978	0.00641	0.00984 ± 0.0023	0.00644 ± 0.0020	<0.0001

[a]: mel - *D. melanogaster*; mau - *D. mauritiana*; sim - *D. simulans*; sec - *D. sechellia*

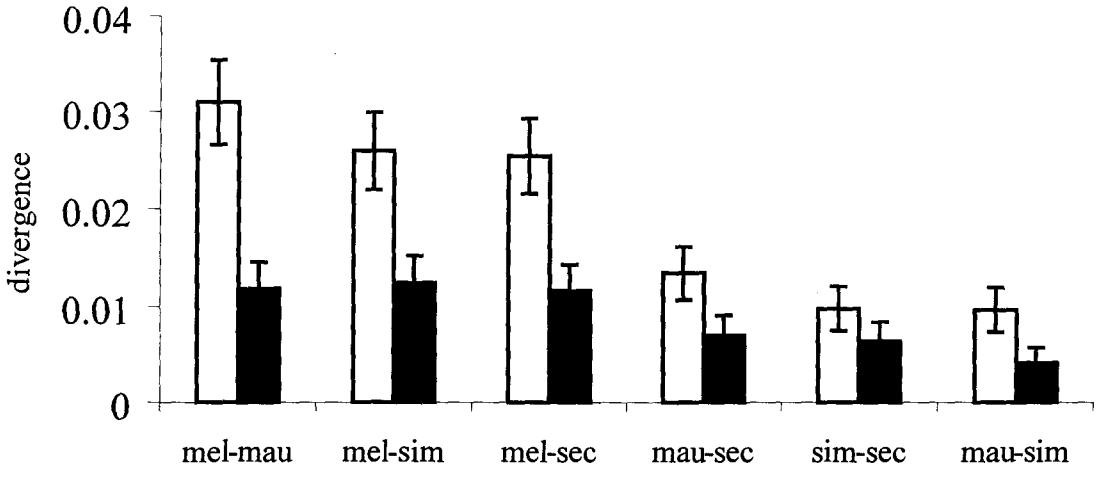


Fig. 1.1 Divergences of sex and non-sex genes between pairs of the *D. melanogaster* Complex. Open bars represent sex genes. Filled bars represent non-sex genes.

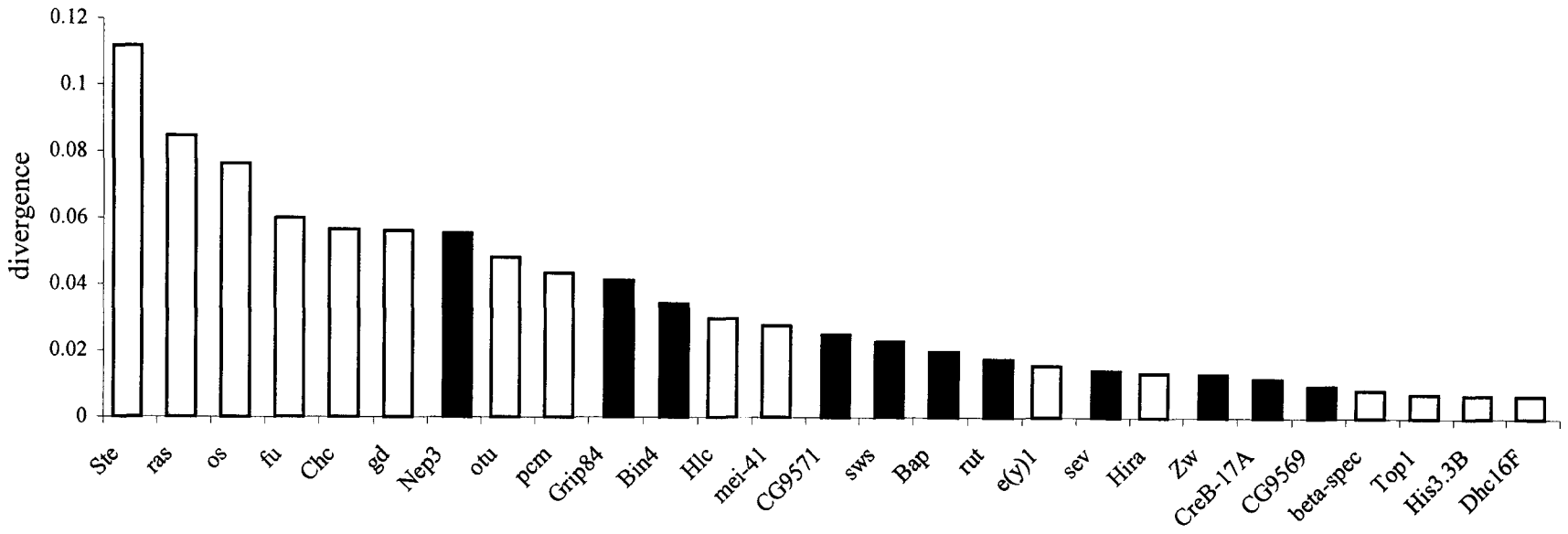


Fig. 1.2a. Divergences of individual genes between *D. melanogaster* and *D. mauritiana.* Open bars represent sex genes. Filled bars represent non-sex genes. Genes whose divergences are zero are not shown.

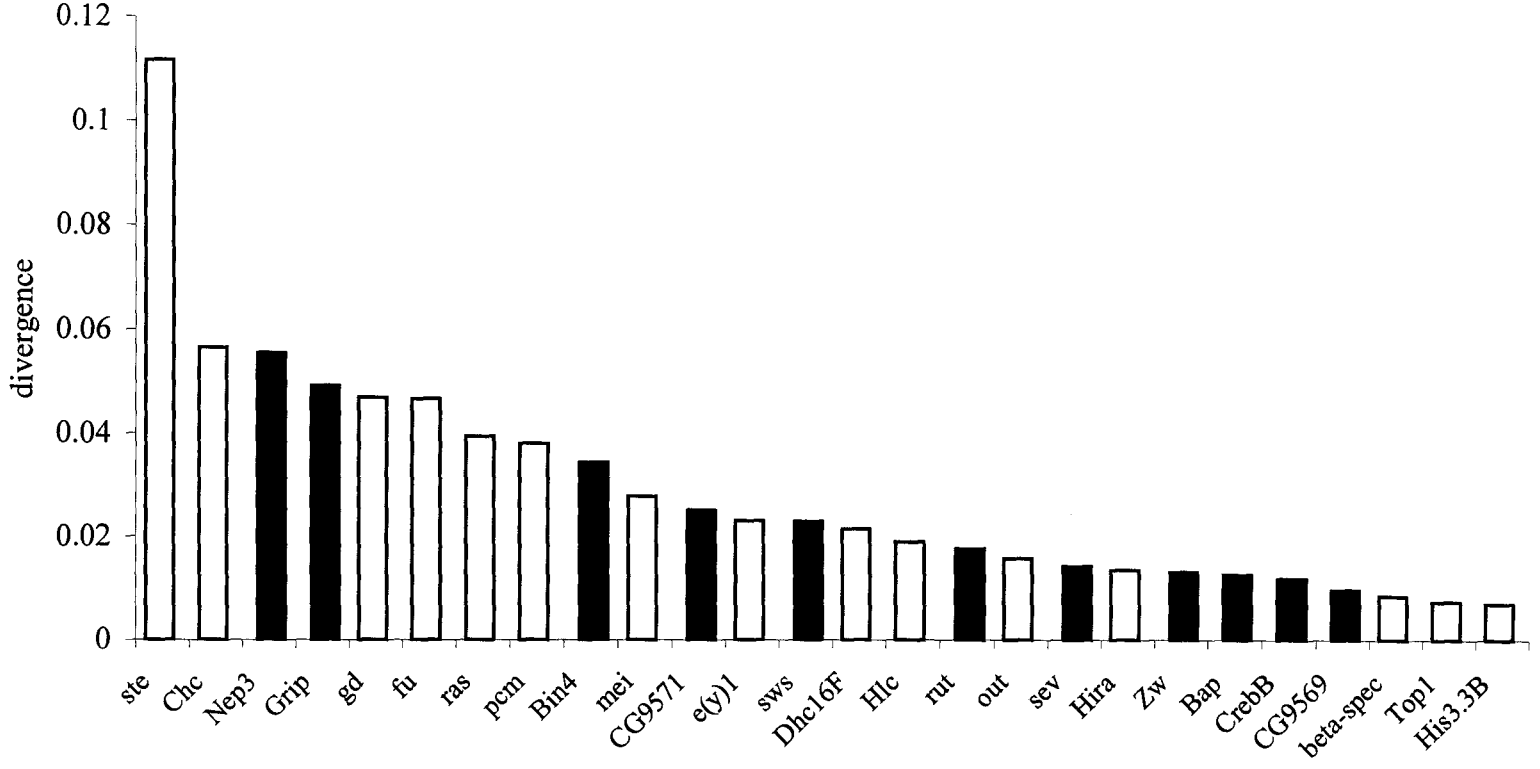


Fig. 1.2b. Divergences of individual genes between D. melanogaster and D. simulans. Open bars represent sex genes. Filled bars represent non-sex genes. Genes whose divergences are zero are not shown.

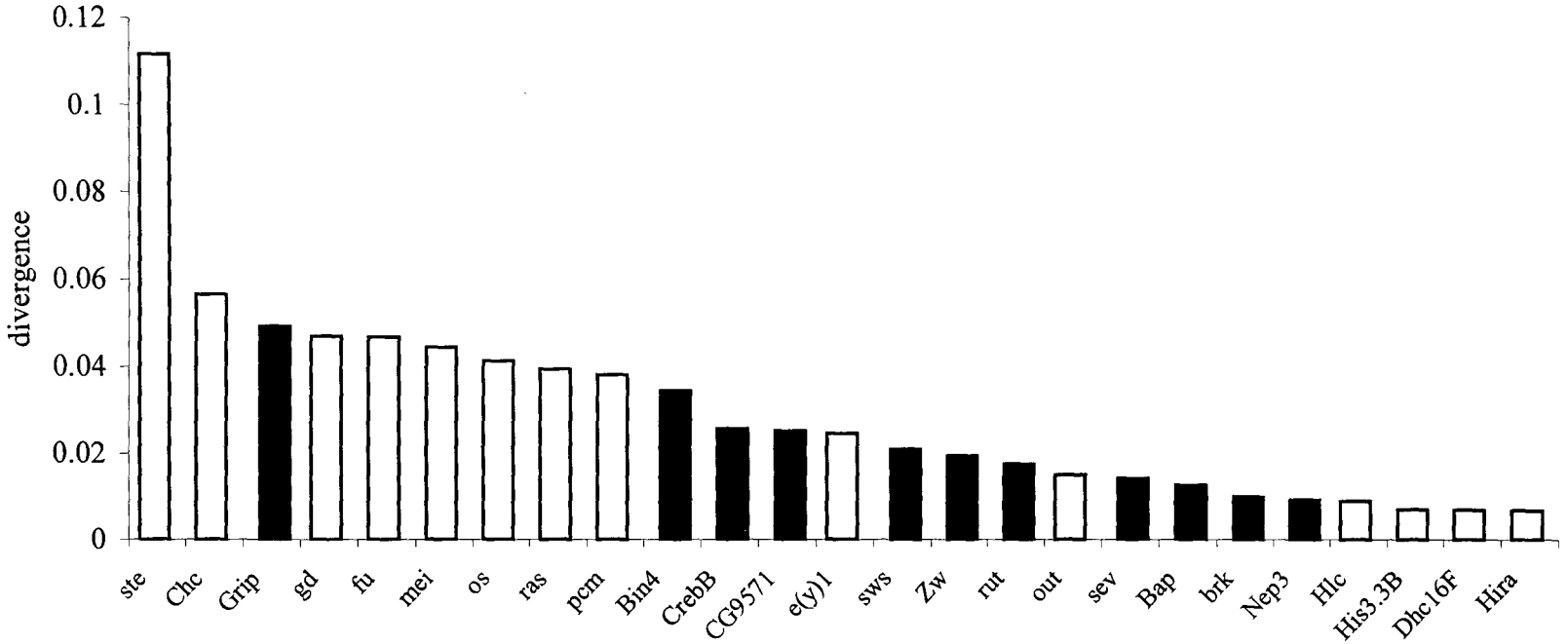


Fig. 1.2c. Divergences of individual genes between *D. melanogaster* and *D. sechellia*. Open bars represent sex genes. Filled bars represent non-sex genes. Genes whose divergences are zero are not shown.

opposed to five out of 189 in non-sex genes, and one polymorphic site in *D. mauritiana* was separately detected in sex genes and non-sex genes. The parameter θ, π, and their standard errors for *D. mauritiana* and *D. simulans* were given (Table 1.2). The intraspecific polymorphisms from coding regions revealed in this study are comparable to estimates from other studies (Begun and Aquadro 1991; Hey and Kliman 1993). The comparable within-species nucleotide variations for sex and non-sex genes suggest that the significantly higher divergence between species for sex genes was not a consequence of increased mutation rate.

2.7 Tests for Neutral Theory of Molecular Evolution

Tajima's test (1989) revealed that neither data for sex genes nor non-sex genes in *D. mauritiana* and *D. simulans* rejected the neutral theory of molecular evolution (Table 1.2). The positive D values reflect high frequency polymorphisms. Most polymorphisms occur more than once in the sample. The second test, HKA test, is based on the prediction that levels of intraspecific variation in regions of the genome are positively correlated with levels of interspecific divergence in the corresponding regions under a neutral, infinite-site model (Kimura 1983). Table 1.3 shows the data used to perform HKA test. In addition to sex and non-sex gene group, one *D. simulans* X-linked gene (*per*) and one autosomal gene (*rosy*), which were obtained from other studies (Aguadro, Lado, and Noon 1988; Begun and Aquadro 1991), were included. Table 1.4 shows the resulting test statistics in *D. simulans*. Neither sex genes nor non-sex genes revealed a significant deviation from neutrality when compared to X-linked genes. However, sex genes show a

Table 1.2. Nucleotide diversity of sex and non-sex genes and Tajima's test of neutrality in *D. simulans* and *D. mauritiana*

Region	Species	π	θ	D
Sex genes	*D. simulans*	0.0011± 0.0008	0.0009 ± 0.0005	0.039
	D. mauritiana	0.0004 ± 0.0003	0.0003 ± 0.0003	0.034
Non-sex genes	*D. simulans*	0.0023 ± 0.0018	0.0013 ± 0.0006	0.161
	D. mauritiana	0.0005 ± 0.0004	0.0003 ± 0.0003	0.067

Table 1.3. Nucleotide data used in the HKA test of neutrality in *D. mauritiana* and *D. simulans*

		Within species			Between species		
Genes	Species	m	k	Effective number of nucleotides surveyed	D	Effective number of nucleotides surveyed	N
Sex genes	D. mel	209		1973			
	D. mau	211	1	1996	56	1984.5	4
	D. sim	209	4	1962	46	1967.5	6
Non-sex genes	D. mel	197		1797			
	D. mau	188	1	1718	21	1757.5	4
	D. sim	189	5	1705	21	1751	6
per[a]	D. mel	39	4	444	21.2	513	35
	D. sim	54	1	582			36
			1				
rosy[b]	D. mel	41	7	450	19.2	477	60
	D. sim	56	2	504			30
			8				

Effective number of nucleotides surveyed by restriction map studies within species follows Hudson (1982) as (2m-k)j. The length of the recognition sequence of restriction enzymes, *j*, is the weighted average of the lengths of the recognition sequences $\sum im_i / \sum m_i$, where m_i is the number of cleavage sites in the sample that correspond to recognition sequences of length *i*. Effective number of nucleotide sites surveyed between species is the mean of the effective number of nucleotides in *D. mauritiana* and *D. simulans*. D is the average pairwise difference of sites between species. N is the number of sequences sampled in each species.

[a] Data from Begun and Aquadro (1991).

[b] Data from Aquadro, Lado and Noon (1988).

Table 1.4. Statistics from HKA test of neutrality in *D. simulans*

gene	Sex genes		Non-sex genes	
	X^2	p	X^2	p
Non-sex genes	1.16	0.28		
per	1.53	0.22	0.001	0.97
rosy	4.49	0.03	1.05	0.31

higher degree of departure from neutrality than non-sex genes. A significant departure was revealed for sex genes versus autosomal *rosy* region. In *D. mauritiana,* HKA test was only available between sex and non-sex genes and the X^2 is 0.41 ($P = 0.52$).

2.8 The Phylogenetic Analysis

The phylogeny of the *D. melanogaster* complex is of general interest. The phylogenetic relationship of these four species has exhibited complicated patterns (Palopoli, Davis, and Wu 1996). The phenogram based on sex gene data is consistent with the phenogram based on non-sex gene data. Fig. 1.3 shows the consensus tree inferred from both sex and non-sex gene data. *D. melanogaster* has consistently been placed as an outgroup of the *simulans* clade. Among the *simulans* clade, *D. mauritiana* and *D. simulans* were clustered first, which is consistent with the phylogeny inferred by the hybrid male sterility gene, *Odysseus* (Ting, Tsaur, and Wu 2000).

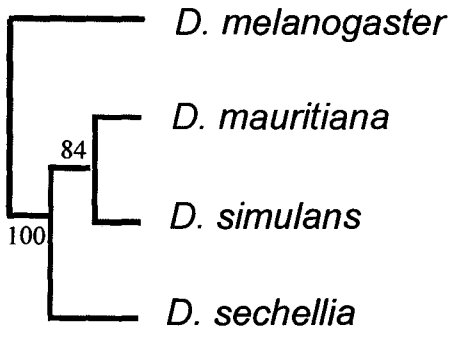


Fig. 1.3. The Phylogeny among the *D. melanogaster* complex, as inferred by the maximum parsimony method. Bootstrapping values are shown.

DISCUSSION

In this study, the restriction-site divergence between sex and non-sex genes on the X chromosome of the *D. melanogaster* complex was examined. The restriction-site polymorphism data for these two classes of genes in *D. mauritiana* and *D. simulans* were also obtained. Sex genes show significantly higher divergence in every pair of species than non-sex genes. In contrast, within-species nucleotide variations for sex and non-sex genes revealed no obvious distinction in *D. mauritiana* or *D. simulans*, suggesting that the disparity in divergence is not a reflection of increased mutation rate in sex genes than non-sex genes.

2.9 The Pattern of Restriction Site Polymorphism and Divergence

The estimates of θ and π for *D. simulans* and *D. mauritiana* are consistent with the previous X-linked studies and revealed relatively similar levels of polymorphism between the two sets of genes in both species. In the restriction site polymorphism, *D. simulans* appears to be more variable than *D. mauritiana*, suggesting that *D. simulans* has a larger effective population size than *D. mauritiana*, which is consistent with the view obtained from other loci (Hey and Kliman 1993).

Tajima's test showed the evolution of sex genes and non-sex genes in *D. simulans* and *D. mauritiana* to be compatible with neutral theory. HKA test reveals no significant departure from neutrality model for all X-linked genes as well. Only sex genes deviate significantly from neutral expectations when compared to autosomal *rosy* gene, which is

possibly partially due to the reduced X-linked nucleotide polymorphism (Begun and Whitley 2000). Since sex genes in this study are a combination of several individual genes, the deviation limited to only a fraction of the genes might not stand out in the result of the whole group. However, the analyses presented in this study emphasize the interspecific divergence and the sample size for intraspecific polymorphism data is small. The failure to reveal a significant departure from the neutral model could be due to a lack of statistical power resulting from the small number of sample size for polymorphism, given the significantly higher divergence for sex genes than non-sex genes. Moriyama and Powell (1996) suggested that Tajima's test may be sensitive to sample size. The test is strong in rejecting neutrality for the large data set. It is obvious that the degree of departure from neutrality expectation for sex genes compared to other genes is higher than non-sex genes *vs.* other gene comparison. It is therefore conceivable to suggest that sex genes have evolved under a different pattern of constraints compared to non-sex genes.

Whether the significantly higher divergence of sex genes compared to non-sex genes is a reflection of the relaxation of selective constrains or positive selection is not clear from our data. Additional sequencing data are required to gain a better understanding of the selective forces. However, the functional role of these genes leads us to tend to the second explanation that these genes are influenced by positive selection. Evidence has accumulated for the presence of positive sexual selection acting on sex genes, as demonstrated in *Acp* gene in *Drosophila* (Tsaur, Ting, and Wu 1998), reproductive genes in mammals (Wyckoff, Wang, and Wu 2000), and the fertilization

gene in marine invertebrate (Kresge, Vacquier, and Stout 2001). Sexual selection can drive the evolution of male reproductive genes through the form of sperm or pollen competition for fertilizations between closely related species (Rieseberg, Desrochers, and Youn 1995; Karr and Pitnick 1999; Van Doorn, Luttikhuizen, and Weissing 2001). Female reproductive traits may be driven through male-female coevolution, as suggested in mammalian egg proteins (Swanson et al. 2001b) and *Drosophila* sperm-storage organ (Miller and Pitnick 2002). These observations indicate that positive sexual selection might have played a critical role in the evolution of sex genes.

2.10 Different Divergence Patterns of Male and Female Genes

Here we broadly defined sex genes as those that are related to reproductive functions. In our data set, one gene is involved in primary sex determination and seven genes are related to spermatogenesis or male tissues, which are referred as male-reproductive genes. Unfortunately, not all of male-reproductive genes are perfectly male limited - some of them show the expression in the reproduction of both sexes. We did not intend to, but there are more genes affecting female reproduction in the chosen region. These data provide useful information for female reproductive genes since relatively little attention has been paid to this group. Female reproductive genes are expected to be also highly diverged because of the correlated evolution of male and female reproduction system due to sexual conflict or coadaptation. Divergences for individual genes reveal a higher evolution rate for male-reproductive genes than female-reproductive genes. These results are consistent as male reproductive genes are expected to be subjected to stronger selection pressures than female reproductive genes.

2.11 The Phylogeny among the *D. melanogaster* Complex

Although the phylogenetic relationship of the *D. simulans* clade has been the subject of many studies, the phylogeny still remains to be an unresolved trichotomy. Various investigations have suggested the discordance between the phylogenies inferred from molecular differentiation and the reproductive divergence in physiology, sexual behavior or morphology, which is possibly the result of the shared ancient polymorphism and/or gene introgression during secondary contact (Palopoli, Davis, and Wu 1996; Ting, Tsaur, and Wu 2000). Even in the molecular phylogenetic trees, the topologies differ from locus to locus (Nei 1987). The combined data across loci should be used in order to overcome the noises of the ancestral polymorphism or gene introgression. In this study, a series of genes were used and supported the hypothesis that *D. mauritiana* and *D. simulans* were more closely related to each other than either was to *D. sechellia*. The same divergence topology was obtained by other studies such as reproductive status comparison (Lachaise et al. 1986), the restriction map analysis of mitochondrial DNA (Solignac, Monnerot, and Mounolou 1986) and the microsatellite-based multilocus examination (Harr et al. 1998).

2.12 The Fast Sex Theory for Hybrid Sterility Component of Haldane's Rule

High divergence of sex genes leads support to the hypothesis that the fast evolution of sex genes may be a predominant force in the production of hybrid sterility and in the explanation of Halane's rule (Singh 2000). Under the fast-sex theory, one could predict the excess of hybrid male sterility in heterogametic-male species and hybrid female sterility in heterogametic-female species. This theory has the same assumption as

the dominance theory that sex genes are on the average recessive as to their hybrid sterility effect on the hybrid genetic background. Two investigations, which have been performed recently to test the large X effect on hybrid sterility, have by the way showed the advantages of the fast-sex theory over the dominance theory. Turelli and Begun (1997) compared the waiting time for the hybrid sterility in *Drosophila* species having large and small X chromosomes. The result showed that *Drosophila* species with large X chromosomes evolved hybrid sterility faster than species with small X chromosomes. This observation is compatible with both the dominance theory and the fast-sex theory that large X chromosome species harbor more recessive X-linked incompatibilities than small X chromosome species. However, a test made in Lepidoptera gave mixed results. Lepidopteran has tiny X chromosome compared to *Drosophila* (Presgraves 2002). If the dominance theory is correct, Lepidopteran should have evolved hybrid sterility much slower than *Drosophila.* But this is not the case. The data suggested that Lepidopteran evolve hybrid sterility as fast as large-X *Drosophila* and nearly twofold than small-X *Drosophila* (Presgraves 2002). The result is expected under the fast-sex theory considering the different accumulation and evolution rate of sex genes on X chromosomes. Wang et al. (2001) have observed an extraordinarily high abundance of sex genes in mammalian X chromosomes. Although there is lack of data comparing the abundance of sex genes between Lepidopteran and *Drosophila*, it is reasonable to assume that Lepidopteran has accumulated more sex genes on the X chromosome than *Drosophila.* But even if the number of sex genes are comparable on the X and autosomes, the rapid evolution of sex genes and not their number per se, may hold the key to the

explanation of Haldane's rule. Swanson et al. (2001a) have observed disproportionately few accessory gland protein genes (*Acps*) on the X than on autosomes in *Drosophila*. As to the test on "unbalanced" hybrid females carrying both X chromosomes from the same species on an otherwise hybrid genetic background, the dominance theory failed to explain this observation, whereas it is not surprising under the fast-sex theory. Female reproductive characters have been suggested to evolve at a rate that is faster than non-reproductive characters but slower than male reproductive characters in this study, probably due to the typically more intense sexual selection on males. As a result, hybrid heterogametic-males display sterility pattern while hybrid females remain fertile in the comparable hybrid genotypes. In the heterogametic-female species, this evolution rate distinction is overcome by the recessivity. In the species lacking a degenerate Y chromosome, since males are not hemizygous for the X chromosome, recessive hybrid steriles are not expected to affect males more than females under the dominance theory. Thus the two sexes are expected to be equal in hybrid fitness, which disagrees with the cross experiment in one such species, Aedes (Presgraves and Orr 1998). The fast-sex theory can extend beyond taxa with a hemizygous sex and attribute the excess of hybrid male sterility in such organisms to the faster accumulation of male steriles.

This fast-sex theory can explain all the exceptions to the dominance theory and it can apply to both male-heterogametic, female-heterogametic as well as to non-heterogametic taxa. However, as a test of its power, we must predict what type of result(s) would be against the fast-sex theory. One such result, obviously, would be a demonstration of the involvement of non-sex genes in the production of species hybrid

sterility. Demonstration of the evolution of hybrid inviability before hybrid sterility would falsify this theory.

CHAPTER 3

An Attempt to Localize Hybrid Sterility Genes between *D. mauritiana* and *D. simulans*

ABSTRACT

The introgression of the *car* region from *D. mauritiana* into *D. simulans* has been shown to have a full effect on hybrid sterility. In the present study, we attempted to further explore this effect by investigating *car* introgression region from *D. mauritiana* in the *D. simulans* genetic background. Thirteen-generations of continuous backcrosses were used to produce 20 individual iso-female lines carrying male sterility factors. The phenotypes of testes for pure species, as well as for hybrids, were examined. The use of molecular markers revealed that a region from 17D to 19A and an unknown region are capable of causing hybrid sterility. The failure in the attempt to trace this unknown region is possibly due to the incorrect selection of morphological marker in every generation. The possible genetic basis of hybrid sterility and the recombination mapping analysis as a general tool were discussed.

One of the best-studied cases of hybrid sterility is between *D. mauritiana* and *D. simulans*. It has been suggested that these two species have diverged at a large number of loci with varied effects in terms of hybrid male sterility (Perez et al. 1993; Cabot et al. 1994; Perez and Wu 1995). Cabot et al. (1994) introgressed the distal one-fourth of the X chromosome from *D. mauritiana* into the *D. simulans* genome. They could not reject the single gene model for hybrid sterility in their low-resolution mapping. However, their high-resolution mapping using a series of molecular markers revealed that a minimum of three distinct factors from *D. mauritiana* were required to confer full hybrid sterility. Each individual factor by itself was relatively ineffective. Davis and Wu (1996) chose the region marked by one of the three morphological markers used by Coyne and Charlesworth (1989) in the same species pair for detailed recombination analysis. Three loci of hybrid sterility were revealed in a presumably single locus. When one of these loci was further dissected, four loci of hybrid sterility were identified. Thus, at least six loci were responsible for hybrid male sterility between this species pair in the region of 3% of the *Drosophila* genome. The number of loci responsible for complete hybrid sterility seems to increase with the ability to decompose the introgression by recombination, as exemplified by above studies in the same region of the same species pair. The density of hybrid male sterility thus becomes a function of resolution of analysis. So far, polygene basis of hybrid sterility has been repeatedly demonstrated in studies with high-resolution. The existence of single gene capable of causing complete sterility can not be ruled out though.

Zeng and Singh (1995) selectively used *D. simulans – mauritiana* hybrid heterozygous females producing fertile and sterile sons in roughly equal proportions to continuously backcross to *D. simulans* males. In this case, it was assumed that either a major hybrid sterility gene, or a tightly linked gene complex with large effect was selected. The recombination analysis was applied and a putative major sterility gene or tightly linked gene complex was mapped at position 62.7 ± 0.66 on the X chromosome of *D. mauritiana* in the *D. simulans* background. In this study, we attempted to finer localize this gene or gene complex by introgression analysis using a series of morphological and molecular markers.

MATERIALS AND METHODS

3.1 Strains and Mutants

Two sibling species, *D. simulans* and *D. mauritiana,* were used. The strains of *D. simulans* carry visible markers of *car* (18D, 1-62.5) and *Bx* (17C, 1-59.4). The cytological locations are the polytene chromosome bands and the recombination locations are the map positions of the same genes in *D. melanogaster*, which differ slightly from those of *D. simulans* but retain the same linear order (Lemeunier and Ashburner 1976). The basic stocks for *D. simulans* are stock 0251.68 (carrying *car*) from the Arizona University stock center, and a *Bx* stock which was kindly provided by J. Coyne. The strain bearing a combination of *car* and *Bx* markers was assembled by recombination from basic stocks. A *D. mauritiana* wild-type strain was originally from Umea stock center. All *Drosophila* stock and crosses were reared in 35ml glass vials with standard banana medium (Table2.1). The temperature was 25°C in an incubator under a 12 hour dark/light cycle.

3.2 Introgression Scheme

We intended to get the introgressions from *D. mauritiana* to the *D. simulans* genome of the region marked by *car* by repeated backcrossing. Flies with *car* mutation display dark-ruby eye color. The introgression from *D. mauritiana* with the wild type car^{+} locus will recover the eyes to normal red color. The experiment was based on the mating scheme shown in Figure 2.1. Hybrid F1 virgin females from female *D. simulans*

Table 2.1 Standard banana medium for *Drosophila*

Water	1800 ml
agar	20 g
bananas	2 medium sized
corn syrup	2 tbsp
brewer's yeast	60 g
tegosept	10 g of methyl p-hydroxy benzoate in 100 ml 95% ethanol

Gradually add agar in boiled water while stirring in a large flask. Mix all remaining ingredients except tegosept in a blender, return to large flask and boil. Cool while stirring to 55 °C. Add 36 ml of tegosept. After careful mixing, pour the media into vials. Cap the vials and place them in cold storage.

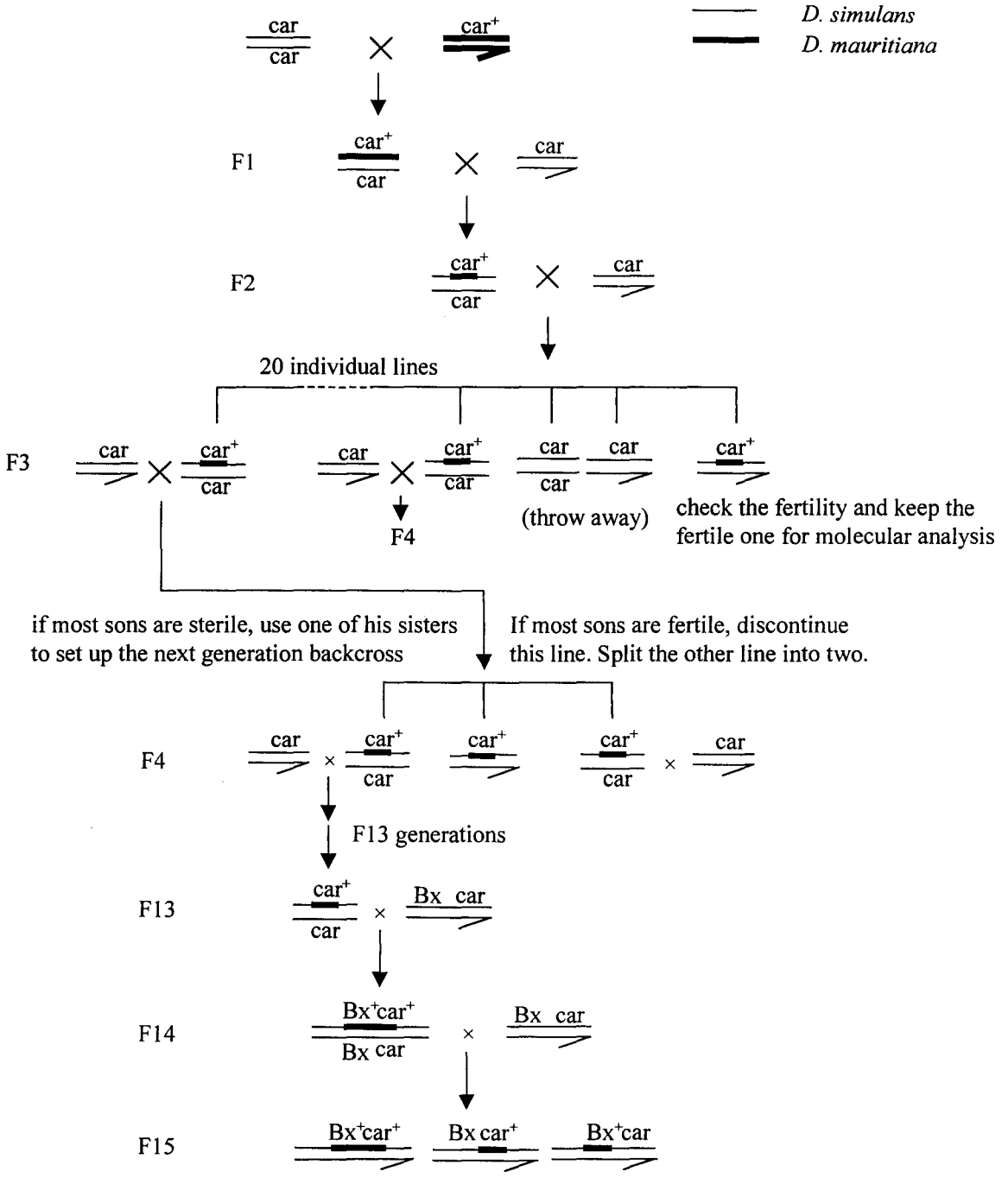


Fig. 2.1 Mating scheme to create *D. mauritiana* introgressions in the *car* region of *D. simulans*

with *car* marker (denoted as sim-*car*) and male *D. mauritiana* (denoted as mau) were mass backcrossed to sim-*car* males. The progenies were aged for 2-3 days to separate individuals with wild type eyes with those harboring mutant eyes. The resulting F2 single virgin females with wild type eyes were then individually backcrossed to *sim-car* males (one female with three to five males) and a number of next generation individuals were generated. 20 such individual lines were maintained for each generation. One of the daughters carrying [car^+] (brackets represent the introgressed chromosome material from *D. mauritiana*) was used to set up next generation backcross. The backcross scheme was kept for 13 generations. As a result, chromosomes of the hybrids will be gradually substituted by those of *sim-car* except for the car^+ locus. From F3 on, individual [car^+] males were dissected to check the presence of motile sperm in the seminal vesicle and the proportion of sterile males was scored. Some [car^+] males in some lines will become fertile due to the absence of the sterility factors by recombination. The lines which produced 50 % or above of sterile [car^+] males were kept for backcrosses. If the line produced more than 50 % of fertile [car^+] males, we assumed that the mother did not carry sufficient sterility factors. In this case, this line was discontinued and the fertile [car^+] males were kept in –20 °C for molecular analysis. Another male-sterile line was split into two to keep the total number of lines to be 20. Fig 2.2 shows the female origin in individual lines. Ultimately, 14 individual iso-female lines of introgressions from *D. mauritiana* in the *D. simulans car* region which produced sterile males were obtained.

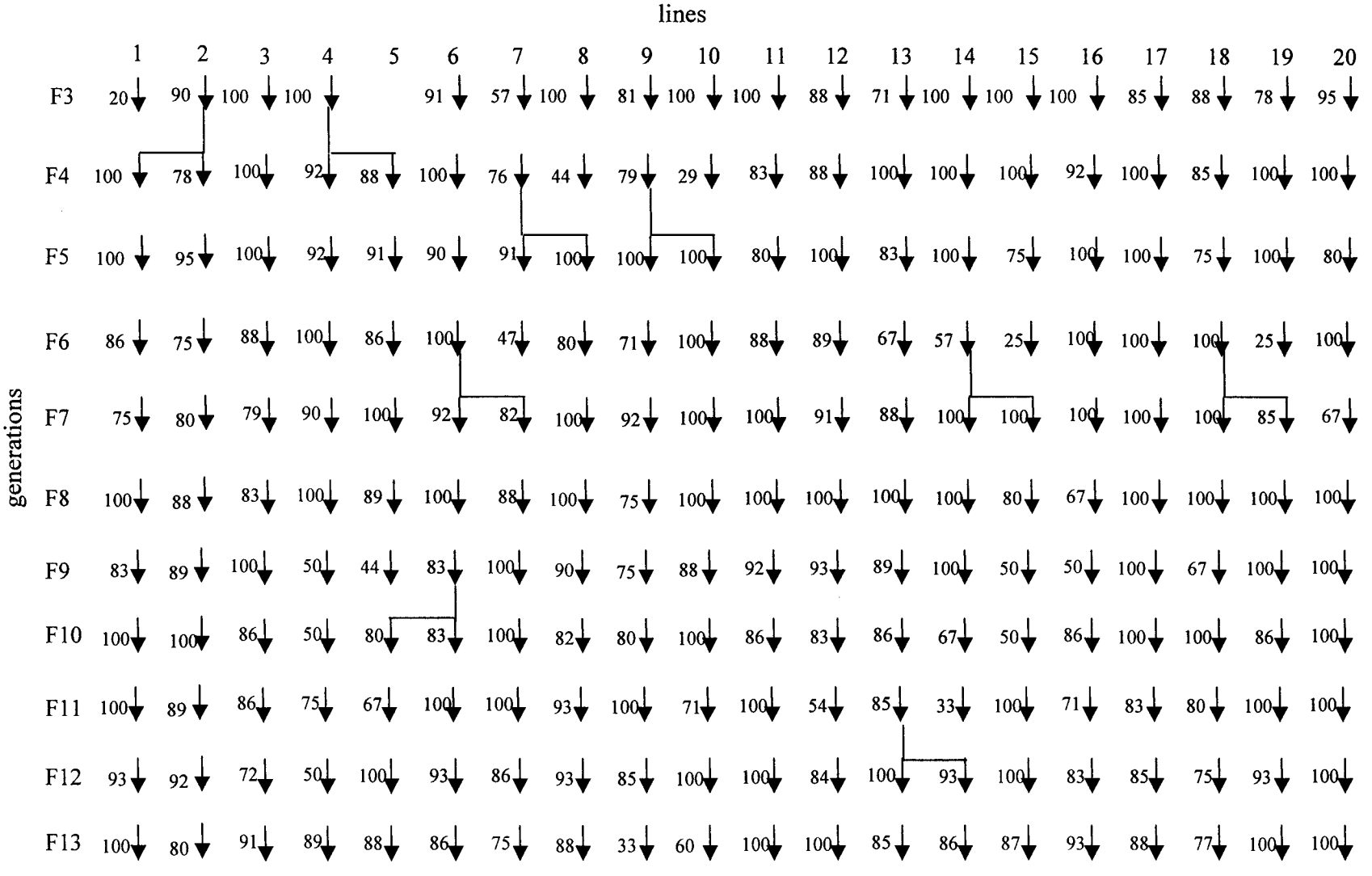


Fig. 2.2. Sterility rates for backcross generations in individual lines. If the rate was below 50%, the original line was discontinued and the next line was split into two.

After 13 generations of backcrosses, *Bx* marker was introduced into the lines for recombination analysis by crossing hybrids with *Bx*- and *car*-marked *D. simulans*. We have assumed that the introgressions have extended to the *Bx* marker, therefore, the Bx^+ region was considered to be from *D. mauritiana*. Three male genotypes were expected: [Bx^+ car^+]/Y, Bx[car^+]/Y and [Bx^+]car/Y. Individual males were scored for the sterility and sterile males were kept in -20 °C for molecular analysis.

3.3 Sterility/Fertility Checking

In each generation from F3 on, 2-3 days old males are checked for the presence of motile sperm in the seminal vesicle. Under a CARL ZWISS dissecting microscope, the reproductive tract is retrieved from a live fly while submerged in a drop of PBS buffer (Table 2.2). Testes and seminal vesicles are then separated from the rest of reproductive tissues and covered by a coverslip. The coverslip was gently squash and the tissues were promptly examined under phase-contrast and Nomarski optics using an AXIOSKOP ZWISS microscope. The males were considered to be fertile if there was any motile sperm. The cytological pictures of the testes were photographed by QuickImage™ 24 in Macintoch IIVX computer connected to ZWISS microscope.

3.4 Molecular Analysis

3.4.1 Development of Molecular Markers

Molecular markers were used to examine the extent of the introgression more precisely. Genes that yield species-diagnostic patterns of RFLP between *D. simulans* and *D. mauritiana* in Chapter 2 were utilized as molecular markers to determine the

Table 2.2 PBS buffer

Ingredients	
NaCl	8 g
KCl	0.2 g
Na_2HPO_4	1.44 g
KH_2PO_4	0.24 g

Dissolve above reagents in 800 ml of distilled H_2O. Adjust the pH to 7.4 with HCl. Add H_2O to 1000 ml. Dispense the solution into aliquots and sterilize them by autoclaving. Store at room temperature.

genotypes of the introgressed hybrids. These markers are *e(y)1* gene digested by MseI (denoted as *e(y)1* / MseI), *Dhc16F* / EcoRV, *os* / HincII, *CrebB* / CfoI, *fu* / MseI, *Cyp18* / ClaI, *pcm* / DdeI, *Bap* / HincII, *CG9569* / XhoI, and *Hlc* / XhoI.

3.4.2 DNA Extraction

Individual fertile males from F3 to F13 and sterile males from F15 were DNA extracted. DNA extraction follows protocol for small scale isolation of *Drosophila* DNA. Add 25 µl homogenization buffer (Table 2.3) to each fly in an Eppendorf tube. Squish the fly thoroughly. Incubate at 70 °C for 20 min. Add 10 µl 8 M Potassium Acetate and incubate on ice for 1 hour. Spin at the maximum speed for 1 min using table centrifuge. Dilute the supernatant with 300 µl ddH_2O and add 300 µl phenol. Shake thoroughly. Spin at maximum speed for 1 min. Transfer the aqueous phase into new tube and add 250 µl phenol and 250 µl chloroform. Spin at maximum speed for 1 min. Transfer the upper layer and add 300 µl chloroform. Spin at maximum speed for 1 min. Collect the upper layer and add 1/10 volume of 3 M sodium acetate and 2.5 times 95 % ethanol. Keep it in –20 °C overnight. Spin at maximum speed for 10 min. Remove the liquid and wash the pellet with 70 % ethanol. Air dry pellet and dissolve it in 25 µl EB buffer (10 mM Tris-HCl, pH8.5) or distilled water.

3.4.3 PCR-RFLP to Determine the Extent of the Introgression

PCR amplification and restriction enzyme digestion for the hybrids follows the protocol in the sex and non-sex gene comparison experiment in chapter 2. The origin of the chromosome segment can be determined by comparing the RFLP pattern of hybrids

with those of pure species. The segment is of *D. mauritiana* origin if the associated marker display the same RFLP pattern with *D. mauritiana*, otherwise, the segment is from *D. simulans*. Each introgression line was checked with many markers to determine the genotype.

Table 2.3 Homogenization buffer

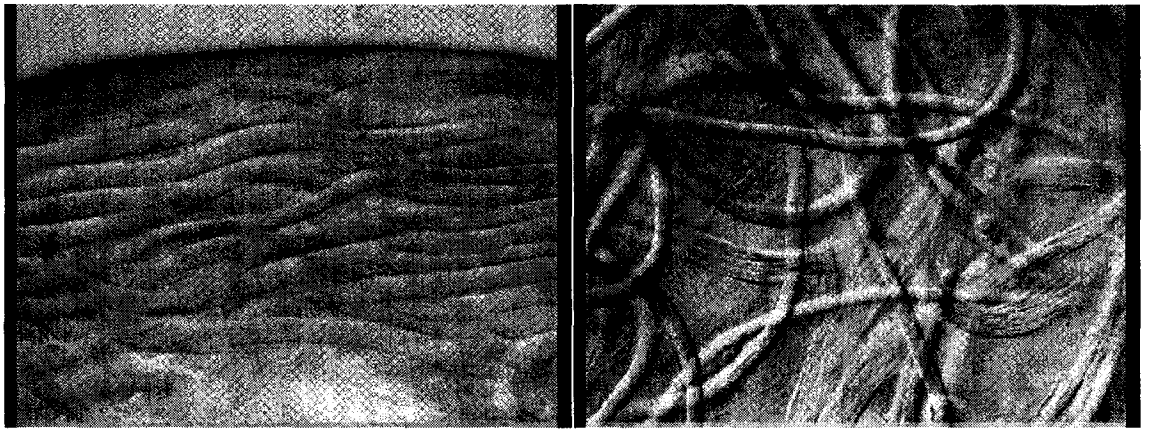
Ingredients	
1 M Tris-Hcl (pH 9.0)	0.5 ml
0.5 M EDTA	0.1 ml
10 % SDS	0.5 ml
Add distilled water to total volume of 50 ml.	

RESULTS

3.5 Spermatogenic Phenotypes

In the normal spermatogenesis process, each cyst of 64 spermatids will transform into a sperm bundle, which holds sperm in orderly parallel arrangement until their individualization in the basal end of the testis. After this, sperm are stored in the seminal vesicle until its utilization in reproduction. The presence of motile sperm in the seminal vesicle and the sperm bundles in the testes were examined for the introgression lines as well as for pure species. Pure species were basically used as reference. The pure species are normally fertile, as evidenced by a large number of motile sperm in the seminal vesicles and their healthy appearance and normal size of their testes. A handful of tightly packed sperm bundles can be observed in the mid-testis regions (Fig 2.3a). Coiling sperm bundles are evident in the basal region of the testes. After a gentle squash, dislodged sperm bundles are released (Fig 2.3b) and coiled individual motile sperm can be found (Fig 2.3c).

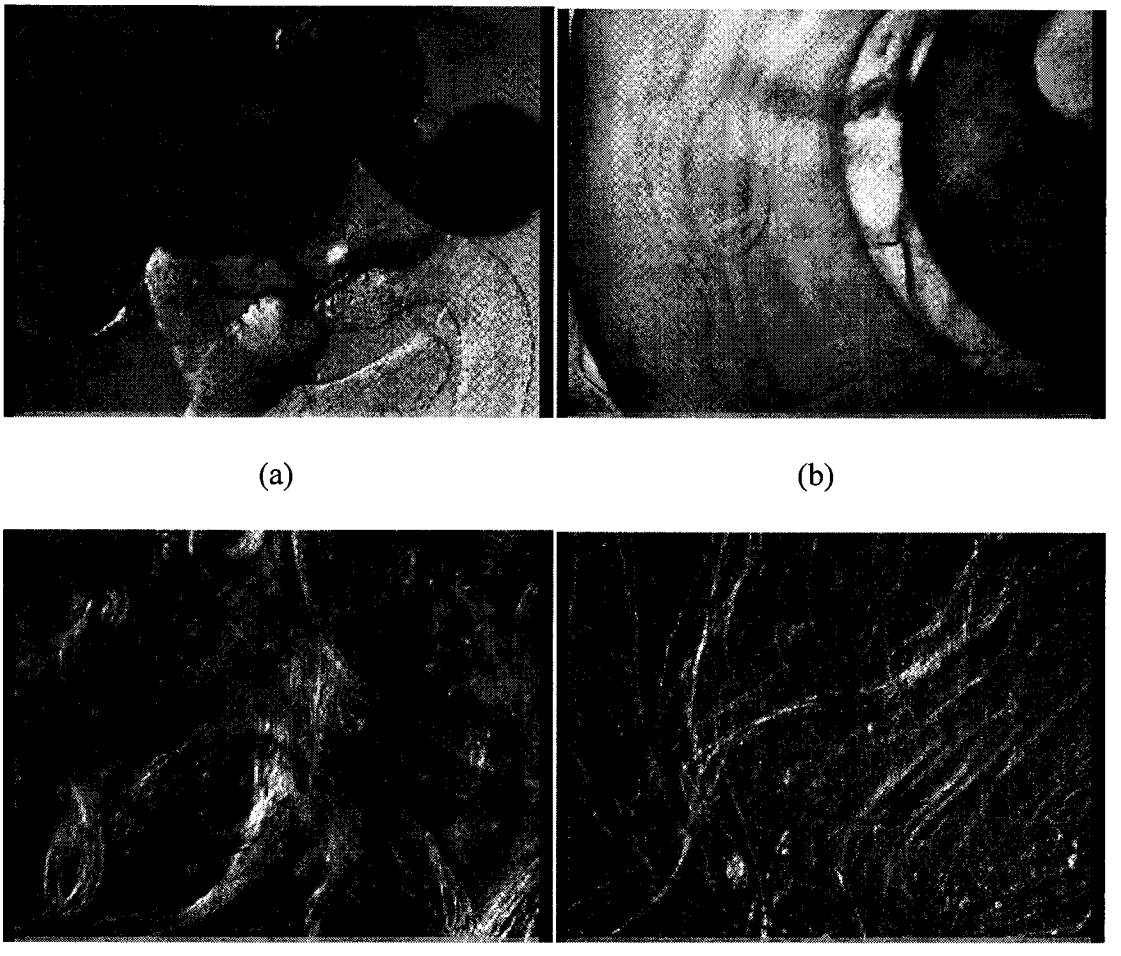
In F1 hybrids, most of the testes and seminal vesicles have normal appearance. Typically, testes with abnormal shape display an enlarged apical end. Some seminal vesicles are relatively small (Fig 2.4a). Most of F1 hybrids exhibit a number of conspicuous sperm bundles which are disheveled in the mid-testis region (Fig. 2.4b), in accord with previous observations (Kulathinal 1996). The presence of coiled sperm



(a) (b)

(c)

Fig. 2.3 Testis phenotypes of wild-type species. (a) sperm bundles in the mid-testis region. (b) sperm bundles released from the testis after being squashed before individualization. (c) individualized coiled sperm.



(a) (b)

(c) (d)

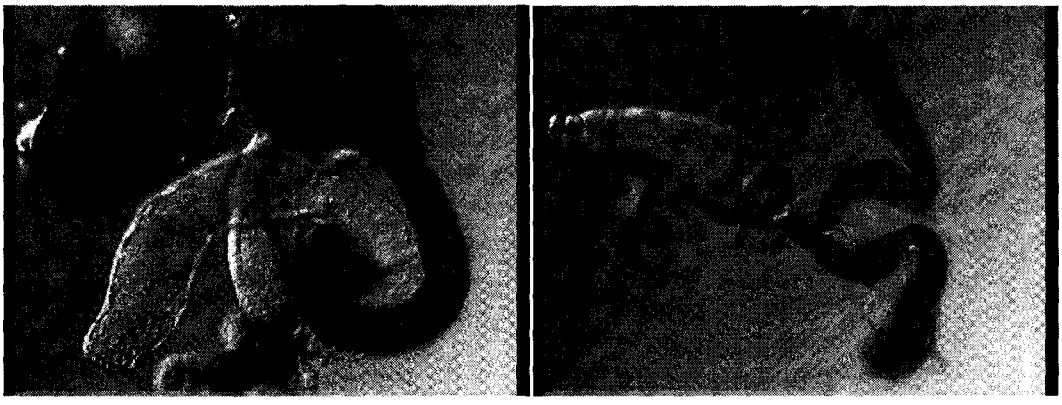
Fig 2.4 The spermatogenic defects in F1 hybrids. (a) Whole mount view for testes and seminal vesicles. Accessory glands and anterior ejaculatory duct were shown. (b) Sperm bundles in mid-testis region. (c) Sperm bundles released from the testis are disheveled. (d) Sperm are in disarray and non-motile.

bundles in the basal region is not apparent. When released from the testes, sperm bundles display a diffuse pattern and are loosely packed (Fig 2.4c). Sperm are relatively small in number and are in disarray (Fig 2.4d). No motile sperm were observed in seminal vesicles, which were instead full of cellular debris.

In the following backcross generations, except for the same abnormal appearance as seen in the F1 hybrids, other types of abnormal testes were observed. Some testes are extremely irregular and have no distinct three parts of the normal testis (apical end of testis, middle of testis and basal region, Fig 2.5a). Some testes are atrophied with small seminal vesicles (Fig 2.5b). Sterile hybrids usually demonstrate a lack of sperm bundles at the mid-testis region. Few testes are completely devoid of sperm bundles. Sperm are in disarray or there is no obvious sperm at all. Some fertile hybrids can be observed to possess considerable number of sperm bundles but sperm bundles were usually defective, as evidenced by their diffuse and thick appearance. The number of motile sperms in fertile hybrids is relatively small.

3.6 Scores for Sterility/Fertility in Hybrids

From F3 on, individual flies were checked for the fertility by examining the presence of motile sperms. All males were aged for 2 or 3 days before dissected for sperm checking. For each line in each generation, we tried to examine approximately 20 individuals. Table 2.4 shows the segregation rate for sterility for hybrid generations. In the first backcross generations, the genetic background of hybrids is mixed so that the hybrid sterility can not be assigned to the action of any sterility factors. Therefore, we



(a) (b)

Fig. 2.5 Abnormal testes in backcross generations. (a) extremely irregular testis. (b) atrophied testes with small seminal vesicles.

Table 2.4 Sterility segragation rate for hybrid generations

line	F3			F4			F5			F6			F7		
	No. of sterility	No. of fertility	sterility rate (%)	No. of sterility	No. of fertility	sterility rate (%)	No. of sterility	No. of fertility	sterility rate (%)	No. of sterility	No. of fertility	sterility rate (%)	No. of sterility	No. of fertility	sterility rate (%)
1	1	4	20	3	0	100	10	0	100	6	1	86	6	2	75
2	19	2	90	18	5	78	20	1	95	6	2	75	4	1	80
3	19	0	100	4	0	100	16	0	100	7	1	88	11	3	79
4	13	0	100	11	1	92	11	1	92	4	0	100	9	1	90
5				14	2	88	10	1	91	12	2	86	6	0	100
6	10	1	91	6	0	100	9	1	90	6	0	100	11	1	92
7	6	3	57	13	4	76	10	1	91	7	8	47	14	3	82
8	3	0	100	4	5	44	14	0	100	12	3	80	11	0	100
9	13	3	81	11	3	79	5	0	100	15	6	71	12	1	92
10	4	0	100	4	10	29	8	0	100	10	0	100	14	0	100
11	2	0	100	10	2	83	8	2	80	14	2	88	8	0	100
12	14	2	88	15	2	88	2	0	100	8	1	89	10	1	91
13	12	5	71	8	0	100	10	2	83	4	2	67	7	1	88
14	4	0	100	4	0	100	11	0	100	4	3	57	16	0	100
15	1	0	100	9	0	100	3	1	75	2	6	25	10	0	100
16	8	0	100	11	1	92	1	0	100	9	0	100	6	0	100
17	11	2	85	8	0	100	6	0	100	11	0	100	15	0	100
18	7	1	88	11	2	85	3	1	75	16	0	100	9	0	100
19	14	4	78	18	0	100	5	0	100	1	3	25	11	2	85
20	18	1	95	12	0	100	4	1	80	14	0	100	2	1	67
Total	179	28	86	194	37	84	166	12	93	168	40	81	192	17	92

Continued

line	F8			F9			F10			F11			F12			F13		
	No. of sterility	No. of fertility	sterility rate (%)	No. of sterility	No. of fertility	sterility rate (%)	No. of sterility	No. of fertility	sterility rate (%)	No. of sterility	No. of fertility	sterility rate (%)	No. of sterility	No. of fertility	sterility rate (%)	No. of sterility	No. of fertility	sterility rate (%)
1	11	0	100	15	3	83	11	0	100	16	0	100	13	1	93	20	0	100
2	14	2	88	8	1	89	14	0	100	8	1	89	11	1	92	16	4	80
3	10	2	83	3	0	100	6	1	86	6	1	86	13	5	72	10	1	91
4	4	0	100	3	3	50	7	7	50	3	1	75	6	6	50	8	1	89
5	8	1	89	4	5	44	8	2	80	4	2	67	5	0	100	14	2	88
6	9	0	100	5	1	83	15	3	83	11	0	100	14	1	93	18	3	86
7	15	2	88	18	0	100	4	0	100	9	0	100	12	2	86	15	5	75
8	16	0	100	9	1	90	9	2	82	13	1	93	13	1	93	15	2	88
9	9	3	75	9	3	75	4	1	80	11	0	100	11	2	85	6	12	33
10	5	0	100	7	1	88	16	0	100	5	2	71	11	0	100	12	8	60
11	5	0	100	11	1	92	6	1	86	12	0	100	9	0	100	11	0	100
12	13	0	100	14	1	93	5	1	83	7	6	54	16	3	84	12	0	100
13	8	0	100	16	2	89	6	1	86	11	2	85	4	0	100	17	3	85
14	12	0	100	16	0	100	4	2	67	2	4	33	13	1	93	6	1	86
15	4	1	80	1	1	50	3	3	50	15	0	100	4	0	100	13	2	87
16	2	1	67	1	1	50	6	1	86	5	2	71	5	1	83	13	1	93
17	2	0	100	15	0	100	12	0	100	10	2	83	11	2	85	7	1	88
18	1	0	100	2	1	67	4	0	100	8	2	80	6	2	75	10	3	77
19	4	0	100	7	0	100	6	1	86	16	0	100	14	1	93	19	0	100
20	3	0	100	2	0	100	11	0	100	11	0	100	12	0	100	19	0	100
Total	159	12	93	166	25	87	157	26	86	183	26	88	191	29	87	261	49	84

focused on the segregations after F8 when the background is sufficiently pure. The average segregation rate for F8 is 93% which means that 93 individuals out of 100 carry the putative sterility factors (denoted as PSF) but seven of them lose the sterility due to the crossover between *car* locus and the PSF. Thus, four genotypes with the wild-type eye color were generated: 93% [car^+ PSF]/Y, 7% [car^+]/Y, 93% [car^+ PSF]/*car* and 7% [car^+]/*car*. We chose one of the females to carry on the backcrosses. There are more females with the genotype [car^+ PSF]/*car* than females with [car^+]/*car*. We have better chance to pick up the females carrying the PSF. If we have picked up the [car^+]/*car* genotype, we will obtain no sterile progeny or low level of sterility. If so, this line is discontinued. In F9, the average male sterility rate is 87%, which indicates that we have picked up the females carrying the PSF from the previous generation. The sterility rates remained at a high level in the following generations. The fertile individuals were subsequently mapped with molecular markers.

Fig 2.6 shows the distribution of segregation rates and the number of individuals with the abnormal testis appearance in 20 individual lines for each generation after F8. In F8, majority of lines has the segregation rates above 80%. There are 12 individuals that have the unhealthy testis appearance, which is 7% of males being checked. They all belong to the lines with more than 80% segregation rates. In F9, 22 individuals (11.5%) have the abnormal testis appearance. The segregation rates for them range from 50% to 100%. Twenty-five (13.7%) F10 males have the abnormal testes. In F11, thirty-two out of total 209 (15.3%) individuals have the abnormal testes. There are thirty-five (15.9%)

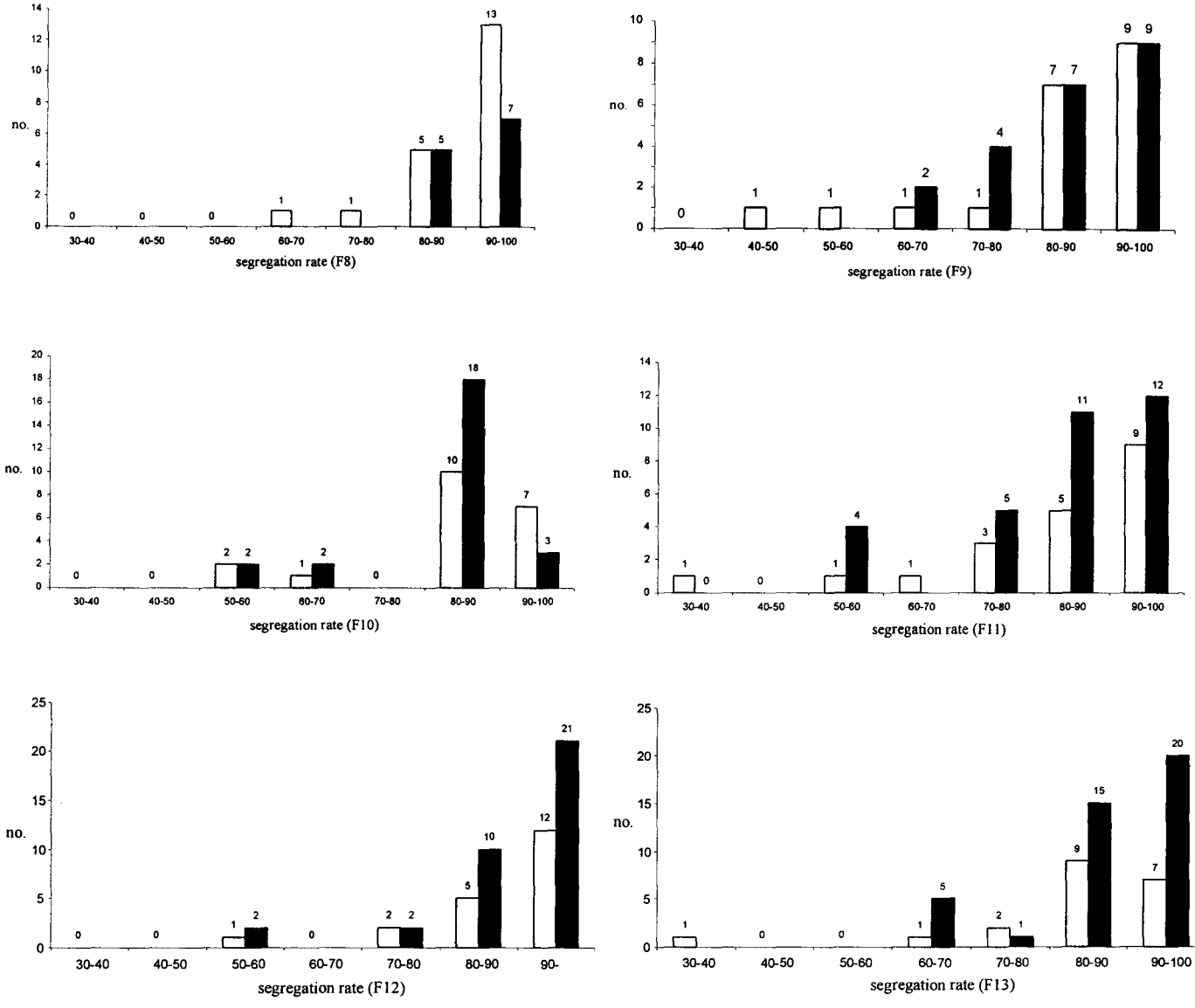


Fig. 2.6 Distribution of segregation rates in individual lines. Open bars represent the number of sterile lines whose segregation rates fall into certain range. Filled bars represent the number of individuals with the abnormal testis appearance.

and forty-one (13.2%) individuals that display the abnormal testes in F12 and F13, respectively. Most of the individuals with the abnormal testes have fallen into the lines that have a high level of segregation rate, suggesting a link between hybrid sterility and alterations of testis appearance in backcross generations.

After 13 generations of backcrossing, we obtained twenty individual lines carrying male-sterile introgressions from *D. mauritiana* in the genome of *D. simulans*. One of the females carrying *car* marker and male-sterility introgression from each line was crossed to *D. simulans* with the combined markers *car* and *Bx* (Fig 2.1). This was designed to shorten the introgressions and thus narrow down the regions for hybrid sterility factors. It was supposed to generate three introgression genotypes for males: [Bx^+ car^+]/Y, [Bx^+] *car*/Y and *Bx*[car^+]/Y, assuming that the introgressions have extended beyond *Bx* marker. By molecular analysis shown below, Bx^+ marker has been shown to be originated from *D. simulans*, that is, the introgressions from *D. mauritiana* do not reach *Bx* locus. Those genotypes are possibly Bx^+[car^+]/Y, Bx^+*car*/Y and *Bx*[car^+]/Y instead. The males were scored for fertility. Table 2.5a and b show the sterility data.

3.7 Molecular Analysis

The fertile males during the backcross generations and sterile males in the recombination generation were subjected to molecular analysis. 151 fertile males and 92 sterile males were collected to map with molecular markers. The hybrid male sterility factor is expected to be located to the segment that is absent in the longest fertile introgression but present in the shortest sterile introgression.

Table 2.5a F15 sterility data ($Bx^+[car^+]$ or $Bx[car^+]$)

line	No. of sterility	No. of fertility	sterility (%)
2	23	2	92
3	19	2	90
5	5	0	100
6	20	1	95
7	11	3	79
8	13	0	100
9	10	2	83
10	8	3	73
11	10	3	77
12	10	0	100
13	11	1	92
15	10	1	91
16	11	5	69
17	7	1	88
18	5	1	83
19	11	4	73
20	9	3	75
total	193	32	86

Table 2.5b F15 Sterility data (Bx^+car)

line	No. of sterility	No. of fertility	sterility (%)
2	11	0	100
3	8	1	89
5	3	0	100
6	9	1	90
7	5	3	63
8	9	1	90
9	4	1	80
10	3	1	75
11	4	3	57
12	1	0	100
13	4	2	67
15	2	0	100
16	6	2	75
17	1	0	100
18	5	0	100
19	5	0	100
20	5	0	100
total	85	15	85

We attempted to use a series of RFLP molecular markers to precisely demarcate the extent of introgressions. Comparison of RFLP patterns from introgression lines with those from the pure species reveals the origin of the chromosome segment. The results are shown in Fig 2.7. There are four individuals (line #16) containing the sterile introgressions covering at least from polytene chromosome band 17D to 19A. In contrast, none of the fertile lines was revealed to contain the introgressions beyond *car* locus. This pattern suggests that at least one sterility factor located at the interval of 17D to 19A.

However, although the morphological marker *car* indicates the existence of the introgression from *D. mauritiana* in the *car* locus in $Bx[car^+]$ or $Bx^+[car^+]$ males, molecular checking failed to detect any introgression from *D. mauritiana* at neighboring loci in the remaining 88 sterile males and 151 fertile males. The sterility effects were unable to be assigned to a specific introgression region.

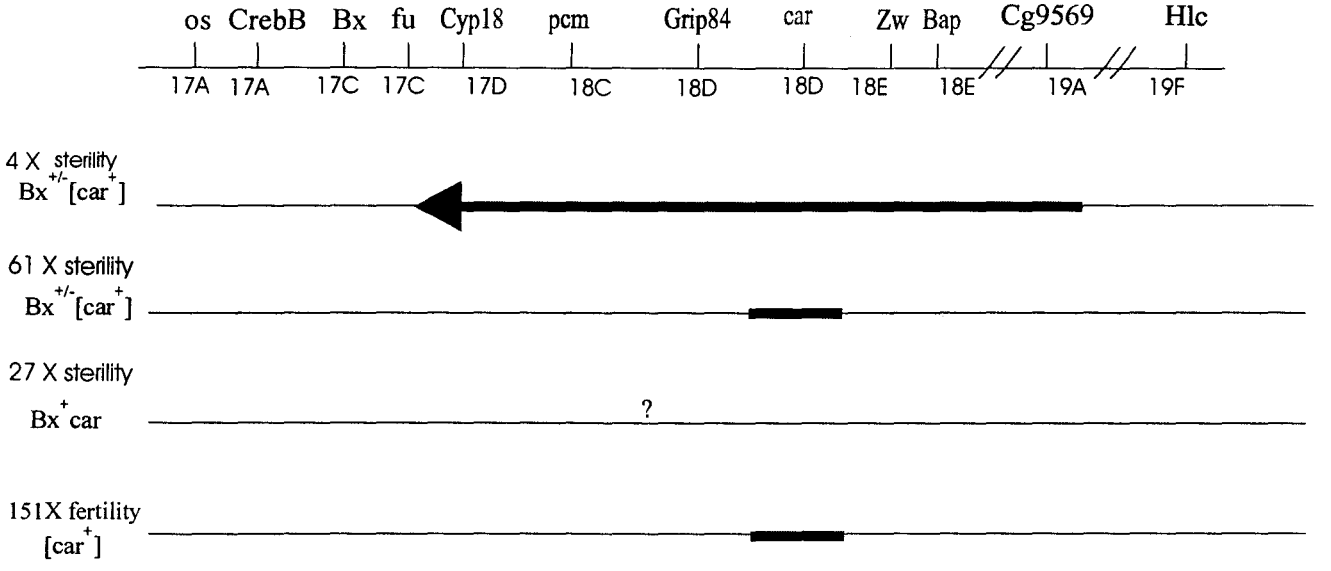


Fig. 2.7 Molecular analysis of sterile and fertile males. The bold lines indicate the chromosome material introgressed from *D. mauritiana*. The arrowhead indicates that the introgression from *D. mauritiana* is beyond a particular molecular marker. The end without the arrowhead indicates that the introgression does not pass the next molecular marker.

DISCUSSION

3.8 Cytological Examination of Sterility

The cytological analysis of sterility demonstrates an interesting phenomenon. In F1 hybrid males, whose genome is composed of half of each species genome, the testes showed a healthier phenotype than later backcross generations that are supposed to contain relatively few candidate genes that could act as hybrid sterility factors. Despite the complete sterility in F1, alterations of the size or the morphology of the testes were less profound. A relatively small introgression thus appears to have a more drastic effect in testis appearance than the combination of two half genomes, despite the fact that the latter contains more foreign genes. This may indicate a balancing epistatic effect of the chromosome within each half genome, suggesting the difference of incompatibilities causing testis phenotype and sterility. However, after the genomic background is pure enough except for the small introgression from other species, more cases of abnormal testis were observed when the introgression had a larger effect in hybrid sterility, as shown in Fig 2.6. Most abnormal testes concentrated in the lines with high proportion of sterility.

3.9 Major Gene or Polygene Basis of Postzygotic Isolation

Although the nature of hybrid sterility has become a subject of extensive studies for decades, a simple but fundamental question, how many genes are required to cause hybrid sterility, has not been answered yet. Many previous studies have tentatively drawn

the conclusion that a single gene is capable of causing hybrid sterility (Wu and Beckenbach 1983; Coyne and Charlesworth 1986; Johnson et al. 1993; Perez et al. 1993). However, evidence has increasingly favored the view of multigenic basis of hybrid male sterility. Single gene of large effect appears not to be a common barrier to genetic exchange between closely related species. As stated above, those presumably major sterility components have been decomposed into several linked loci with small effects. Even in the study with sufficient resolution, the presumably major hybrid sterility genes, *Ods*, turned out to be ineffective in causing sterility without the cointrogression of neighbor genes (Perez and Wu 1995). A likely explanation is that since neighbor interacting genes are so closely linked, they were always to be found together in low resolution. The major sterility effect found by Zeng and Singh (1995) was located 6 cM away from *forked* locus (1-56.7). This rough recombination localization could correspond to a big region based on an average of 256 Kb per centimorgan on the X chromosome (Lefevre 1971). So far, there is lack of conclusive evidence for the presence of major sterility genes. Palopoli and Wu (1994) estimated that there are at least 40 loci harboring the hybrid sterility effect on the X chromosome alone between *D. simulans* and *D. mauritiana*.

Wright (1982) proposed that there are two types of possible interactions among multiple loci to cause hybrid sterility. Firstly, multiple genes could act additively to determine the sterility phenotype (Naveira and Fontdevila 1986, 1991). Secondly, complex gene interactions, such as epistasis, producing a much greater effect than any individual gene alone, could underlie the basis of sterility. The second view is more

important since a lot of studies have suggested the involvement of complex gene interactions. For example, three distinct factors from *D. mauritiana* were revealed to confer full sterility in *D. simulans*, each individual factor relatively ineffective by itself in causing sterility (Cabot et al. 1994).

The stage of divergence has been suggested to be an important component in the genetics of species differences (Powell 1997). Major hybrid sterility genes are possibly more common between species at their late stage of divergence because genes continue to accumulate differences between species after speciation. Such examples came from gene transformation studies in which several genes have exhibited strong effects on gene expression between divergent species (Brady and Richmond 1990; Seeger and Kaufman 1990). No such genes are evident in the introgression studies between incipient species, in which candidate genes could be more likely responsible for the initial establishment of reproductive isolation. The age of the species pair is not only reflected by the effect of sterility genes, but also by the number of genes. Only a modest number of genes were involved between Bogota and USA subspecies of *D. pseudoobscura* (Orr and Irving 2001). Whereas, a large number of genes are suggested to be involved between *D. simulans* and *D. mauritiana* (Palopoli and Wu 1994). Our results also suggested the involvement of at least two regions in the hybrid sterility in this pair. Unfortunately, in the recombination mapping study, it is impossible to differentiate genes involved in the speciation itself from those that might have diverged after the speciation event. It is more likely that all sterility genes detected are not required for the initial expression of sterility.

3.10 Development of Molecular Markers

The resolution of introgression analysis has been substantially increased but is limited by the availability of molecular markers. The lack of ability to distinguish the effect of one gene from that of several closely linked genes greatly hampers the progress of such studies. In the *Odysseus* introgression study between *D. mauritiana* and *D. simulans*, this gene can only be defined by boundary to be in an interval of 500 Kb (Perez et al. 1993), in which more recombinants can be defined if more markers in this region are available. Genes that show the distinct digestion patterns between two species can be utilized as molecular markers for the search for hybrid sterility genes. For example, in chapter 2, we have developed 10 RFLP markers between *D. mauritiana* and *D. simulans* and two of them are located in the interval of *Odysseus* gene. RFLP markers are more efficient than Southern blotting markers and SSCP markers (single stranded conformation polymorphism) because of their simple and quick operation. The development of markers is essential to reveal ultimately the number, effect and nature of hybrid sterility genes.

3.11 Reasons Not to Use Recombination Mapping as a General Tool

As mentioned above, many of the previous studies have revealed the polygene basis of hybrid sterility and most of chromosomes have been suggested to have some effect. Based on the study with only morphological markers, our results support the view that more than one region is capable of causing hybrid sterility. The sterility factors responsible for sterility of $Bx^{+}[car^{+}]$ or $Bx[car^{+}]$ in Table 2.5a are different from factors in Table 2.5b since the introgressions in Table 2.5a result in ruby eyes while the introgressions in Table 2.5b display the normal eye color. Thus, it appears that at least

two regions are capable of causing hybrid sterility from *D. mauritiana* in the *D. simulans* background. Unfortunately, we failed to provide detail information on the nature of these regions in the molecular analysis. Among the introgressions causing ruby eyes, the sterility penetrance ranges from 69% to 100% with a mean of 86% (Table 2.5a). Three out of twenty lines are 100% penetrant, suggesting the presence of sterility gene(s) with complete effect in the introgression. These three lines have distinct recombination processes (Fig 2.2), therefore, it is likely that they differ from each other the length or the crossover point of introgressions. The average sterility penetrance of introgressions causing normal eyes is 85%, with the distribution from 57% to 100% as shown in Table 2.5b. There are eight lines showing the complete sterility penetrance. However, recombinants of partial sterility are common. As suggested by Perez and Wu (1995), the prevalence of incomplete sterility penetrance within a certain marker genotype in recombination mapping produces unreliable results. For example, when 92% of the genotype in line #2 (Table 2.5a) were sterile, it is not clear whether most of the males carry a gene that has 92% sterility penetrance or whether 92% of the males carry a sterility gene of complete penetrance.

The uncertainty of the interpretation for recombination mapping was exemplified by *Ods* studies. *Ods* was firstly defined as a single gene with full sterility effect, located at cytological interval of 16D (Perez et al. 1993). However, when the introgression containing *Ods* were subdivided into mini-introgressions, the lines containing introgression from 16C to 16F only showed 39% sterility penetrance. Only lines with the introgression from 12B to 16D displayed complete sterility (Fig 2 in Perez and Wu

1995). Moreover, the sterility gene was assumed to be located in an interval that was contained in the shortest sterile introgression but was absence in the longest fertile introgression. This assumption excludes sterility genes with small effect present in the fertile introgression. Since epistatic interactions of weak effect sterility genes have been suggested underlying the genetics of hybrid sterility, this kind of study would underestimate the number of sterility genes.

The other problem in recombination mapping study is the selection of morphological markers. In this study, the *car* locus, affecting pigment cells of eyes thus displaying ruby color, was selected in every generation. However, in the molecular analysis, no introgressions from *D. mauritiana* near *car* locus were detected in most of ruby-eye males. Pigment cells are affected by a number of genes, such as *ade2*, *ca*, *cm*, *g*, and *pur*, and the resulting mutant males are eye color defective. One possible interpretation for the molecular analysis is that other genes affecting eye color were selected as the introgression instead of *car* locus in most of lines. Only four lines harbored the introgression from *car* locus. Our results suggest that there are sterility gene(s) with full effect in the interval of 17D to 19A and at least another sterility region exists in lines carrying other eye-defection genes. This is consistent with the view that the genetic differentiation is extensive, even between closely related species.

GENERAL DISCUSSION

In the incipient stage of speciation, hybrid sterility is the first trait to evolve, as suggested by the predominant appearance of hybrid sterility in species hybridizations. The understanding of the genetic basis of hybrid sterility is essential to finally untangle the genetics of speciation. The nature of genes underlying the genetic basis of the hybrid sterility component of Haldane's rule and the localization of sterility genes were two main subjects in this study. The fast evolution of sex genes, as a general group, corroborates previous studies on individual sex genes. The forces driving the evolution of sex genes are not clear in this study. The faster evolution of such genes might have little or no effect on fitness and randomly drift to fixation. The elevated sequence divergence at some sex determination genes has been suggested to be due to the relaxation of selection constraints (Kulathinal 2001). Alternatively, the fast evolution of sex genes might be a reflection of sexual selection and male-female coevolution, which has been revealed in a number of individual sex genes. Our study combined a large data set of sex genes. Although the Tajima's and KHA test did not reveal the obvious sign of positive selection, sex genes appear to evolve as a class of genes distinctly from those that are not related to sex and reproduction. The results of this study provide support for the fast-sex theory as an explanation for the hybrid sterility component of Haldane's rule in all taxa. This explanation makes up for the deficiency of the dominance theory and the fast-male

theory. It is important in the understanding of the genetics of speciation because Haldane's rule appears to be related to the onset of postzygotic isolation (Coyne 1992).

The attempt to localize sterility genes near *car* locus failed in finely mapping but a region ranging from 17D to 19A and another unknown region were revealed to display full effect to cause hybrid sterility between *D. mauritiana* and *D. simulans*. Wu and colleagues (Perez et al. 1993; Cabot et al. 1994; Palopoli and Wu 1994; Perez and Wu 1995; Davis and Wu 1996) are the leading group to take advantage of molecular markers to increase the resolution of conventional recombination study. They have shown that at least five X-chromosome regions are capable of causing hybrid sterility between *D. simulans* and *D. mauritiana* (Fig 3.1) and complex conspecific epistasis were suggested between sterility factors. The overall distribution of sterility factors was surveyed as well. Hollocher and Wu (1996) studied the density of sterility factors and concluded that there is approximately one sterility factor per 1% of autosomes and two per 1% of the X chromosome. True et al. (1996) respectively replaced 87 P element-marked regions from the *D. mauritiana* genome in a homozygous *D. simulans* background (Fig 3.1). A greater effect for introgression from the X chromosome than that of autosomes was revealed. The outcomes of the studies on genetics of hybrid sterility can be summarized as: (1) Hybrid male sterility results from the incompatibility between genes from two species, which all function properly in their original pure species; (2) The most common form of interaction between heterospecific chromosomes is the X-autosome interaction (Dobzhansky 1936;

Fig 3.1 Hybrid sterility regions on the X chromosome between *D. simulans* and *D. mauritiana* detected by high-resolution mapping

Cabot et al. (1994), mau → sim

Davis and Wu (1996), mau → sim

Palopoli and Wu (1994), sim → mau

Zeng and Singh (1995), mau → sim

Perez et al. (1993), mau → sim

Ture et al. (1996), mau → sim

Vigneault and Zouros 1986; Heikkinen and Lumme 1991; Zeng and Singh 1993b; Orr and Irving 2001). Other interactions such as X-Y interaction (Orr 1987), Y-autosome interaction (Pantazidiz, Galanopoulos, and Zouros 1993) also have been suggested. (3)Although a couple of studies have suggested the involvement of major sterility genes (Perez, 1993; Zeng and Singh 1993), many more results are compatible with the hypothesis that male sterility is due to complicated interactions of a large number of genes, each with partial or even no effect (Naveira 1992; Cabot et al. 1994; Palopoli and Wu 1994); (4) Most studies so far have focused on genes that reduce the hybrid fertility when placed in genetic background of other species. Little attention has been paid to the corresponding component of another species. The only example is the study on the introgression of the Y chromosome of *D. arizonae* that resulted in male sterility in a *D. mojavensis* genetic background (Pantazidis and Zouros 1988). The hybrid male fertility was restored when fourth chromosome from *D. arizonae* was cointrogressed. By introgression and mapping experiments, Pantazidis, Galanopoulos and Zouros (1993) suggested that this rescuing effect behaved as a single Mendelian factor, Sperm Motility Factor (SMF). The sterility interaction was interpreted as the interspecific incompatibility between the Y chromosome from *D. arizonae* and an automosomal SMF from *D. mojavensis*. Increased resolution mapping is required to rule out the possibility of this factor to be tightly linked gene complex, though. (5) X chromosome was showed to have a disproportional great effect on male sterility (Dobzhansky 1936; Coyne 1984; Orr 1987; Orr 1989a; Coyne, Rux and David 1991).

While our understanding of the genetic basis of hybrid sterility has progressed greatly, details of the genetic structure of hybrid sterility still remain rudimentary. The recombination mapping study seems to be in a dilemma to provide further information. At the current stage, we have not reached the consensus unequivocally concerning the underlying architecture of hybrid sterility. Until critical transformation experiments which require the full identification of tightly linked conspecific factors are performed, all sterility genes mentioned above remain only candidate steriles.

REFERENCE

Adams, M. D. et al. 2000. The genome sequence of *Drosophila melanogaster*. Science 287: 2185-2195.

Aguadro, C. F., K. M. Lado, and W. A. Noon. 1988. The *rosy* region *of Drosophila melanogaster* and *Drosophila simulans*. I. Contrasting levels of naturally occurring DNA restriction-map variation and divergence. Genetics 119: 875-888.

Begun, D. J., and C. F. Aquadro. 1991. Molecular population genetics of the distal portion of the X chromosome in *Drosophila*: Evidence for genetic hitchhiking of the *yellow-achaete* region. Genetics 129: 1147-1158.

Begun, D. J., and P. Whitley. 2000. Reduced X-linked nucleotide polymorphism in *Drosophila simulans*. Proc. Natl. Acad. Sci. USA 97: 5960-5965.

Betancourt, A. J., D. C. Presgraves, and W. J. Swanson. 2002. A test for faster X evolution in *Drosophila*. Mol. Biol. Evol. 19: 1816-1819.

Bock, I. R. 1984. Interspecific hybridization in the genus *Drosophila*. Evolutionary Biology 18: 41-70.

Brady, J. P., and R. C. Richmond. 1990. Molecular analysis of evolutionary changes in the expression of *Drosophila esterases*. Proc Natl Acad Sci U S A. 87: 8217-8221.

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., J. Coyne, and N. Barton. 1987. The relative rates of evolution of sex chromosomes and autosomes. Am. Nat. 130: 113-146.

Civetta, A., and R. S. Singh. 1995. High divergence of reproductive tract proteins and their association with postzygotic reproductive isolation in *Drosophila melanogaster* and *Drosophila virilis* group species. J Mol Evol. 41: 1085-1095.

Civetta, A., and R. S. Singh. 1998a. Sex and speciation: genetic architecture and evolutionary potential of sexual versus non-sexual traits in the sibling species of the Drosophila melanogaster complex. Evolution 52: 1080-1092.

Civetta, A., and R. S. Singh. 1998b. Sex-related genes, directional sexual selection, and speciation. Mol. Biol. Evol. 15: 901-909.

Civetta, A., and R. S. Singh. 1999. Broad-sense sexual selection, sex gene pool evolution, and speciation. Genome 42: 1033-1041.

Coulthart, M. B., and R. S. Singh. 1988. High level of divergence of male-reproductive-tract protein, between *Drosophila melanogaster* and its sibling species, *D. simulans*. Mol. Biol. Evol. 5: 182-191.

Coyne, J. A. 1984. Genetic basis of male sterility in hybrids between two closely related species of *Drosophila*. Proc. Natl. Acad. Sci. USA 81: 4444-4447.

Coyne, J. A. 1985. The genetic basis of Haldane's rule. Nature 314: 736-738.

Coyne, J. A. 1992. Genetics of sexual isolation in females of the *Drosophila simulans* species complex. Genet Res. 60: 25-31.

Coyne, J. A. and B. Charlesworth. 1986. Location of an X-linked factor causing sterility in male hybrids of *Drosophila simulans* and *D. mauritiana*. Heredity 57: 243-246.

Coyne, J. A., and B. Charlesworth. 1989. Genetic analysis of X-linked sterility in hybrids between three sibling species of *Drosophila*. Heredity 62: 97-106.

Coyne, J. A., and M. Kreitman. 1986. Evolutionary genetics of two sibling species, *Drosophila simulans* and *D. sechellia*. Evolution. 40: 673-691.

Coyne, J. A., J. Rux, and J. R. David. 1991. Genetics of morphological differences and hybrid sterility between *Drosophila sechellia* and its relatives. Genet Res 57: 113-22.

David, A. W. and C. I. Wu. 1996. The broom of the sorcerer's apprentice: the fine structure of a chromosomal region causing reproductive isolation between two sibling species of *Drosophila.* Genetics 143: 1287-1298.

Dobzhansky, T. 1936. Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. Genetics 21:113-135.

Dobzhansky, T. 1937. Genetics and the origin of species. New York: Columbia University Press. Pp364.

Eberhard, W. G. 1985. Sexual selection and animal genitalia. Harvard University Press, Cambridge, Mass.

Felsenstein, J. 1995. PHYLIP (Phylogeny Inference Package) version 3.57c. Distributed by the author at http://evolution.genetics.washington.edu/phylip.html Department of Genetics, University of Washington, Seattle.

Ferris, P. J., C. Pavlovic, S. Fabry, and U. W. Goodenough. 1997. Rapid evolution of sex-related genes in *Chlamydomonas*. Proc. Natl. Acad. Sci. USA 94: 8634-8639.

Galtier, N., and M. Gouy. 1995. Inferring phylogenies from DNA sequences of unequal base compositions. Proc Natl Acad Sci U S A. 92: 11317-11321.

Haldane, J. B. S. 1922. Sex ratio and unisexual sterility in animal hybrids. J. genet. 12: 101-109.

Harr, B., S. Weiss, J. R. David, G. Brem, and C. Schlotterer. 1998. A microsatellite-based multilocus phylogeny of the Drosophila melanogaster species complex. Curr. Biol. 8: 1182-1186.

Heikkinen, E. and J. Lumme. 1991. Sterility of male and female hybrids of *Drosophila virilis* and *Drosophila lummei*. Heredity 67: 1-11.

Hey, J., and R. M. Kliman. 1993. Population genetics and phylogenetics of DNA sequence variation at multiple loci within the *Drosophila melanogaster* species complex. Mol. Biol. Evol. 10: 804-822.

Hilton, H., R. M. Kliman, and J. Hey. 1994. Using hitchhiking genes to study adaptation and divergence during speciation within the *Drosophila melanogaster* species complex. Evolution. 48: 1900-1913.

Hudson, R. R. 1982. Estimating genetic variability with restriction endonucleases. Genetics 100: 711-719.

Hudson, R. R., M. Kreitman, and M. Aguade. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics 116: 153-159.

Johnson, N. A., H. Hollocher, E. Noonburg, and C. I. Wu. 1993. The effects of interspecific Y chromosome replacements on hybrid sterility within the *Drosophila simulans* clade. Genetics 135: 443-453.

Karr, T. L., and S. Pitnick. 1999. Sperm competition: defining the rules of engagement. Curr. Biol. 9: R787-790.

Kelly, J. K. and M. A. Noor. 1996. Speciation by reinforcement: a model derived from studies of *Drosophila*. Genetics 143: 1485-1497.

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge.

Kliman, R. M. and J. Hey. 1993. DNA sequence variation at the *period* locus within and among species of the *Drosophila melanogaster* complex. Genetics 133: 375-387.

Kliman, R. M. et al. 2000. The population genetics of the origin and divergence of the *Drosophila simulans* complex species. Genetics 156: 1913-1931.

Kresge, N., V. D. Vacquier, and C. D. Stout. 2001. Abalone lysin: the dissolving and evolving sperm protein. Bioessays 23: 95-103.

Kulathinal, R. J. 1996. Cytological characterization of hybrid male sterility among sibling species of the *Drosophila melanogaster* complex. Master Thesis. Biology Department, McMaster University.

Kulathinal, R. J. 2001. Molecular evolution of sex determination genes in closely related species of the *Drosophila melanogaster* subgroup. Ph. D. thesis. Biology Department, McMaster University.

Lachaise, D. L., M. L. Cariou, J. R. David, F. Lemeunier, and M. Ashburner. 1988. Historical biogeography of the D. melanogaster species subgroup. Evol. Biol. 22: 159-226.

Lachaise, D. L., J. R. David, F. Lemeunier, L. Tsacas, and M. Ashburner. 1986. The reproductive relationships of *Drosophila sechellia* with *D. mauritiana, D. simulans*, and *D. melanogaster* from the Afrotropical region. Evolution 40:262-271.

Lee, Y. H., T. Ota, and V. D. Vacquier, 1995. Positive selection is a general phenomenon in the evolution of abalone sperm lysin. Mol. Biol. Evol. 12: 231-238.

Lefevre, G. 1971. Salivary chromosome bands and the frequency of crossing over in *Drosophila melanogaster*. Genetics 67: 497-513.

Lindsley, D. L., and K. T. Tokuyasu. 1980. Spermatogenesis. In M. Ashburner and T. R. F. Wright, eds. The genetics and biology of *Drosophila*. Pp141-225. Academic Press, New York.

Makalowski, W. and M. S. Boguski. 1998. Evolutionary parameters of the transctibed mammalian genome: an analysis of 2,820 orthologous rodent and human sequences. Proc. Natl Acad. Sci. USA 95: 9407-9412.

Miller, G. T. and S. Pitnick. 2002. Sperm-female coevolution in Drosophila. Science. 298: 1230-1233.

Moriyama, E. N. and J. R. Powell. 1996. Intraspecific nuclear DNA variation in *Drosophila.* Mol. Biol. Evol. 13: 261-277.

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

Naveira, H. F. 1992. Location of X-linked polygenic effects causing sterility in male hybrids of *Drosophila simulans* and *D. mauritiana.* Heredity 68: 211-217.

Naveira, H., and A. Fontdevila. 1986. The evolutionary history of *Drosophila bussatii* XII. The genetic basis of sterility in hybrids between *D. buzzatii* and its sibling *D. serido* from Argentina. Genetics 114: 841-857.

Naveira, H., and A. Fontdevila. 1991. The evolutionary history of *Drosophila buzzatii* XXI. Cumulative action of multiple sterility factors on spermatogenesis in hybrids of *D. buzzatii* and *D. koepferae*. Heredity 67: 57-72.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Nei, M., and W. H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci U S A 76: 5269-5273.

Nei, M., and F. Tajima. 1981. DNA polymorphism detectable by restriction endonucleases. Genetics. 97: 145-163.

Nei, M., and J. Zhang. 1998. Molecular origin of species. Science 282: 1428-1429.

Nurminsky, D. I., N. V. Nurminskaya, D. De Aguiar, and D. L. Hartl. Selective sweep of a newly evolved sperm-specific gene in *Drosophila*. Nature 396: 572-575.

Noor, M. A. 1995. Incipient sexual isolation in *Drosophila pseudoobscuraa bogotana* Ayala & Dobzhansky. Pan-Pacif. Entomol. 71: 125-129.

Orr, H. A. 1987. Genetics of male and female sterility in hybrids of *Drosophila - pseudoobscura* and *D. persimilis*. Genetics 116: 555-563.

Orr, H. A. 1989a. Genetics of sterility in hybrids between two subspecies of *Drosophila*. Evolution 43: 180-189.

Orr, H. A. 1989b. Localization of genes causing postzygotic isolation in two hybridizations involving *Drosophila pseudoobscura*. Heredity 63: 231-237.

Orr, H. A. 1992. Mapping and characterization of a "speciation gene" in *Drosophila*. Genet. Res. 59: 73-80.

Orr. H. A. 1993a. A mathematical model of Haldane's rule. Evolution 47: 1606-1611.

Orr, H. A. 1993b. Haldane's rule has multiple genetic causes. Nature 361: 532-533.

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

Orr, H. A. 1997. Haldane's rule. Annu. Rev. Ecol. Syst. 28: 195-218.

Orr, H. A., and J. A. Coyne. 1989. The genetics of postzygotic isolation in the *Drosophila virilis* group. Genetics 121: 527-537.

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.

Palopoli, M. F., A. W. Davis, and C. I. Wu. 1996. Discord between the phylogenies inferred from molecular versus functional data: uneven rates of functional evolution or low levels of gene flow? Genetics. 144: 1321-1328.

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: 1-13.

Pantazidis, A. C., V. K. Galanopoulos, and E. Zouros. 1993. An autosomal factor from *Drosophila arizonae* restores normal spermatogenesis in *D. mojavensis* males carrying the *arizonae* Y chromosome. Genetics 134: 309-318.

Pantazidis, A. C., and E. Zouros. 1988. Location of an autosome factor causing sterility in *D. mojavensis* males carrying the *D. arizonensis* Y chromosome. Heredity 60: 299-304.

Perez, D. E., C. I. Wu, N. A. Johnson, and M. L. Wu. 1993. Genetics of reproductive isolation in the *Drosophila simulans* clade: DNA-marker assisted mapping and characterization of a hybrid male sterility gene, *Odysseus* (*Ods*). Genetics 134: 261-275.

Perez, D. E., and C. I. Wu. 1995. Further characterization of the *Odysseus* locus of hybrid sterility in *Drosophila*: one gene is not enough. Genetics 140: 201-206.

Powell J. R. 1997. Progress and prospects in evolutionary biology: The *Drosophila* model. Oxford University press, New York.

Presgraves, D. C. 2002. Patterns of postzygotic isolation in Lepidoptera. Evolution 56: 1168-1183.

Presgraves, D. C. and H. A. Orr. 1998. Haldane's rule in taxa lacking a hemizygous X. Science 282: 952-954.

Rieseberg, L. H., A. M. Desrochers, and S. J. Youn. 1995. Interspecific pollen competition as a reproductive barrier between sympatric species of *Helianthus* (Asteraceae). Am. J. Bot. 82: 515-519.

Sawamura, K., A. W. Davis, and C. I. Wu. 2000. Genetic analysis of speciation by means of introgression into *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA 97: 2654-2655.

Schaeffer, S. W., and E. L. Miller. 1991. Nucleotide sequence analysis of *Adh* gene estimates the time of geographic isolation of the Bogota population of *Drosophila pserdoobscura*. Proc. Natl. Acad. Sci. USA 88: 6097-6101.

Schafer, U. 1978. Sterility in *Drosophila hydei* × *Drosophila neohydei* hybrids. Genetica 49: 205-214.

Seeger, M. A. and T. C. Kaufman. 1990. Molecular analysis of the bicoid gene from *Drosophila pseudoobscura*: identification of conserved domains within coding and noncoding regions of the bicoid mRNA. EMBO J 9: 2977-2987.

Singh, R. S. 2000. Toward a unified theory of speciation. In Evolutionary genetics: From molecules to morphology. R.S. Singh and C.B. Krimbas (eds.), Cambridge University Press. Cambridge.

Singh, R. S. and R. J. Kulathinal. 2000. Proceedings of Fukuoka international symposium on population genetics: Sex gene pool evolution and speciation: a new paradigm. Genes Genet. Syst. 75: 119-130.

Solignac, M., M. Monnerot, and J. C. Mounolou. 1986. Mitochondrial DNA evolution in the *melanogaster* species subgroup of *Drosophila.* J. Mol. Evol. 23: 31-40.

Swanson, W. J., A. G. Clark, H. M. Waldrip-Dail, M. F. Wolfner, C. F. Aquadro. 2001a. Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila.* Proc. Natl. Acac. Sci. USA 98: 7375-7379.

Swanson, W. J., and V. D. Vacquier. 2002. The rapid evolution of reproductive proteins. Nature Reviews Genetics 3: 137-144.

Swanson, W. J., Z. Yang, M. F. Wolfner, and C. F. Aquadro. 2001b. Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. Proc. Natl. Acad. Sci. USA 98: 2509-2514.

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

Ting, C. T., S. C. Tsaur, and C. I. Wu. 2000. The phylogeny of closely related species as revealed by the genealogy of a speciation gene, *Odysseus*. Proc. Natl. Acad. Sci. USA 97: 5313-5316.

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

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.

Van Doorn G. S., P. C. Luttikhuizen, and F. J. Weissing. 2001. Sexual selection at the protein level drives the extraordinary divergence of sex-related genes during sympatric species. Proc. R. Soc. Lond. B 268: 2155-2161.

Vigneault, G. and E. Zouros. 1986. The genetics of asymmetrical male sterility in *Drosophila mojavensis* and *Drosophila arizonensis* hybrids: interactions between the Y-chromosome and autosomes. Evolution 40: 1160-1170.

Wang, P. J., J. R. McCarrey, F. Yang, and D. C. Page. 2001. An abundance of X-linked genes expressed in spermatogonia. Nat Genet. 27: 422-426.

Wright, S. 1982. The shifting balance theory and macroevolution. Annu Rev Genet. 16:1-19.

Wu, C. I. and A. T. Beckenbach. 1983. Evidence for extensive genetic differentiation between the sex-ratio and the standard arrangement of Drosophila pseudoobscura and the identification of hybrid sterility factors. Genetics 105: 71-86.

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. and M. F. Palopoli. 1994. Genetics of postmating reproductive isolation in animals. Annu. Rev. Genet. 27: 283-308.

Wyckoff, G. J., W. Wang, C. I. Wu. 2000. Rapid evolution of male reproductive genes in the descent of man. Nature 403: 304-309.

Vigneault, G. and E. Zouros. 1986. The genetics of asymmetrical male sterility in *Drosophila mojavensis* and *Drosophila arizonensis* hybrids: interactions between the Y-chromosome and autosomes. Evolution 40: 1160-1170.

Zeng, L. W., and R. S. Singh. 1993a. A combined classical genetic and high resolution two-dimensional electrophoretic approach to the assessment of the number of genes affecting hybrid male sterility in *Drosophila simulans* and *Drosophila sechellia*. Genetics 135: 135-147.

Zeng, L. W., and R. S. Singh. 1993b. The genetic basis of Haldane's rule and the nature of asymmetric hybrid male sterility among *Drosophila simulans*, *Drosophila mauritiana* and *Drosophila sechellia*. Genetics 134: 251-260.

Zeng, L. W., and R. S. Singh. 1995. A general method for identifying major hybrid male sterility genes in *Drosophila*. Heredity. 75: 331-341.

Appendix A: Sex and non-sex genes and their gene expression patterns/gene product functions in *D. melanogaster*

full name (*symbol*)	GenBank Accession No.	primers (5'-3')	CDS* (amplified segment)	gene expression patterns / gene product functions
sex genes				
outstretched (*os*)	AF231684	ccaggactcggccatagatat actggcttggatacctcgttg	2684-2833, 3023-3757 (2740-3642)	involved in primary sex determination
Dynein heavy chain at 16F (*Dhc16F*)	AE003507	ccttgtggaactatcgcatcg tttctgttctcattgcatccg	6736-6885, 6946-8185, 8250-8429 (6748-8422)	expressed in testis
Histone H3.3B (*His3.3B*)	X81205	aggctcctcgtaagcagctag tccttaggcatgattgtcacg	76-486 (119-446)	expressed in adult (testis), ovary (ovary) and other tissues
raspberry (*ras*)	AE003451	aggatgggctgagttgtaagg caccaatgacctgtagttcgg	1315-1478, 1571-1721, 1916-2567 (1423-2275)	expressed in oogenesis, adult (ovary), adult, S (testis)
ovarian tumor (*otu*)	X13693	gcctccagtttgttccgtgtg cgatgatagcggaatggcaac	155-2590 (266-1242)	expressed in adult, S (testis), oogenesis, adult (cystocyte, germarium, oocyte, ovary)
discs large 1 (*dlg1*)	M73529	agcgagaagaacctggagaac gttatagaaggagttggcggg	381-3263 (1020-1664)	expressed in adult (ejaculatory bulb, male accessory gland), adult, S (testis), oogenesis

Stellate (*Ste*)	X15899	gtgcaggatacgttcaaccag cagcagcgagaagaagatgtc	455-955 (537-923)	involved in spermatogenesis
Clathrin heavy chain (*Chc*)	Z14133	gcgggagaaggtgaatgatac acggtgttgatgtagggaatg	288-5324 (407-1309)	mutants affect primary spermatocyte cyst, spermatid, spermatozoon
Hira	AF031081	gtggtcatctggaatctgctg cccagaccgttcacatagttc	99-3284 (216-1252)	expressed in oogenesis
Gastrulation-defective (*gd*)	AF056311	aagcccaccacgaactatctg tgggagcctttactaccttgg	63-1649 (263-1328)	expressed in oogenesis
meiotic 41 (*mei-41*)	U34925	tggactgctggagaccctaag gctgctagaacaagattggcc	2318-4267 (2749-3936)	expressed in oogenesis, adult (ovary)
scully (*scu*)	Y15102	aaggagctgggcgacaaggtg ccggatctagggcatcatgcg	1049-1807 (1175-1813)	expressed in gonad and other tissues
ariadne (*ari-1*)	X98309	ccctgttggcacgagtatctg ttggcctcgtcctcgtcatag	3230-3879 (3269-3834)	expressed in gonad and other tissues
pacman (*pcm*)	AJ242529	acatgaacggcattgtccaca tcgatgttcttgaggtcgctg	112-4953 (215-1538)	expressed in ovary
fused (*fu*)	X80468	tatacaaggcgacacgcaagg tcagcttttcctcattggtgg	904-1014, 1087-2456 (956-2449)	mutants affect blastoderm, ocellus, ovary and others
Topoisomerase1 (*Top1*)	M74557	caagagcagcagttcgtcgtc tcgtcgtcatcgtcgttatcc	259-3177 (513-1325)	expressed in adult (ovary)

Helicase (*Hlc*)	AF017777	cggtttatcgttttcctcagc cagtcccgtaaggtgatcgag	55726-56372 (55809-56255)	expressed in adult (ovary) and other tissues
beta-Spectrin (*beta-Spec*)	M92288	atagtcgatggcaatgcttcc tctttgggtgtgtagggcttc	129-7004 (528-1270)	expressed in oogenesis and other tissues
enhancer of yellow 1 (*e(y)1*)	AE003506	agatcagtgcccaaatcaagc tgatgtcgtttcgctagttgg	103-253, 318-1003 (131-1006)	mutations affect the ovary and the oocyte and are recessive female sterile
non-sex genes				
Cyclic-AMP response element binding protein B (*CrebB-17A*)	AE003508	atggacaacagcatcgtcgag atccttggtagccacttctgg	1-384 (1-384)	RNA polymerase II transcription factor, expressed in adult (adult head)
Beadex (*Bx*)	AJ010387	ggacattaccaaaaccgaacc ctcaagtagtcccgcttgcag	276-1217 (287-718)	mutations affect the dorsal compartment wing
Cytochrome P450-18a1 (*Cyp18a1*)	AE003509	cttgtgggcgagtgagtggtg aactgtggaaggatcagcgcc	1759-1880, 1958-2387, 2491-2755 (1780-2755)	cytochrome P450
Protein kinase-like 17E (*bin4*)	AF096866	ccaacatcaaacgccacctac ttgtagatggcaatcctttgg	102-1565 (1174-1423)	protein kinase, mutants affect anterior fascicle
gamma-tubulin ring protein 84 (*Grip84*)	AF118379	agccgcgattatcttctgcac gtacgcactgctggtgttcgt	282-2741 (284-1088)	structural protein of cytoskeleton, mutants are recessive lethal
Tim9b	AF150104	ctgtgtttcagccgatgcgtg ccgcctctttcagtcgctcct	185-280, 343-579 (224-520)	protein translocase
Zwischenferment (*Zw*)	AE003512	tggaaccgcgtgattatcgag gccacttgtaggagccggagt	3877-4954 (3881-4922)	glucose-6-phosphate 1-dehydrogenase

Merlin (*Mer*)	U49724	tcgttttgcaggtgctccatg gggctgcgcgagtcttggtac	3451-4107, 4167-4376, 4474-4710 (3583-4685)	cytoskeletal protein binding
Cdc42	U11824	atgcaaaccatcaagtgcgtg ccttttctttgtgggctctgg	875-1162, 1224-1511 (875-1493)	RHO small monomeric GTPase, expressed in embryo (somatic mesoderm et al.)
CG9569	AE002611	ctacttgccgctctacgagga gcaggagtcggttaggttgag	1-1117 (312-1101)	G-protein coupled receptor
CG9571	AE002611	gaagcacaactcaggaagcag gatatgaaggggctttgattg	1-783 (84-738)	transcription factor
swiss cheese (*sws*)	Z97187	ctttattgatgtcctgtcggg tccagtccaaggcatagtcac	490-4767 (1230-2359)	expressed in adult (adult head)
sevenless (*sev*)	J03158	aggctaatgtgagtgctgtgc ataggcagtggacaatcttcg	9420-14080 (11564-12782)	protein tyrosine kinase, expressed in larva (imaginal disc), adult (adult head)
brinker (*brk*)	AB023583	aactactcccacagcaatcgg acctccacatcctcgttctcc	533-2647 (891-1854)	expressed in embryo (ectoderm, embryonic/larval midgut, endoderm)
carnation (*car*)	AF133260	gttattgtgctggacgagacc gtgcatctggacgtgaatttc	64-1917 (171-1071)	mutants affect pigment cell
minibrain (*mnb*)	X70794	cttacgaccacgaggagcagt ctcgtattggttgcctcgtct	2031-3650 (2376-3319)	expressed in embryo (supraoesophageal ganglion, ventral nerve cord)
rutabaga (*rut*)	AE003497	ccacgttacacttctccccac ttgtggtgactgctgttgttg	33478-34658 (33514-34508)	expressed in larva (corpora pedunculata)
Beta Adaptin (*Bap*)	X75910	accaccaagaagggcgagatc ttgtcaatgcgctctgcgtac	91-2856 (115-1361)	product involved in non-selective vesicle coating

amnesiac (*amn*)	AE003513	tttttatccggctgctgtggc gagtgctcttgggttctcgcg	2569-3111 (2589-3081)	neuropeptide hormone, signal transducer, mutations affect the Kenyon cell
Neprilysin 3 (*Nep3*)	AE002611	caattatctcgtctggcaggc gagtgtgcgtccgttcatctt	4495-5324 (4504-5305)	endothelin-converting enzyme, metallopeptidase

*: CDS shown may not include entire CDS of the gene. Only those that were amplified were shown.

Appendix B: Digestion data for sex genes ("0" and "1" represent the absence and presence of the restriction site, respectively.)

Sex genes	d ± S.D.						enzymes (r)	sites*	mel	mau	sim	sec
	mel-mau	mel-sim	mel-sec	mau-sec	mau-sim	sim-sec						
Hlc	0.02986 ± 0.01846	0.01892 ± 0.01389	0.00890 ± 0.00907	0.01892 ± 0.01389	0.00958 ± 0.00980	0.00882 ± 0.00897	AluI(AG/CT) (4)	372	1	1	1	1
								422	1	1	1	1
							CfoI(GCGC) (4)	28	1	1	1	1
								115	1	1	1	1
							MboI(/GATC) (4)	45	1	1	1	1
								225	1	1	1	1
								264	1	1	1	1
								430	1	1	1	1
							MseI(T/TAA) (4)	77	1	1	1	1
								179	1	0	0	1
							MspI(C/CGG) (4)	63	1	1	1	1
								87	1	1	1	1
							PvuII(CAG/CTG) (6)	421	0	1	1	1
							XhoI(C/TCGAG) (6)	70	1	0	1	1
ras	0.08468 ± 0.0400	0.03935 ± 0.0246	0.03935 ± 0.0247	0.01403 ± 0.01445	0.01403 ± 0.01445	0	AccI(GT/MKAC) (5.33)	789	1	1	1	1
							AluI(AG/CT) (4)	415	1	1	1	1
								676	1	1	1	1
							BclI(T/GATCA) (6)	805	1	1	1	1
							DdeI(C/TNAG) (4)	9	1	1	1	1
								108	1	1	1	1
							HaeII(RGCGC/Y) (5.33)	203	1	0	0	0
								521	1	0	0	0
							MseI(T/TAA) (4)	277	0	1	0	0
								300	1	0	0	0
								552	0	1	0	0
								817	1	0	1	1
Dhc16F	0.00686 ± 0.00699	0.02144 ± 0.01311	0.00686 ± 0.00701	0	0.01343 ± 0.00986	0.01343 ± 0.00987	AccI(GT/MKAC) (5.33)	865	1	1	1	1
							BclI(T/GATCA) (6)	1042	1	1	1	1
							DdeI(C/TNAG) (4)	1410	1	1	1	1
							HinfI(G/ANTC) (4)	800	0	1	1	1
								1414	1	1	1	1
							PvuII(CAG/CTG) (6)	44	1	1	1	1
								405	1	1	1	1
							EcoRV(GAT/ATC) (6)	54	1	1	1	1
								1467	0	0	1	0
								1567	1	1	0	1
							MseI(T/TAA) (4)	523	1	1	1	1
								970	1	1	1	1
								1486	1	1	1	1

Sex genes	d ± S.D.						enzymes (r)	sites	mel	mau	sim	sec
	mel-mau	mel-sim	mel-sec	mau-sec	mau-sim	sim-sec						
pcm	0.04357 ± 0.01826	0.03802 ± 0.01701	0.03802 ± 0.01702	0.03111 ± 0.01262	0.03111 ± 0.01263	0	AccI(GT/MKAC) (5.33)	800	0	1	1	1
								1290	1	1	1	1
							BamHI(G/GATCC) (6)	824	0	1	1	1
								1024	1	1	1	1
							BclI(T/GATCA) (6)	869	1	1	1	1
								1271	1	1	1	1
							DdeI(C/TNAG) (4)	105	0	0	0	0
								709	1	1	1	1
								994	1	1	1	1
								1274	0	1	0	0
							EcoRI(G/AATTC) (6)	1083	1	0	0	0
							HaeII(RGCGC/Y) (5.33)	1118	1	0	0	0
								1259	1	1	1	1
							HinfI(G/ANTC) (4)	361	1	1	1	1
								376	1	1	1	1
							XhoI(C/TCGAG) (6)	650	0	1	1	1
								729	1	1	1	1
							PvuII(CAG/CTG) (6)	250	0	1	1	1
							PstI(CTGCA/G) (6)	96	1	1	1	1
								1068	1	1	1	1
e(y)1	0.01549 ± 0.01128	0.02301 ± 0.01385	0.02454 ± 0.01481	0.00835 ± 0.00848	0.00782 ± 0.00794	0.01653 ± 0.01205	AluI(AG/CT) (4)	64	1	1	1	1
								298	1	1	1	1
								448	0	1	0	0
								634	1	0	1	1
							BclI(T/GATCA) (6)	43	1	1	1	1
							HaeII(RGCGC/Y) (5.33)	523	1	1	1	1
								688	1	1	1	1
							HinclI(GTY/RAC) (5.33)	96	1	1	1	1
								427	1	1	1	1
							MseI(T/TAA) (4)	135	1	1	1	1
								374	1	1	1	0
								509	1	0	0	0
								671	1	0	0	0
							Sau96I(G/GNCC) (4)	307	1	1	1	1
								443	1	1	1	1
								486	1	1	1	1
								532	1	1	1	1
								832	0	0	1	0

Sex genes	d ± S.D.						enzymes (r)	sites	mel	mau	sim	sec
	mel-mau	mel-sim	mel-sec	mau-sec	mau-sim	sim-sec						
ari	0	0	0	0	0	0	AccI(GT/MKAC) (5.33)	354	1	1	1	1
							PvuII(CAG/CTG) (6)	378	1	1	1	1
								411	1	1	1	1
							Sau96I(G/GNCC) (4)	515	1	1	1	1
							BclI(T/GATCA) (6)	407	1	1	1	1
							EcoRV(GAT/ATC) (6)	480	1	1	1	1
							HinfI(G/ANTC) (4)	202	1	1	1	1
							MspI(C/CGG) (4)	251	1	1	1	1
							AluI(AG/CT) (4)	379	1	1	1	1
								398	1	1	1	1
								412	1	1	1	1
fu	0.05996 ± 0.03485	0.04666 ± 0.03041	0.04666 ± 0.03042	0.01527 ± 0.01576	0.01527 ± 0.01577	0	AccI(GT/MKAC) (5.33)	1326	1	1	1	1
							BclI(T/GATCA) (6)	652	1	1	1	1
								1111	1	1	1	1
							HincII (GTY/RAC) (5.33)	1326	1	1	1	1
							MseI (T/TAA) (4)	759	0	1	0	0
								859	1	0	0	0
								1259	0	1	1	1
								1456	1	0	0	0
							PvuII(CAG/CTG) (6)	759	1	1	1	1
							KpnI(GGTAC/C) (6)	477	1	0	0	0
His3.3B	0.00695 ± 0.00704	0.00695 ± 0.00705	0.00695 ± 0.00706	0	0	0	AccI(GT/MKAC) (5.33)	128	1	1	1	1
							AluI(AG/CT) (4)	16	1	1	1	1
								151	1	1	1	1
							CfoI(GCGC) (4)	45	1	1	1	1
								289	1	1	1	1
								305	1	1	1	1
							DdeI(C/TNAG) (4)	184	1	1	1	1
								247	1	0	0	0
								322	1	1	1	1
							HaeII(RGCGC/Y) (5.33)	44	1	1	1	1
							HincII(GTY/RAC) (5.33)	128	1	1	1	1
							MboI(/GATC) (4)	110	1	1	1	1
								201	1	1	1	1
							MspI(C/CGG) (4)	55	1	1	1	1
								88	1	1	1	1

Sex genes	d ± S.D.						enzymes (r)	sites	mel	mau	sim	sec
	mel-mau	mel-sim	mel-sec	mau-sec	mau-sim	sim-sec						
os	0.07638 ± 0.03962	0	0.04121 ± 0.02613	0.05456 ± 0.03038	0.01278 ± 0.01319	0.04205 ± 0.02647	EcoRV(GAT/ATC) (6)	250	0	0	0	1
								882	1	1	1	1
							HaeII(RGCGC/Y) (5.33)	554	1	1	1	1
								605	1	1	1	1
							HinclI(GTY/RAC) (5.33)	438	1	1	1	0
								698	0	1	0	0
								750	0	0	0	1
							XhoI(C/TCGAG) (6)	330	1	1	1	1
							KpnI(GGTAC/C) (6)	599	1	1	1	1
							MspI(C/CGG) (4)	286	1	1	1	1
Top1	0.00736 ± 0.00747	0.00736 ± 0.00748	0	0.00736 ± 0.00750	0	0.00736 ± 0.00752	AccI(GT/MKAC) (5.33)	536	1	1	1	1
							BclI(T/GATCA) (6)	688	1	1	1	1
							CfoI(GCGC) (4)	46	1	1	1	1
								250	1	1	1	1
								754	1	1	1	1
							HinfI(G/ANTC) (4)	527	1	1	1	1
							MseI(T/TAA) (4)	372	1	1	1	1
								477	1	1	1	1
								564	1	1	1	1
							MspI(C/CGG) (4)	377	1	1	1	1
								465	1	1	1	1
								498	1	1	1	1
							PvuII(CAG/CTG) (6)	395	1	1	1	1
							Sau96I(G/GNCC) (4)	139	1	1	1	1
								153	1	1	1	1
								653	0	1	1	0
beta-Spectra	0.00835 ± 0.00848	0.00835 ± 0.00849	0	0.00835 ± 0.00849	0	0.00835 ± 0.00849	AluI(AG/CT) (4)	335	1	1	1	1
								483	0	1	1	0
							CfoI(GCGC) (4)	54	1	1	1	1
								174	1	1	1	1
								195	1	1	1	1
							MboI(/GATC) (4)	63	1	1	1	1
								526	1	1	1	1
								540	1	1	1	1
							MseI(T/TAA) (4)	212	1	1	1	1
							PstI(CTGCA/G) (6)	694	1	1	1	1
							XhoI(C/TCGAG) (6)	673	1	1	1	1

Sex genes	d ± S.D.						enzymes (r)	sites	mel	mau	sim	sec
	mel-mau	mel-sim	mel-sec	mau-sec	mau-sim	sim-sec						
otu	0.04822 ± 0.03103	0.01565 ± 0.01612	0.01502 ± 0.01556	0.04889 ± 0.03100	0.0343 ± 0.02602	0.04889 ± 0.03101	AccI(GT/MKAC) (5.33)	877	1	1	1	1
							CfoI(GCGC) (4)	194	1	1	1	1
							MboI(/GATC) (4)	466	1	1	1	1
								544	1	0	0	0
							MspI(C/CGG) (4)	680	1	1	1	1
							PvuII(CAG/CTG) (6)	650	0	1	0	0
								751	1	0	1	1
							RsaI(GT/AC) (4)	390	1	1	1	1
							TaqI(T/CGA) (4)	134	1	1	1	1
Chc	0.05651 ± 0.02843	0.05651 ± 0.02844	0.05651 ± 0.02845	0	0	0	AccI(GT/MKAC) (5.33)	750	0	1	1	1
								866	1	0	0	0
							BglII(A/GATCT) (6)	147	1	1	1	1
								756	1	1	1	1
							CfoI(GCGC) (4)	182	1	1	1	1
								532	0	1	1	1
							MseI(T/TAA) (4)	228	1	1	1	1
								492	1	0	0	0
							ClaI(AT/CGAT) (6)	497	1	1	1	1
							DdeI(C/TNAG) (4)	79	1	1	1	1
							HinclI(GTY/RAC) (5.33)	750	1	0	0	0
								866	0	1	1	1
							HinfI(G/ANTC) (4)	766	1	1	1	1
							RsaI(GT/AC) (4)	543	1	1	1	1
Hira	0.01343 ± 0.00975	0.01343 ± 0.00976	0.00677 ± 0.00686	0.00642 ± 0.0065	0	0.00642 ± 0.0065	CfoI(GCGC) (4)	174	1	1	1	1
								639	1	1	1	1
							DdeI(C/TNAG) (4)	626	1	1	1	1
							HinclI(GTY/RAC) (5.33)	100	1	1	1	1
							HinfI(G/ANTC) (4)	225	1	1	1	1
								732	0	1	1	0
							KpnI(GGTAC/C) (6)	205	1	1	1	1
							MboI(/GATC) (4)	424	1	1	1	1
								790	1	1	1	1
							MseI(T/TAA) (4)	407	1	1	1	1
							MspI(C/CGG) (4)	209	1	1	1	1
								309	0	1	1	1
								571	1	1	1	1
							PstI(CTGCA/G) (6)	321	1	1	1	1
								820	1	1	1	1
							PvuII(CAG/CTG) (6)	318	1	1	1	1
							RsaI(GT/AC) (4)	206	1	1	1	1
								539	1	1	1	1

Sex genes	d ± S.D.						enzymes (r)	sites	mel	mau	sim	sec
	mel-mau	mel-sim	mel-sec	mau-sec	mau-sim	sim-sec						
dlg1	0	0	0	0	0	0	AccI(GT/MKAC) (5.33)	243	1	1	1	1
							BclI(T/GATCA) (6)	77	1	1	1	1
							CfoI(GCGC) (4)	253	1	1	1	1
								452	1	1	1	1
							ClaI(AT/CGAT) (6)	54	1	1	1	1
							HaeII(RGCGC/Y) (5.33)	451	1	1	1	1
							HincII(GTY/RAC) (5.33)	243	1	1	1	1
							HinfI(G/ANTC) (4)	283	1	1	1	1
							MboI(/GATC) (4)	57	1	1	1	1
							MseI(T/TAA) (4)	518	1	1	1	1
							MspI(C/CGG) (4)	116	1	1	1	1
gd	0.05615 ± 0.02769	0.04686 ± 0.02537	0.04686 ± 0.02538	0.00999 ± 0.01023	0.00999 ± 0.01024	0	CfoI(GCGC) (4)	337	1	1	1	1
								696	1	1	1	1
							ClaI(AT/CGAT) (6)	555	1	0	1	1
							DdeI(C/TNAG) (4)	820	1	1	1	1
							EcoRV(GAT/ATC) (6)	600	0	1	0	0
								838	1	1	1	1
							MseI(T/TAA) (4)	500	1	0	0	0
								600	0	1	1	1
								809	1	1	1	1
							MspI(C/CGG) (4)	223	1	1	1	1
								1015	1	1	1	1
							PvuII(CAG/CTG) (6)	817	1	1	1	1
							RsaI(GT/AC) (4)	250	0	1	1	1
								296	1	0	0	0
								927	1	1	1	1
mei	0.02776 ± 0.02229	0.02776 ± 0.02230	0.04437 ± 0.02969	0.04437 ± 0.02970	0.02776 ± 0.02233	0.01349 ± 0.01481	ClaI(AT/CGAT) (6)	341	1	1	1	1
							HaeII(RGCGC/Y) (5.33)	139	1	1	1	0
							MboI(/GATC) (4)	854	1	1	1	1
							MspI(C/CGG) (4)	514	1	0	0	0
								740	1	1	1	1
								940	0	1	0	0
								990	0	0	1	1
							PstI(CTGCA/G) (6)	371	1	1	1	1
							TaqI(T/CGA) (4)	969	1	1	1	1
							XhoI(C/TCGAG) (6)	284	1	1	1	1

Sex genes	d ± S.D.						enzymes (r)	sites	mel	mau	sim	sec
	mel-mau	mel-sim	mel-sec	mau-sec	mau-sim	sim-sec						
Ste	0.1116 ± 0.05129	0.1116 ± 0.05130	0.1116 ± 0.05131	0	0	0	CfoI(GCGC) (4)	254	0	1	1	1
							ClaI(AT/CGAT) (6)	73	1	1	1	1
							PvuII(CAG/CTG) (6)	331	0	1	1	1
							TaqI(T/CGA) (4)	74	1	1	1	1
							RsaI(GT/AC) (4)	117	1	1	1	1
								187	0	1	1	1
							DdeI(C/TNAG) (4)	250	1	0	0	0
							HaeII(RGCGC/Y) (5.33)	253	0	1	1	1
							HincII(GTY/RAC) (5.33)	270	1	0	0	0
							HinfI(G/ANTC) (4)	126	1	1	1	1
							MspI(C/CGG) (4)	67	1	0	0	0
								364	1	1	1	1
scu	0	0	0	0	0	0	AluI(AG/CT) (4)	5	1	1	1	1
							BclI(T/GATCA) (6)	197	1	1	1	1
							HaeII(RGCGC/Y) (5.33)	521	1	1	1	1
							HinfI(G/ANTC) (4)	573	1	1	1	1
							PstI(CTGCA/G) (6)	70	1	1	1	1

*: The sites indicate the positions of the restriction sites on the amplified segments. Restriction sites for *D. melanogaster* are known (albeit some differences in my case) as the genome has already been sequenced. The determination of sites for the other species were made based on the known sites of *D. melanogaster*.

Appendix C: Digestion data for non-sex genes ("0" and "1" represent the absence and presence of the restriction site, respectively.)

nonsex genes	d ± S.D.						enzymes (r)	sites*	mel	mau	sim	sec
	mel-mau	mel-sim	mel-sec	mau-sec	mau-sim	sim-sec						
Grip84	0.04137 ± 0.02271	0.04917 ± 0.02480	0.04917 ± 0.02481	0.00709 ± 0.00815	0.02221 ± 0.01514	0.01408 ± 0.01154	AluI(AG/CT) (4)	194	1	0	0	0
								250	0	1	1	1
								535	0	1	0	1
								600	0	0	1	0
								668	1	0	0	0
							HinfI(G/ANTC) (4)	607	1	1	1	1
								857	0	0	1	1
							Sau96I(G/GNCC) (4)	469	1	1	1	1
								563	1	1	1	1
								571	1	1	1	1
								821	0	1	1	1
							MspI(C/CGG) (4)	632	1	1	1	1
								754	1	1	1	1
CreB-17A	0.01176 ± 0.01504	0.01176 ± 0.01505	0.02564 ± 0.02404	0.01299 ± 0.01716	0	0.01299 ± 0.01716	BamHI(G/GATCC) (6)	355	1	1	1	1
							HinclI(GTY/RAC) (5.33)	329	1	1	1	1
							MboI(/GATC) (4)	356	1	1	1	1
							MspI(C/CGG) (4)	179	1	1	1	0
								316	1	0	0	0
							PstI(CTGCA/G) (6)	295	1	1	1	1
							RsaI(GT/AC) (4)	170	1	1	1	1
							Sau96I(G/GNCC) (4)	184	1	1	1	1
								313	1	1	1	1
Bap	0.01980 ± 0.01297	0.01259 ± 0.01014	0.01259 ± 0.01015	0.01980 ± 0.01297	0.01980 ± 0.01298	0	AccI(GT/MKAC) (5.33)	181	1	1	1	1
							BamHI(G/GATCC) (6)	375	1	1	1	1
							BglII(A/GATCT) (6)	1344	1	1	1	1
							DdeI(C/TNAG) (4)	526	1	1	1	1
								858	1	1	1	1
								1042	1	1	1	1
								1137	1	1	1	1
							HinclI(GTY/RAC) (5.33)	139	1	1	1	1
								1186	1	0	1	1
							MseI(T/TAA) (4)	239	1	1	1	1
								476	0	1	1	1
								536	1	0	0	0
								590	1	1	1	1
								1124	1	1	1	1
							TaqI(T/CGA) (4)	23	1	1	1	1
								490	1	1	1	1
								779	1	1	1	1
							XbaI(T/CTAGA) (6)	778	1	1	1	1
							Sau96I(G/GNCC) (4)	447	1	1	1	1
								462	1	1	1	1
								630	1	1	1	1
								882	1	1	1	1

nonsex genes	d ± S.D.						enzymes	sites	mel	mau	sim	sec
	mel-mau	mel-sim	mel-sec	mau-sec	mau-sim	sim-sec						
amn	0	0	0	0	0	0	DdeI(C/TNAG) (4)	146	1	1	1	1
								222	1	1	1	1
							HinfI(G/ANTC) (4)	340	1	1	1	1
							MspI(C/CGG) (4)	8	1	1	1	1
							PvuII(CAG/CTG) (6)	143	1	1	1	1
							Sau96I(G/GNCC) (4)	316	1	1	1	1
								350	1	1	1	1
							TaqI(T/CGA) (4)	101	1	1	1	1
								138	1	1	1	1
Zw	0.01308 ± 0.01065	0.01308 ± 0.01066	0.01948 ± 0.01305	0	0.00671 ± 0.00765	0.00671 ± 0.00765	AluI(AG/CT) (4)	95	1	1	1	1
								175	1	1	1	1
								343	1	1	1	1
								680	1	1	1	1
								770	1	1	1	1
								884	1	1	1	1
							BclI(T/GATCA) (6)	212	1	1	1	1
							ClaI(AT/CGAT) (6)	106	1	1	1	1
							DdeI(C/TNAG) (4)	58	1	1	1	1
								172	1	1	1	1
							HinfI(G/ANTC) (4)	505	1	1	1	1
								802	1	1	1	1
							MseI(T/TAA) (4)	100	0	0	0	1
								719	1	0	0	0
							PstI(CTGCA/G) (6)	304	1	0	0	0
							PvuII(CAG/CTG) (6)	94	1	1	1	1
Cdc42	0	0	0	0	0	0	AccI(GT/MKAC) (5,33)	65	1	1	1	1
							AluI(AG/CT) (4)	443	1	1	1	1
								491	1	1	1	1
								503	1	1	1	1
							HinfI(G/ANTC) (4)	329	1	1	1	1
							MboI(/GATC) (4)	134	1	1	1	1
								192	1	1	1	1
							MspI(C/CGG) (4)	175	1	1	1	1
							Sau96I(G/GNCC) (4)	158	1	1	1	1
								549	1	1	1	1
							XhoI(C/TCGAG) (6)	589	1	1	1	1
Bx	0	0	0	0	0	0	HincII(GTY/RAC) (5,33)	155	1	1	1	1
							Sau96I(G/GNCC) (4)	116	1	1	1	1
								117	1	1	1	1
								289	1	1	1	1
							MspI(C/CGG) (4)	21	1	1	1	1
								87	1	1	1	1
								162	1	1	1	1
								171	1	1	1	1
								267	1	1	1	1

nonsex genes	d ± S.D.						enzymes	sites	mel	mau	sim	sec
	mel-mau	mel-sim	mel-sec	mau-sec	mau-sim	sim-sec						
Mer	0	0	0	0	0	0	AccI(GT/MKAC) (5.33)	1062	1	1	1	1
							DdeI(C/TNAG) (4)	618	1	1	1	1
							HincII(GTY/RAC) (5.33)	1062	1	1	1	1
							HinfI(G/ANTC) (4)	209	1	1	1	1
								543	1	1	1	1
								1090	1	1	1	1
							MspI(C/CGG) (4)	760	1	1	1	1
								848	1	1	1	1
								947	1	1	1	1
							PstI(CTGCA/G) (6)	891	1	1	1	1
								1071	1	1	1	1
							TaqI(T/CGA) (4)	94	1	1	1	1
								351	1	1	1	1
								395	1	1	1	1
								1063	1	1	1	1
CG9565	0.05544 ± 0.03323	0.05544 ± 0.03323	0.00919 ± 0.01108	0.04139 ± 0.02678	0	0.04139 ± 0.02678	BclI(T/GATCA) (6)	667	1	0	0	1
							BglII(A/GATCT) (6)	144	1	1	1	1
								614	0	1	1	0
							HincII(GTY/RAC) (5.33)	518	1	1	1	1
							MspI(C/CGG) (4)	58	1	1	1	1
								209	1	1	1	1
								479	1	1	1	1
							MseI(T/TAA) (4)	162	0	0	0	1
								612	1	0	0	1
								738	1	1	1	1
							PstI(CTGCA/G) (6)	572	1	1	1	1
CG9569	0.00953 ± 0.01156	0.00962 ± 0.01176	0	0.00953 ± 0.01156	0.02064 ± 0.01833	0.02084 ± 0.01843	AluI(AG/CT) (4)	518	1	1	1	1
								534	1	1	1	1
							HinfI(G/ANTC) (4)	5	1	1	1	1
							MseI(T/TAA) (4)	182	1	1	1	1
							MboI(/GATC) (4)	221	1	0	1	1
								460	1	1	1	1
								730	1	1	1	1
							PvuII(CAG/CTG) (6)	517	1	1	1	1
							XhoI(C/TCGAG) (6)	79	1	1	0	1
CG9571	0.02509 ± 0.02339	0.02509 ± 0.02340	0.02509 ± 0.02341	0	0	0	DdeI(C/TNAG) (4)	407	1	1	1	1
								494	1	1	1	1
								642	1	1	1	1
							HinfI(G/ANTC) (4)	405	1	0	0	0
								500	0	0	0	0
								590	1	1	1	1
							XhoI(C/TCGAG) (6)	617	1	1	1	1
Tim9	0	0	0	0	0	0	Sau96I(G/GNCC) (4)	175	1	1	1	1
							DdeI(C/TNAG) (4)	266	1	1	1	1

nonsex genes	d ± S.D.						enzymes	sites	mel	mau	sim	sec
	mel-mau	mel-sim	mel-sec	mau-sec	mau-sim	sim-sec						
mnb	0	0	0	0	0	0	ClaI(AT/CGAT) (6)	498	1	1	0	0
							DdeI(C/TNAG) (4)	300	1	1	1	1
							HincII(GTY/RAC) (5.33)	390	1	1	1	1
							HinfI(G/ANTC) (4)	698	1	1	1	1
							MseI(T/TAA) (4)	43	1	1	1	1
								139	1	1	1	1
							PstI(CTGCA/G) (6)	894	1	1	1	1
							PvuII(CAG/CTG) (6)	276	1	1	1	1
							RsaI(GT/AC) (4)	209	1	1	1	1
								716	1	1	1	1
							XhoI(C/TCGAG) (6)	93	1	1	1	1
car	0	0	0	0	0	0	AccI(GT/MKAC) (5.33)	340	1	1	1	1
							BglII(A/GATCT) (6)	467	1	1	1	1
							CfoI(GCGC) (4)	157	1	1	1	1
								490	1	1	1	1
							ClaI(AT/CGAT) (6)	351	1	1	1	1
							DdeI(C/TNAG) (4)	604	1	1	1	1
							HincII(GTY/RAC) (5.33)	340	1	1	1	1
							HinfI(G/ANTC) (4)	537	1	1	1	1
								743	1	1	1	1
							PstI(CTGCA/G) (6)	439	1	1	1	1
							PvuII(CAG/CTG) (6)	169	1	1	1	1
								436	1	1	1	1
rut	0.01753 ± 0.01816	0.01753 ± 0.01817	0.01753 ± 0.01818	0	0	0	AccI(GT/MKAC) (5.33)	501	1	1	1	1
							CfoI(GCGC) (4)	627	1	1	1	1
								877	0	1	1	1
							HincII(GTY/RAC) (5.33)	501	1	1	1	1
								561	1	1	1	1
							HinfI(G/ANTC) (4)	200	0	0	0	0
								292	1	1	1	1
								634	1	1	1	1
							MseI(T/TAA) (4)	298	1	1	1	1
							PvuII(CAG/CTG) (6)	397	1	1	1	1
								819	1	1	1	1
							RsaI(GT/AC) (4)	792	1	1	1	1
sws	0.02291 ± 0.01692	0.02291 ± 0.01693	0.02095 ± 0.01542	0.02291 ± 0.01695	0	0.02291 ± 0.01697	AccI(GT/MKAC) (5.33)	627	1	1	1	1
								967	1	1	1	1
							BclI(T/GATCA) (6)	324	1	1	1	1
							BglII(A/GATCT) (6)	529	1	1	1	1
							DdeI(C/TNAG) (4)	679	1	0	0	0
							MseI(T/TAA) (4)	270	1	0	0	1
							MspI(C/CGG) (4)	160	1	1	1	1
								464	1	1	1	1
							RsaI(GT/AC) (4)	173	1	1	1	1
								573	0	0	0	1
								1079	1	1	1	1
							TaqI(T/CGA) (4)	99	1	1	1	1

nonsex genes	d ± S.D.						enzymes	sites	mel	mau	sim	sec
	mel-mau	mel-sim	mel-sec	mau-sec	mau-sim	sim-sec						
Cyp18	0	0	0	0	0	0	AccI(GT/MKAC) (5.33)	442	1	1	1	1
								592	1	1	1	1
							BclI(T/GATCA) (6)	242	1	1	1	1
							HincII(GTY/RAC) (5.33)	442	1	1	1	1
								592	1	1	1	1
							MspI(C/CGG) (4)	223	1	1	1	1
								652	1	1	1	1
								679	1	1	1	1
							PstI(CTGCA/G) (6)	87	1	1	1	1
								698	1	1	1	1
brk	0	0	0.0099 ± 0.01016	0.0099 ± 0.01017	0	0.0099 ± 0.01019	BclI(T/GATCA) (6)	916	1	1	1	1
							DdeI(C/TNAG) (4)	57	1	1	1	1
								457	0	0	0	0
								651	1	1	1	1
							HinfI(G/ANTC) (4)	740	1	1	1	1
							MspI(C/CGG) (4)	291	1	1	1	1
							PstI(CTGCA/G) (6)	160	1	1	1	1
							PvuII(CAG/CTG) (6)	54	1	1	1	1
								164	0	0	0	1
							TaqI(T/CGA) (4)	441	1	1	1	1
sev	0.01414 ± 0.0146	0.01414 ± 0.0147	0.01414 ± 0.0148	0	0	0	AccI(GT/MKAC) (5.33)	519	1	0	0	0
								861	1	1	1	1
								1028	1	1	1	1
							BclI(T/GATCA) (6)	1186	1	1	1	1
							DdeI(C/TNAG) (4)	98	1	1	1	1
								148	0	0	0	0
								806	1	1	1	1
							HincII(GTY/RAC) (5.33)	408	1	1	1	1
							TaqI(T/CGA) (4)	382	1	1	1	1
Bin4	0.03437 ± 0.03663	0.03437 ± 0.03664	0.03437 ± 0.03665	0	0	0	DdeI(C/TNAG) (4)	79	1	0	0	0
							MspI(C/CGG) (4)	45	1	1	1	1
								132	1	1	1	1
							PstI(CTGCA/G) (6)	96	1	1	1	1

*: The sites indicate the positions of the restriction sites on the amplified segments. Restriction sites for *D. melanogaster* are known (albeit some differences in my case) as the genome has already been sequenced. The determination of sites for the other species were made based on the known sites of *D. melanogaster*.