Sex-specific selection and sexual antagonism in the fruit fly, *Drosophila melanogaster*

Jack D. Hesketh
PhD, Evolutionary Biology
University College London

Declaration

I, Jack D. Hesketh, confirm that the work present in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Males and females differ in their reproductive roles, and as a consequence each sex is subject to divergent selection pressures to optimise its own reproductive success. Due to the shared genome between males and females these selection pressures frequently act on shared phenotypic traits. Divergent selection can favour the invasion of sexually antagonistic alleles which increase the fitness of one sex at the detriment of the other. Sexual antagonism can be subsequently resolved through the evolution of sex-specific gene expression, allowing the sexes to diverge phenotypically. While sexual dimorphism is common, recent evidence shows that antagonistic genetic variation continues to segregate in populations of many organisms. The basis of sexual antagonism remains poorly understood.

I first present empirical data on the interaction between sexual antagonism and genetic drift in small populations that had independently evolved under standardised conditions. I demonstrated that these experimental populations of *Drosophila melanogaster* had diverged in male and female fitness, with some populations showing increased male but decreased female fitness, while other populations showed the reverse pattern.

I also exploited a sample of nine genomes that belonged to three fitness classes (low male/high female, high male/low female, intermediate in both sexes) to test the association between the sexually dimorphic trait wing morphology (size and shape) and fitness in both sexes. I found that wing morphology significantly affected the fitness of both males and females, but to a differing degree in each sex. In males wing shape rather than wing size was especially important. I found evidence that there was appropriate genetic architecture for the existence of sexual antagonism, and for divergent selection on aspects of wing morphology. I place all of my findings in the context of variation in sexually dimorphic traits and sex-specific fitness.

Acknowledgements

This thesis would not have been possible were it not for numerous people who made both professional and personal contributions, and I would like to thank them here:

Max Reuter and Kevin Fowler, for their supervision, guidance, patience, and for getting up in the middle of the night to help toss on flies.

All the members of the Reuter Laboratory for their support: Charles Mullon, Sara Fuentes Perez, and Julie Collet.

Members of other laboratories that we work alongside: Aaron Towlson, and Rebecca Finlay.

My parents for their general support throughout.

My housemates and good friends for their kindness.

My cat, Pluto

Plus, too many others to mention.

1

General introduction

1.1 Sexual selection and sex-specific selection

Since Darwin (1874) first started methodically recording the phenotypes of numerous species, it has been apparent that males and females frequently have different life histories, primarily owing to the differences in the types of selective pressure operating on mating and reproduction. Bateman (1948) conducted seminal work that has had a huge impact on the understanding of the reproductive pressures for males and females. Using the *D. melanogaster* insect model, Bateman (1948) found that male reproductive success varied substantially more than that of females. For example, his results showed that only 4% of females failed to reproduce at all, whereas the corresponding figure for males was 21%. Furthermore, some males were reproductively much more successful than females, and produced almost three times as many offspring as the most successful females. In summary, Bateman (1948) demonstrated that male reproductive success increased proportionally with the number of matings a male could acquire. Conversely, when females access to multiple mating partners such that they were not sperm limited, then their reproductive success was mostly limited by the ability to produce eggs and did not increase beyond one mating.

Bateman went on to hypothesize that the imbalance between male and female reproductive potential could be explained by differences in the energetic investment of each sex in their gametes. Males typically invest relatively little energy in the production of each sperm, and as such their reproductive potential is not limited by the production of gametes but by their ability to acquire mates with eggs to fertilize. Conversely, females invest substantial energy in the production of each of their eggs, and therefore a female's reproductive success is primarily determined by her ability to acquire resources and convert them into eggs. Trivers (1972) extended the general principle of unequal investment between males and females and argued that the asymmetry in investment in gametes leads to a more fundamental divergence in sex roles. For females, where fitness is limited by gamete production and offspring care, the high investment into offspring favors selection for characteristics that make this investment most efficient. For males, in contrast, where fitness is mostly limited by success in competition with other males, selection is dominated by pressures favouring traits that facilitate the acquisition of matings and success in fertilisation. This means that the evolution of the female phenotype will to a large degree be shaped by fecundity selection, whereas male evolution tends to be dominated by sexual selection.

Fecundity selection favors traits that will increase a female's potential reproductive capacity. A typical example of such a trait is size. In the overwhelming majority of sexually reproducing species females benefit from larger size (Anderson 1994). Large size allows for increased egg production and storage capacity. In species such as mammals, where females not only need energy for egg production but also for gestation and lactation, the energy demands are further increased by maternal care. As such, experimental evidence suggests that size correlates positively with fecundity in numerous mammalian species (e.g. Boyce and Boyce 1988; Purvis and Harvey 1995).

Sexual selection in males will favour traits that increase the number of successful matings and fertilisations that a male can achieve. Darwin (1874), who first described sexual selection as an evolutionary force in its own right, distinguished two different modes of selection, intrasexual selection, and intersexual selection. Intrasexual selection describes competition between males, often through direct encounters, to increase their reproductive success. Here, males often benefit from the evolution weapons to aid direct encounters with other males. For example, in the rhinoceros beetle, *T. dichotomus,* one of the best known horned beetles, mating success in males has been shown to increase with horn size (Emlen 2008). Intersexual selection describes interaction between the sexes whereby a female will typically choose a male to copulate with. Here, males frequently benefit from elaborate behavioral display and elaborate morphological traits to increase their chances of being chosen by females (e.g. Prum 1990). As a consequence of these selective pressures on males, selection for reproduction in males is credited with the evolution of some of nature’s most extravagant structures.

The sex-specific pressures of fecundity and sexual selection are complemented by those arising from natural selection. In most cases, natural selection will act on both sexes equally. Differences in the survival strategy of males and females are rare, especially for higher animals (Darwin 1874). Unless each of the sexes occupies a distinct ecological niche, many of the fundamental pressures for an organism to survive will be identical between the sexes. Importantly, natural selection often opposes the sex-specific pressures of fecundity and sexual selection. This is very obvious for some of the exaggerated sexual ornaments often found in males, that frequently compromise the viability of their bearers (Anderson 1994).

1.2 Sexual dimorphism

The sex-specific selection pressures described above often exert divergent selection on traits that are shared between males and females. As a result, we expect that male and female phenotypes should reflect the sex-specific balance between different forces acting in members of that sex, notably that between sex-specific selection pressures and natural selection. The results of this sexually divergent selection are visually evident in many species, where we observe sexually dimorphic phenotypes. A typical example comes from the red deer (*Cervus elaphus*). The male stags require large antlers to fight and/or intimidate other males, in order to maximize their access to females and hence reproductive success. Conversely, the female hinds gain no advantage from carrying around this large weapon, and so possess significantly smaller antlers. Sexual dimorphism encompasses any phenotypic differences, which exist between the sexes of a given species. These differences are often more subtle than in our deer example, but this demonstrates the evolutionary result of adaptation to selection pressures for each sex to fulfill its own lifetime role.

One of the most common forms of sexual dimorphism is sexual size dimorphism (SSD), whereby one sex is larger than the other. As with all forms of sexual dimorphism, SSD is largely the result of males and females adapting to their own sex-specific roles (e.g. Greenwood and Adams 1987; Anderson 1994; Fairbairn 1997). SSD varies substantially between different species in both direction and magnitude (Blanckenhorn et al. 2006). Assuming similar natural selection against increased size in both sexes, the direction and extent of SSD reflects the relative intensity of fecundity and sexual selection. Females should be larger where the intensity of fecundity selection for larger size in females exceeds the intensity of sexual selection for larger size in male, and males should be larger where the reverse is true. The differences in SSD between mammalian species and insect species serves as a good example of how this balance of selective forces affects the direction of SSD. In mammals, SSD is typically male-biased (Alexander 1979; Weckerly 1998). This pattern can be explained by sexual selection on males (Darwin 1874), and is increasingly pronounced for mammalian species where males provide minimal parental care. Where this is the case, a male's reproductive success can directly increase by competing for matings against other males (Trivers 1972) and large size can provide a substantial advantage in this intrasexual competition. Numerous studies support the idea that intense intrasexual competition among males is the principal driver of male-biased SSD in mammals. Several comparative studies have

shown that the smaller the proportion of males that mate in a given population, then the greater the level of male biased SSD. In other words, where mating systems are more polygynous then more extreme SSD is more observed. Correlations between the level of polygyny and the scale of SSD have been demonstrated in mammals in general (Alexander 1979), but also specifically for primates (Cluttonbrock and Harvey 1977), ungulates (Geist 1974), and pinnipeds (Lindenfors et al. 2002). Although the bulk of empirical evidence suggests that male-male competition is the principal driver of male-biased SSD in mammals (Isaac et al. 2005), it is likely that other selective pressures are involved too. In fact some studies have shown that, in rare cases, female fecundity can decrease with increased body size (Boyce and Boyce 1988). Comparative studies and theoretical models (e.g.Charnov 1993) predict that because growing to a large size requires time and energy, there is a trade-off where reproductive success is constrained beyond a certain size, as it is not energy efficient to be extremely large.

The male-biased SSD driven by sexual selection in mammals contrasts with female-biased SSD in many insects. Here the larger size of females occurs despite the persistence of strong sexual selection among males and is the result of intense fecundity selection in female size. Unlike in mammals, the number of offspring that a female insect can produce (so long as she is not sperm limited) is directly proportional to the number of eggs that she can produce, and this is highly correlated with body size (Knight and Robertson 1957). Moreover, female fecundity is much more closely related to body size than are the major components of male reproductive success, such as mating ability. An important comparative study by Honek (1993) examined literature published for 57 different insect species. It showed that the increase in fecundity with body weight was similar for most taxa. Honek 1993 recorded a 0.95% increase in median fecundity, for every 1% increase in dry body mass. Furthermore, the number of ovarioles increased at 0.81% for every 1% increase in body mass, providing a direct connection between the rate of egg production and female body size.

Although males are almost always the smaller of the two sexes in insect male-male competition generally favors larger size for many of the same proximate reasons as in mammals. Specifically, larger males tend to experience greater success in male-male competition or female choice (Anderson 1994), and can further increase the fecundity of mating partners via larger nuptial gifts (Stillwell and Fox 2007). The field cricket (*Gryllus bimaculatus*) provides a neat example of large male advantage for an insect.

This species transfers spermatophores, a capsule that contains spermatozoa, to inseminate females. Simmons (1988) demonstrated that large males could generate spermatophores at relatively low cost compared with smaller males. Females would more readily mate with larger males, relative to smaller males. Larger males were more sexually competent, in that they were more successful in attaching spermatophores once mounted.

However, there are exceptions to the general pattern of a male size advantage in insect species. Selection is not always straightforward, and numerous ecological factors may generate selection towards reduced male size. A general theory for a small-male advantage, applicable to many invertebrate species, is the 'Ghiselin-Reiss small-male hypothesis' (Ghiselin 1974). It suggests that smaller males can have a reproductive advantage in mating systems that are dominated by scramble competition, whereby access to females is shared equally among males. In species where scramble competition is prevalent male reproductive success is often largely determined by their encounter rate with females. Under these conditions smaller males are expected to be able to dedicate the maximum possible time to finding females, as such males feed less frequently owing to lower energy requirements. Support for this type of small-male advantage comes from an example in the water strider (*A. remigis*), a scramble competitor that mates multiple times. Blanckenhorn et al. (1995) used a series of controlled feeding experiments, in which they tested the mating success of both sexes. Mating success and mating duration increased with the amount of available food for males and females. However, when food was limited, male body size was negatively correlated with both the number of mating attempts and mating success, providing direct empirical support for a small-male advantage as postulated by the Ghiselin-Reiss hypothesis.

1.3 Sexual antagonism

Above I have described the fundamental selection that generates sexual dimorphism. I have argued that the balance of selective forces between the sexes determines both the magnitude and direction of sexual selection. Although sexual dimorphism is both widespread and often extreme in its magnitude, research suggests that it cannot freely evolve, and as a result both males and females are frequently displaced from their optimal phenotype. This is elegantly illustrated by a study of Chippindale et al. (2001) who investigated male and female fitness variation in a laboratory population of *D.*

melanogaster. A random sample of forty haploid genomes was expressed in both males and females. That is one copy of chromosomes X, II and III, paired with a random genetic complement. The fitness of each genome, in each sex, was measured at both the larval and adult life stages. At the larval stage, survival was highly positively correlated between the sexes. This suggests that genomes vary in their performance in fitness, but because the sex roles at the larval stage are identical, genomes that make high quality male larvae also make high quality female larvae. In adult flies, however, where the roles of the sexes diverge to optimise the reproductive success of each sex, Chippindale et al. (2001) observed a significant negative correlation between male and female fitness across the forty genomes. Thus, genomes that have a high fitness when expressed in males tend to perform badly in females, and vice versa.

Chippindale et al.'s results are consistent with the presence of genetically determined phenotypic variation, where some genomes express more masculinised features in both sexes whereas others express more feminised features in both sexes. A more masculine genome will be advantageous when expressed in a male, but deleterious when expressed in a female whereas a more feminine genome will have the opposite fitness effects. These so-called "sexually antagonistic" fitness effects occur because the phenotypic effects of these genotypes are not sex-specific, implying that the traits underlying these effects are not dimorphic enough to maximise both male and female fitness. The existence of sexually antagonistic variation for fitness is not restricted to laboratory populations of fruitflies but has been detected across a range of different taxa (e.g. Fedorka and Mousseau 2004; Brommer et al. 2007; Foerster et al. 2007) implying that antagonism is a widespread phenomenon. Theoretically, sexual antagonism (SA) can be resolved by the evolution of differential gene expression. This will break up previously existing genetic correlations and allow each sex to evolve towards its own phenotypic optimum and lead to the sexual dimorphism that we observe in most sexual species. Hence, SA is a built-in evolutionary conflict that precedes the evolution of all sexual dimorphism that originates from new mutations.

1.4 Dynamics of SA alleles

At the level of alleles, sexually antagonistic selection means that selection in one sex favors the fixation of one allele, whilst selection in the other sex favors fixation of another allele. Population genetic models have been used to understand the conditions for which a sexually antagonistic mutation can invade a population (Rice 1984).

Fundamentally, a new SA mutation will invade when the fitness benefit that it confers to one sex outweighs the fitness cost that it confers to the other sex, meaning that its net fitness benefit is positive across the sexes. These same genetic models have also been used to explain the conditions under which SA selection can lead to stable polymorphism. The dynamics of this can be explained with a simple verbal model (Rice and Chippindale 2001) Take for example a male-benefit allele, which for simplicity we will assume is co-dominant. When rare, such an allele will most commonly be expressed in the heterozygous state, but at this low frequency it confers a greater advantage to males than disadvantage to females, and will consequentially increase in frequency. Over time, this causes the allele to be expressed more commonly in the homozygous state. However, increasingly frequent homozygous expression of the allele can subsequently reduce the allele frequency in two ways; 1) by increasing the negative dosage females receive, or 2) by overshooting the optimum expression levels for males. The effect of both scenarios can make the female disadvantage increase, relative to male advantage, such that overall selection switches against the allele, resulting in a frequency reduction, and it becoming more commonly expressed in the polymorphic state. This type of persistent polymorphism is a likely byproduct of sexually antagonistic selection. Gavrilets and Rice (2006) have since extended the basic premise of these simple selection dynamics using a Wright-Fisher model to generate deterministic mathematical models, which predict the rate at which SA alleles will be lost or fixed in a given population. They predict that the range of parameters that allow for stable polymorphism of sexually antagonistic alleles is large, if the cost:benefit ratio of expression between the sexes is relatively high, thereby preventing fixation.

Evidence from both theoretical (Rice 1984; Gavrilets and Rice 2006) and empirical (Gibson et al. 2002) studies also suggests that the X chromosome (in systems of XY chromosomal sex determination) should be enriched with polymorphisms that have sexually antagonistic fitness effects. This prediction is based on the observation that invasion conditions on the X are less stringent than those for the autosomes when male-beneficial antagonistic mutations are recessive and female beneficial mutations dominant (but see Fry 2010). We can extend the illustration used above to explain this effect. If we first consider a rare recessive mutation with a small benefit to males, and large cost to females. The average selection on an autosomal allele of this type would not be favored by selection. However, if X-linked this allele can theoretically increase in frequency, as it is commonly expressed in males, whilst being masked in females

owing to it being recessive. As the allele accumulates, and it begins to get expressed in the homozygous state in females, counter-selection in females will ensure the spread of the allele is stopped. Alternatively, we can also consider a dominant mutation that benefits females and costs males. If X-linked, this allele can increase in frequency when rare because it is more often expressed in diploid females than in hemizygous males. The general effect of these two scenarios theoretically promotes the establishment of antagonistic variation for loci on the X chromosome over a broader range of cost:benefit ratios than would be possible for autosomal loci (Rice 1984). This theoretical prediction is also backed up by empirical data collected using the *D. melanogaster* model system, which suggests that 97% of all genome-wide sexually antagonistic fitness variation is located on the X-chromosome (Gibson et al. 2002).

To date, the overwhelming majority of theory on sexual antagonism has used deterministic population genetic models and hence ignored the effects of genetic drift. However, a recent study by Connallon and Clark (2012) demonstrated that genetic drift can have important consequences for the level of antagonistic polymorphism that is observed in wild populations. Specifically, they show that the effective population size, and genetic drift substantially affect the statistical frequency distributions of alleles that are subject to antagonistic selection. This means that the random sampling of alleles causes gene frequencies to fluctuate, and causes loss of genetic variation by speeding up the fixation or loss of alleles. As a result, the effects of genetic drift should theoretically oppose the balancing selection that can be caused by sexually antagonistic fitness effects. One of the key implications of these findings is that the standing genetic variation in a given population will depend on the relative intensity of both sexually antagonistic selection, and its interplay with genetic drift. In Chapter 2 of this thesis I describe an empirical investigation on the effects of genetic drift in sexually antagonistic fitness variation using the *D. melanogaster* model species. I present empirical data on the interaction between sexual antagonism and genetic drift in populations that have independently evolved under standardised conditions. I demonstrate that small experimental populations of *Drosophila melanogaster* have diverged in male and female fitness, with some populations showing increased male, but decreased female fitness, while other populations show the reverse pattern. The fitness divergence between populations is consistent with the differentiation in reproductive fitness being driven by genetic drift in sexually antagonistic alleles. I discuss the implications of my results with respect to the maintenance of antagonistic

variation in subdivided populations and consider the wider implications of drift in fitness-related genes.

1.5 Traits that contribute towards SA fitness variation

In chapters 3 and 4, I turn my attention to understanding the morphological traits that contribute to both sex-specific fitness and sexual antagonism in *D. melanogaster*. In order to get a full understanding of antagonistic evolution, it is important to understand SA variation at the level of the genotype through to that of the phenotype in its fitness effects. So far, such an integrated understanding of antagonism has not been possible. Most studies that have investigated the SA fitness effects of specific traits have used species for which reliable fitness measures are difficult to obtain (Cox and Calsbeek 2009). Other studies, usually under laboratory conditions, have been able to obtain a good understanding of SA variation for fitness, but have not extensively investigated the phenotypic basis of this these effects (Chippindale et al. 2001; Rowe and Day 2006; Innocenti and Morrow 2010). To really advance our understanding of SA we need to bridge this information gap, with studies that can measure both fitness and trait variation in an accurate and repeatable way. Experiments that have managed to repeatedly detect SA fitness effects have used classic fruitfly genetic techniques to manipulate entire haploid genomes (a complete set of chromosomes X, II, and III) and multiply them identically. These haploid genomes can then be expressed in both male and female genetic backgrounds, and assayed for their lifetime fitness effects in each sex. In one particular study this system of genetic manipulation has been used to successfully measure both sex-specific fitness, and its association with a specific trait (Long and Rice 2007). This experiment measured the fitness effects associated with the sexually dimorphic trait of locomotion in both male and female *D. melanogaster*. It was established that the rate of adult locomotion was both highly correlated between the sexes, and had opposite fitness effects in males and females. Male locomotion correlated positively with fitness, whereas females suffered a fitness cost with increased rates of locomotion. This demonstrated that loci which determine the rate of locomotion are among those that underlie negative genetic correlations for fitness.

Here I chose to investigate the SA fitness effects associated with the size and shape of *D. melanogaster* wings. *D. melanogaster* serves as an excellent model for the study of wing morphology. In particular the shape and size of *D. melanogaster* wings are easily measured by using the wing vein intersections as landmarks from which to compare

samples using geometric morphometrics (Mardia 1998). As is typical in flies, *D. melanogaster* shows pronounced female biased SSD. There is a broad consensus that females are the larger sex due to selection for increased fecundity (Knight and Robertson 1957). Large female size is related to increased fecundity (Bouletreaumerle et al. 1982). The consensus view of the literature is that male mating success, i.e. the specific ability of acquiring matings, is positively affected by increased size. This size related mating success is largely determined by components of male courtship, many of which are associated with variation in male size (Partridge & Farquhar 1983).

The general pattern of sexual shape dimorphism (SShD) is highly conserved across the *D. melanogaster* subgroup (Gidaszewski et al. 2009), with male wings being relatively shorter and broader than female wings (Gilchrist et al. 2000; Gidaszewski et al. 2009; Abbott et al. 2010). In studying variation of traits which are intrinsically connected in their morphology, such as wing size and shape, it is important to consider allometric effects. Allometry describes the relationship with body size of various other life history components, including shape, anatomy, physiology, and behavior. However, it is most used in the study of shape variation with changes in size. Allometry has been identified by numerous studies as playing an important role in sexual shape dimorphism (SShD) (Gould 1966; Klingenberg 1996). In order to directly test how much of the variation in SShD is due to allometric effects studies have separated the effects of allometric (shape variation correlated with size), and non-allometric variation (shape variation not correlated with size). Gidaszewski (2009) tested the wing size and shape of nine different species from the *D. melanogaster* subgroup to estimate the average proportion of variation in wing shape that was due to allometric variation with wing size at approximately 50%. This provided support for the general idea that factors other than allometry are likely to contribute to SShD of *D. melanogaster* wings.

For wing shape to be a sexually selectable trait with measurable fitness variation, it was also important that it was not subject to any severe developmental constraints. Experimental evidence suggested that the wings of *D. melanogaster* are free of any absolute developmental constraints. Mezey and Houle (2005) directly estimated the dimensionality of genetic variation in wing shape, using 20 different aspects of wing shape to test for bidirectional constraints on the evolution of *D. melanogaster* wings. Their results showed significant additive genetic variation for the vast majority of wing shape aspects that they measured. They concluded that no *D. melanogaster* wing shape

is beyond the reach of evolution. Further to this lack of developmental constraint, experiments that have applied artificial selection on *D. melanogaster* wing shape consistently show a strong evolutionary response to selection (Weber 1990; Houle et al. 2003), suggesting that there is extensive standing variation for wing shape, and therefore scope for selection to generate sexually divergent phenotypes. In combination these results suggested that selection is likely to be an important architect of *D. melanogaster* wing shape.

One of the principal reasons for choosing to measure SA fitness effects of wing morphology was the existence of sexual dimorphism for both wing size and wing shape. This suggests that selection on overall wing morphology is now, or was in the past, subject to opposing selection pressures in males and females. Two recent studies suggest that *D. melanogaster* wings are indeed still subject to sexually antagonistic selection between the sexes. Both Prasad et al. (2007) and Abbott et al. (2010), used *D. melanogaster* selection lines that had been maintained exclusively in males, for 25 and 70 generations respectively. This male-limited evolution meant that these selection lines were subject only to directional selection towards male-specific optimal phenotypes and could evolve free from interference due to potentially opposing selection in females. As a result they evolved masculinised male wing morphology that was consistent with the direction of sexual dimorphism. These results provided direct evidence that male wing shape was subject to divergent selection between the sexes, such that under normal evolutionary conditions males are displaced from their optimal wing size and shape. In studies that follow in this thesis, I demonstrate that the fitness of both male and female *D. melanogaster* is affected by wing size and shape. My evidence suggests that wing shape is more important for determining male fitness, whereas wing size is a more important in determining female fitness. I studied the fitness effects of male wings in detail, in the expectation that male morphology was under more intense sexual selection than that of females, and should therefore provide measurable morphological variation. I found evidence that appropriate genetic architecture existed for the sexes to be genetically restricted in their evolution of sex-specific wing morphology and that wings were likely subject to sexually divergent selection.

1.6 Thesis structure

My PhD project was funded by the Biotechnology and Biological Sciences Research Council (UK) and was performed under the supervision of Dr Max Reuter and Professor Kevin Fowler. All experiments were conducted by the author, with assistance from other members of the Reuter laboratory on those occasions where the logistical scale of the experiments made this necessary.

1.6.1 Chapter 2

In this first empirical chapter of this thesis I present findings from a large scale fitness assay of more than 70 genomes sampled from 4 independently evolving selection lines. Each selection line had been reared for more than 80 generations at a low (n = 100) population size. I tested the larval and adult fitness of each sex under assay conditions that matched their rearing regime in their selection line of origin. Using this fitness data I tested for the interaction between sexual antagonism and genetic drift in these small independently evolving populations. I demonstrate that small experimental populations of *D. melanogaster* have diverged in male and female fitness, with some populations showing increased male, but decreased female fitness, while other populations show the reverse pattern. A revised version ('major revision') of the material presented in this chapter is now published in the journal *Evolution*.

1.6.2 Chapter 3

Here I took a trait-focused approach to the study of male fitness and exploited a sample of genomes that conferred a distinct pattern of SA fitness variation. Of the nine genomes that I sampled, three conferred low male/high female fitness, three conferred high male/low female fitness, and three conferred an intermediate level of fitness in both sexes. I monitored the fitness effects of these genomes when expressed in males. I measured male wing morphology (size and shape), and male fitness (estimated via a measure of mating success), for males expressing these genomes after being raised in each of three different larval density environments. I demonstrated that male fitness varies significantly with wing morphology, specifically showing that wing shape is frequently more important for male fitness than wing size.

1.6.3 Chapter 4

Here, I extended my analysis of the fitness effects associated with male wing morphology by collecting a synonymous data set for females expressing the same set of nine genomes as used in chapter 3. That is, I measured female fitness, wing size, and wing shape for the same nine genomes across three equivalent larval density treatments. I also added a third component, development time, which was tested across both sexes. Using the new data on females in combination with the data acquired in chapter 3 on males, I investigated the relationship of fitness, wing morphology, and development time between the sexes. I specifically tested for evidence of SA selection on wing size and shape. Plus I tested for evidence that males may be subject to selection for decreased development rate owing to selection for a more accurate morphology. Here I provided evidence that the genetic architecture, in the form of intersexual genetic correlations, exists to restrict the sexually divergent evolution of wing morphology. Furthermore I provide evidence that suggests wing morphology is indeed subject to divergent selection pressures in males and females.

1.6.4 Chapter 5

Here I discuss how my findings across the three empirical chapters relate to each other. Also I outline the direction of future experiments and describe recent experiments undertaken in our laboratory to further investigate wing morphology between the sexes of *D. melanogaster*.

1.6.5 Appendix 1

This describes a series of preliminary experiments designed to inform the experimental designs used in chapters 3 and 4. Specifically, I tested factors relating to wing morphology, larval development time, and larval mortality.

1.6.6 Appendix 2

This consists of a draft manuscript of a study to which I contributed substantially at the stages of experimental design and data collection. The study documents a change in the genetic correlation between male and female fitness in a replicate of the LHm population and hence a possible case of resolving sexual antagonism.

1.7 References

Abbott JK, Bedhomme S, Chippindale AK (2010) Sexual conflict in wing size and shape in *Drosophila melanogaster*. J Evolution Biol 23(9): 1989-1997.

Alexander RD, J. L. Hoogland, R. D. Howard, K. M. Noonan & P. W. Sherman. (1979) Sexual dimorphism and breeding systems in pinnipeds, ungulates, primates, and humans. In: Evolutionary Biology and Human Social Behavior: An Anthropological Perspective (N. A. Chagnon & W. G. Irons eds.). 402-435. Duxbury Press, North Scituate.

Anderson M (1994) Sexual selection: Princeton University Press.

Bateman AJ (1948) Intra-Sexual Selection in *Drosophila*. Heredity 2(3): 349-368.

Blanckenhorn WU, Preziosi RF, Fairbairn DJ (1995) Time and Energy Constraints and the Evolution of Sexual Size Dimorphism - to Eat or to Mate. Evol Ecol 9(4): 369-381.

Blanckenhorn WU, Stillwell RC, Young KA, Fox CW, Ashton KG (2006) When Rensch meets Bergmann: Does sexual size dimorphism change systematically with latitude? Evolution 60(10): 2004-2011.

Bouletreaumerle J, Allemand R, Cohet Y, David JR (1982) Reproductive Strategy in *Drosophila melanogaster* - Significance of a Genetic-Divergence between Temperate and Tropical Populations. Oecologia 53(3): 323-329.

Boyce CCK, Boyce JL (1988) Population Biology of Microtus-Arvalis .1. Lifetime Reproductive Success of Solitary and Grouped Breeding Females. J Anim Ecol 57(3): 711-722.

Brommer JE, Kirkpatrick M, Qvarnstrom A, Gustafsson L (2007) The Intersexual Genetic Correlation for Lifetime Fitness in the Wild and Its Implications for Sexual Selection. Plos One 2(8).

Charnov EL (1993) Is Maximum Sustainable-Yield Independent of Body Size for Mammals (and Others). Evol Ecol 7(3): 309-311.

Chippindale AK, Gibson JR, Rice WR (2001) Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. Proc Natl Acad Sci USA 98(4): 1671-1675.

Cluttonbrock TH, Harvey PH (1977) Primate Ecology and Social-Organization. J Zool 183(Sep): 1-39.

Connallon T, Singh ND, Clark AG (2012) Impact of Genetic Architecture on the Relative Rates of X versus Autosomal Adaptive Substitution. Mol Biol Evol 29(8): 1933-1942.

Cox RM, Calsbeek R (2009) Sexually Antagonistic Selection, Sexual Dimorphism, and the Resolution of Intralocus Sexual Conflict. Am Nat 173(2): 176-187.

Darwin C (1874) The Descent of Man and Selection in Relation to Sex (2nd edn.), London (1874); Murray J, editor.

Emlen DJ (2008) The Evolution of Animal Weapons. Annu Rev Ecol Evol S 39: 387-413.

Fairbairn DJ (1997) Allometry for sexual size dimorphism: Pattern and process in the coevolution of body size in males and females. Annu Rev Ecol Syst 28: 659-687.

Fedorka KM, Mousseau TA (2004) Female mating bias results in conflicting sex-specific offspring fitness. Nature 429(6987): 65-67.

Foerster K, Coulson T, Sheldon BC, Pemberton JM, Clutton-Brock TH et al. (2007) Sexually antagonistic genetic variation for fitness in red deer. Nature 447(7148): 1107-U1109.

Fry JD (2010) The Genomic Location of Sexually Antagonistic Variation: Some Cautionary Comments. Evolution 64(5): 1510-1516.

Gavrilets S, Rice WR (2006) Genetic models of homosexuality: generating testable predictions. P Roy Soc B-Biol Sci 273(1605): 3031-3038.

Geist V (1974) Relationship of Social Evolution and Ecology in Ungulates. Am Zool 14(1): 205-220.

Ghiselin M (1974) The Economy of Nature and the Evolution of Sex. University of California Press, 1974 - 346 pages.

Gibson JR, Chippindale AK, Rice WR (2002) The X chromosome is a hot spot for sexually antagonistic fitness variation. P Roy Soc B-Biol Sci 269(1490): 499-505.

Gidaszewski NA, Baylac M, Klingenberg CP (2009) Evolution of sexual dimorphism of wing shape in the *Drosophila melanogaster* subgroup. Bmc Evol Biol 9.

Gilchrist AS, Azevedo RBR, Partridge L, O'Higgins P (2000) Adaptation and constraint in the evolution of *Drosophila melanogaster* wing shape. Evol Dev 2(2): 114-124.

Gould SJ (1966) Allometry and Size in Ontogeny and Phylogeny. Biol Rev 41(4): 587-&.

Greenwood PJ, Adams J (1987) Sexual Selection, Size Dimorphism and a Fallacy. Oikos 48(1): 106-108.

Honek A (1993) Intraspecific Variation in Body Size and Fecundity in Insects - a General Relationship. Oikos 66(3): 483-492.

Houle D, Rowe L (2003) Natural selection in a bottle. Am Nat 161(1): 50-67.

Houle D, Mezey J, Galpern P, Carter A (2003) Automated measurement of *Drosophila* wings. Bmc Evol Biol 3.

Innocenti P, Morrow EH (2010) The Sexually Antagonistic Genes of *Drosophila melanogaster*. Plos Biol 8(3).

Isaac NJB, Jones KE, Gittleman JL, Purvis A (2005) Correlates of species richness in mammals: Body size, life history, and ecology. Am Nat 165(5): 600-607.

Klingenberg CP (1996) Multivariate allometry, Pages 23–49 in L. F. Marcus, M. Corti, A. Loy, G. J. P. Naylor, and D. E. Slice, eds. Advances in morphometrics. New York, Plenum Press.

Knight GR, Robertson A (1957) Fitness as a Measurable Character in *Drosophila*. Genetics 42(4): 524-530.

Lindenfors P, Tullberg BS, Biuw M (2002) Phylogenetic analyses of sexual selection and sexual size dimorphism in pinnipeds. Behav Ecol Sociobiol 52(3): 188-193.

Long TAF, Rice WR (2007) Adult locomotory activity mediates intralocus sexual conflict in a laboratory-adapted population of *Drosophila melanogaster*. Proc Roy Soc B-Biol Sci 274(1629): 3105-3112.

Mardia Da (1998) Statistical Shape Analysis. John Wiley, Chichester in July 1998.

Mezey JG, Houle D (2005) The dimensionality of genetic variation for wing shape in *Drosophila melanogaster*. Evolution 59(5): 1027-1038.

Partridge L, and Farquhar, M (1983) Lifetime mating success of male fruitflies (*Drosophila melanogaster*) is related to their size. Animal Behaviour 31, 871-877.

Prasad NG, Bedhomme S, Day T, Chippindale AK (2007) An evolutionary cost of separate genders revealed by male-limited evolution. Am Nat 169(1): 29-37.

Prum RO (1990) Phylogenetic Analysis of the Evolution of Display Behavior in the Neotropical Manakins (Aves, Pipridae). Ethology 84(3): 202-231.

Purvis A, Harvey PH (1995) Mammal Life-History Evolution - a Comparative Test of Charnovs Model. J Zool 237: 259-283.

Rice WR (1984) Sex-Chromosomes and the Evolution of Sexual Dimorphism. Evolution 38(4): 735-742.

Rice WR, Chippindale AK (2001) Intersexual ontogenetic conflict. J Evolution Biol 14(5): 685-693.

Rowe L, Day T (2006) Detecting sexual conflict and sexually antagonistic coevolution. Philos T Roy Soc B 361(1466): 277-285.

Simmons LW (1988) Male Size, Mating Potential and Lifetime Reproductive Success in the Field Cricket, *Gryllus bimaculatus* (Degeer). Anim Behav 36: 372-379.

Stillwell RC, Fox CW (2007) Environmental effects on sexual size dimorphism of a seed-feeding beetle. Oecologia 153(2): 273-280.

Trivers RL (1972) Mother-Offspring Conflict. Am Zool 12(4): 648-648.

Weber KE (1990) Selection on Wing Allometry in *Drosophila melanogaster*. Genetics 126(4): 975-989.

Weckerly FW (1998) Sexual-size dimorphism: Influence of mass and mating systems in the most dimorphic mammals. J Mammal 79(1): 33-52.

2

Genetic drift in antagonistic genes leads to divergence in sex-specific fitness between experimental populations of *Drosophila melanogaster*

A revised manuscript describing this material has been re-submitted for publication in *Evolution*

2.1 Introduction

By definition, males and females of any species with separate sexes differ phenotypically. This sexual dimorphism varies in degree and reflects the different roles of males and females in reproduction. Males typically invest less in parental care and achieve higher fitness by increasing their number of mating partners. Females, in contrast, usually provide higher levels of care and their fitness is primarily constrained by their capacity to produce and provide for offspring (Bateman 1948). The differences in sex roles cause divergent selection on morphological, physiological and behavioral characters, with males being subject to sexual selection and females experiencing selection for increased fecundity (Andersson 1994).

Although phenotypic differences between males and females are ubiquitous and sometimes striking, a growing body of work shows that the evolution of sexual dimorphism is often incomplete. A number of studies published over the past years have shown that 'sexual antagonism' (also called 'intra-locus sexual conflict'; Rice and Chippindale 2001; Bonduriansky and Chenoweth 2009) persists in many populations (e.g., Chippindale et al. 2001; Foerster et al. 2007; Mainguy et al. 2009; Svensson et al. 2009; Delph et al. 2011a). Sexual antagonism arises when alleles that benefit the fitness of one sex cause deleterious effects in the other sex. Sexually antagonistic alleles are viewed as a first evolutionary step towards the evolution of sexual dimorphism. Theory predicts that divergent selection on male and female phenotypes can favor the invasion of antagonistic mutations, as long as their benefit in one sex outweighs their cost in the other (Rice 1984). Once antagonistic variation segregates in the population, selection will theoretically favor the invasion of modifiers that limit the expression of antagonistic alleles to the favored sex, thus resolving antagonism and allowing the sexes to diverge phenotypically (Lande 1980; Rice 1984).

Population genetic models not only predict the invasion of sexually antagonistic mutations but also indicate that such alleles can be maintained in stable polymorphism (Kidwell et al. 1977; Gavrilets and Rice 2006; Fry 2010). The degree to which such polymorphism will persist depends on the dynamics of resolution. This could potentially be slow if it relies on the occurrence of rare events such as gene duplication (Connallon and Clark 2011) and could indeed be hampered altogether by deleterious pleiotropic effects of sex-specific gene expression (Mank et al. 2008). Irrespective of

the exact dynamics of its resolution, the fact that antagonistic variation has been found in a number of animal and plant populations indicates that detectable levels of such variation persist in many populations. As such, sexual antagonism is one of the major forces that potentially contribute to the maintenance of genetic variation for fitness (Chippindale et al. 2001; Bonduriansky and Chenoweth 2009).

When considering the role of antagonism in maintaining fitness variation it is important to realize that most existing theory of antagonism uses deterministic models that effectively assume populations of infinite size (e. g., Rice 1984; Gavrilets and Rice 2006). Similarly, the empirical studies demonstrating the existence of antagonistic variation are based on studies in large, outbred populations comprising many hundreds or thousands of individuals (the fruitfly population used in studies by Rice and co-workers (e.g., Chippindale et al. 2001), for example, is maintained at a population size of about 1,800 breeding adults). This means that antagonism has so far only been investigated under conditions where the evolutionary dynamics are dominated by the selective forces generated by antagonistic fitness effects. In contrast, the role of genetic drift has hitherto been ignored. Random changes in allele frequency are expected to affect the evolution of any phenotypic trait in small or subdivided populations. However, drift is likely to play a particularly prominent role in the evolution of sexually antagonistic traits (Connallon and Clark 2012). Due to their opposing effects on the fitness of males and females, the net selection pressure on sexually antagonistic mutations is often small, even when the sex-specific effects are large. Consequently the force of selection acting on antagonistic mutations, be it directional or balancing, is often weak and easily overcome by genetic drift when effective population sizes are small (Connallon and Clark 2012).

Genetic drift is also a very interesting phenomenon to investigate in the context of sexually antagonistic selection. Sexual antagonism is seen as a possible force maintaining genetic variation for fitness in populations (Patten et al. 2010; Connallon and Clark 2012) and it is therefore important to assess its evolutionary interplay with genetic drift as a force depleting genetic variation. Furthermore, as sexual antagonism is driven by selection on males and females and therefore by definition closely related to fitness, studying genetic drift in antagonistic variation will reveal how the sex-specific fitness of populations is shaped by random changes in gene frequencies. Due to the frequently weak net selection on antagonistic alleles, drift may cause differentiation in

sex-specific fitness between populations that would be highly unlikely to occur with genetic variation under sexually concordant selection. By affecting the overall productivity of populations, genetic drift in antagonistic alleles could have significant consequences for the long-term viability and survival of populations.

Here we describe experimental results on the effect of genetic drift on sex-specific fitness in four replicate populations of the fruitfly, *Drosophila melanogaster*. These were established from a large, outbred laboratory population and subsequently have undergone more than eighty generations of independent evolution under standardized conditions at a population size of one hundred individuals (50 males and 50 females). Due to the low number of individuals and the imposition of discrete generations by the maintenance regime, the effective size of each of these experimental populations is significantly smaller than that of the stock population as well as that of the laboratory populations previously used to study sexual antagonism in *D. melanogaster*. This material therefore constitutes an ideal opportunity to investigate how random drift affects the genetic architecture of fitness in finite populations.

We measured male and female larval and adult fitness of replicate genotypes sampled from each of the replicate populations. We found that populations had significantly diverged in both male and female average fitness. The patterns of divergence, however, differed between life stages. In the larval stage, populations had diverged independently of sex, whereas we found that divergence in adult fitness was sex-specific. For total fitness (the product of larval and adult fitness), populations did not differ in their overall mean, while their sex-specific total fitness varied greatly. We furthermore show that differences in sex-specific total fitness between populations are due to divergence along a sexually antagonistic fitness cline, ranging from high male/low female fitness to low male/high female fitness. The patterns of fitness variation we observed between populations are consistent with differentiation in reproductive fitness of the sexes being driven by genetic drift in sexually antagonistic alleles. We discuss the implications of our results with respect to antagonistic variation in the Dahomey base stock, from which our populations were derived, and more generally for the evolution of sexually antagonistic alleles in subdivided populations.

2.2 Methods and Materials

2.2.1 Experimental populations

We used four experimental populations in this study. These represent a subset of experimental populations originally established as part of a larger study (Reuter et al. 2008, 1:1 lines). The four populations were derived from the outbred and laboratory-adapted Dahomey wild-type stock and each was founded by 50 virgin males and 50 virgin females. Subsequently, populations were maintained independently under a standardized rearing regime as described in Reuter et al. (2008). Briefly, larvae were reared in culture bottles at constant densities (300 larvae per 65ml of food), eclosing adults were collected as virgins and new adult populations of 50 males and 50 females established in cages supplied with food (yeast paste) and oviposition media. Adult populations were allowed to interact and mate over a period of four days. Eggs for the subsequent generation were collected over the last 24 hours of the interaction period. All fly cultures were maintained at 21 degrees C throughout the life cycle.

The populations were established from the large and genetically diverse Dahomey stock. Due to their small numerical size and the imposition of a rearing regime with discrete generations, the replicate populations are expected to undergo increased levels of drift. Based on standard population genetic models (Crow and Kimura 1970, p. 350), the effective size of populations with 50 males and 50 females is $N_e = 100$. However, this calculation ignores the fact that the number of matings is finite. Using a more refined model (Balloux and Lehmann 2003, Eq. 7) and empirical estimates of the frequency of double matings in the experimental conditions under which the lines evolved (Table 1 in Reuter et al. 2008), the effective size of the populations is predicted to be in the order of $N_e \approx 80$.

When analyzed here, the populations had undergone approximately eighty generations of experimental evolution under the conditions described above. We then measured male and female larval and adult fitness of 81 haploid genomes ('hemiclones') sampled randomly from the four populations (18-22 from each population).

2.2.2 Sampling and amplification of hemiclones

Hemiclones were sampled and multiplied using 'cytogenetic cloning' (see Rice 1996; Abbott and Morrow 2011 for descriptions of approach and Fig. 1 for a schematic

representation of the crossing scheme used here). Haploid genomes were extracted from the populations by mating randomly sampled males to females of the 'Clone Generator' (CG) stock (Fig. 1A) and back-crossing single male offspring once again to females of the CG stock (Fig. 1B). The genotype of CG females (compound X, Y chromosome, homozygous viable translocation of chromosomes II and III, see (Rice 1996)) ensures paternal transmission of the X chromosome to sons and co-segregation of the paternal second and third chromosomes, making it possible to produce many males carrying an identical set of randomly sampled X, II and III (the fourth chromosome, which carries only about 0.5% of the coding genes in *D. melanogaster* , is ignored here for pragmatic reasons). The haploid complement of X, 2nd and 3rd chromosomes is hereafter referred to as a Target Genome (TG) and, due to the absence of male recombination in *Drosophila*, can be maintained and multiplied by crossing their male carriers to CG females.

2.2.3 Expression of TGs in males and females

To assay their effect on male and female fitness, TGs were expressed in an outbred genetic background in males and females, complemented with genomes randomly sampled from their population of origin. To express TGs in a female background, males carrying a TG in a CG background were crossed with multiple virgin females of their corresponding population of origin (Fig. 1C, right-hand side). Half of the females produced in this cross inherited an identical target genome from the father (the other half received the target X together with the eye-color marked translocation of chromosomes II and III), complemented with different maternal genomes. To express TGs in males, males carrying a TG in clone generator background were mated to multiple females of a stock carrying a compound X chromosome [C(1)DX], a Y chromosome and autosomes of the TG's population of origin (Fig. 1C, left-hand side). The compound X of the females ensured paternal transmission of the X chromosome. Accordingly, half of the emerging males produced in this cross inherited an identical target genome from the father (with others receiving the target X and the II-III translocation), complemented with different Y chromosomes and autosomes contributed by the mothers.

2.2.4 Larval fitness assay

We measured larval fitness of TGs as survival under conditions comparable to the rearing regime under which the flies had evolved. For this purpose, the crosses to

express TGs in males and females (described above) were set up in small cages supplied with grape juice plates for egg-laying. Eggs laid on the grape juice plates were incubated until first instar larvae hatched. In parallel, cultures were similarly maintained of a competitor strain marked with a recessive eye color mutant, *sparkling poliert* in an outbred genetic background sampled from the four experimental evolution lines (Fig. 1). Cultures of competing larvae were then set up for each TG by transferring 150 larvae from the TG expression cross and 150 eye color-marked competitor larvae to a 190ml food bottle containing 65ml of media. Flies eclosing from the cultures were counted under cold anesthesia. Larval fitness was calculated as the proportion of wildtype flies of the sex under investigation, minus the expected proportion of 1/8 (this was the expectation because half of the larvae transferred were descendants of the expression cross, of which half were of the desired sex, again half of which were of the desired genotype.

2.2.5 Male adult fitness assay

Male adult fitness was measured as fertilization success under competitive conditions that were similar to the populations' rearing regime. For each fitness assay, we set up a cage containing a target of 15 males sharing a particular TG, 35 eye color-marked competitor males and 50 eye color-marked females. In a small number of cases (~10% of genomes) fewer than 15 TG males were available for a target genome. In these cases we added a further complement of competitor males to attain a total of 50 males. All flies used were virgin and, in effect, between 1 and 3.5 days of age (they had matured between 1 and 5 days at 18 degrees C). In line with the conditions of the rearing regime, flies were allowed to interact in the cages for four days and cages were supplied daily with fresh grape juice plates (as oviposition media) and *ad libitum* yeast paste. After the end of day 4, males were discarded and females were placed individually into yeasted vials to lay eggs for a further 3 days. Females were then discarded and their progeny left to develop. Upon eclosion, progeny were scored for eye color and counted. The mating success of target males (probability of a female mating with a wildtype rather than eye color-mutant male) was estimated from the scores obtained from the fifty females of an assay, using a Bayesian procedure described in the Appendix of Reuter et al. (2008). This estimation takes into account the fact that different numbers of matings with males of the same phenotype produce batches of offspring all of the same eye color (see Reuter et al. 2008). This mating probability was divided by the expected probability

under random mating (number of TG males in the cage/50) to obtain a measure of male adult fitness.

2.2.6 Female adult fitness assay

Female adult fitness was measured as egg laying rate under competitive conditions similar to those of the rearing regime under which the populations evolved. For each fitness assay, a cage with a grape juice egg laying plate and ad libitum yeast paste was set up, containing a target of 15 females sharing a particular TG, 35 eye color -marked competitor females and 50 eye color-marked males. Again, in rare cases (<5% of genomes) fewer than 15 TG females were available and numbers were boosted with a further complement of competitor females to attain a total of 50 females. All flies used were virgin and between, in effect, 1 and 3.5 days of age (1-5 days at 18 degrees C). Flies were allowed to interact for three days. At the end of the third day, females were isolated in individual vials and allowed to lay eggs for one day (equivalent to the last 24 hours of the rearing cycle). After this period, females were discarded and their progeny left to develop. Upon eclosion, offspring were counted. The average fertility of wildtype females in the cage, divided by the average fertility of the eye color-marked competitors in the cage, was used as the female adult fitness measure.

2.2.7 Total and relative fitness

We calculated relative fitness of the genomes for each sex and life stage (larva, adult) separately by dividing individual fitness values by the average fitness across all populations. Total fitness values of individual genomes were calculated for each sex separately by multiplying the relative, sex-specific values of larval and adult fitness.

2.2.8 Statistical analysis

In order to ensure sufficient quality of our dataset we removed target genomes for which we deemed fitness data unreliable. Thus, we removed from the analysis one target genome for which fewer than 140 flies eclosed in one of the larval fitness assays, six target genomes for which fewer than ten adult TG males or females entered the fitness assay and two further target genomes that presented outlier values for one of the fitness measures (defined here as differing by more than 2.5 standard deviations from the population average).

We analyzed the data of our experiments using standard parametric statistics in R (R Development Core Team 2006). In analyses of variance, population was modeled as a fixed effect. Although differences between populations would potentially be more appropriately represented as a random variable, the low number of populations analyzed (four) did not allow for a reliable estimation of between-population variances and covariances (Crawley 2002, p. 670). Principal Component Analyses were performed based on covariances.

For all analyses we verified that the distribution of the data matched the assumption of the tests used. Where required, we transformed the data and indicate so when reporting the results. We also confirmed that despite being based on the same individuals, our measures of larval and adult fitness were independent. It is conceivable that genomes with high larval fitness would experience greater larval competition and accordingly show lowered adult fitness. Due to the large excess of larval growth media used here, such effects are unlikely and we can formally rule them out because adult fitness in both sexes was uncorrelated with the total number of flies eclosing from the larval growth cultures (Pearson's product moment correlation; males: $r=0.16$, $t_{75}=1.36$, $P=0.18$; females: $r=0.08$, $t_{75}=0.67$, $P=0.51$).

2.3 Results

We obtained estimates for the four fitness components (male and female larval and adult fitness) for a total of 77 and an average of 19.3 target genomes per population (i.e., 17, 19, 20 and 21 replicate genomes for the four populations).

We first assessed differences between populations in sex-specific relative fitness by analyzing larval and adult data separately. We performed ANOVAs of larval and adult fitness with population, sex, and their interaction as independent factors. For larval fitness, we found that populations had significantly diverged in fitness (population term; Table 2.1), but that this divergence did not differ between the sexes (population-by-sex term; Table 2.1). For adult fitness, we observed significant between-population divergence in fitness across both sexes (population term; Table 2.2). However, the degree and direction of divergence between populations also differed strongly between the sexes (population-by-sex term; Table 2.2). As male and female relative fitness values in the larval and adult stages are all standardized to an average of unity, the sex effect is not significant in either analysis.

We also applied the same ANOVA model to total fitness (Fig. 2C). This analysis showed that across both life stages, populations had diverged in a sex-specific manner and differed strongly in average sex-specific total fitness (population-by-sex term Table 2.3). In contrast, neither populations nor the sexes differed in average total fitness (population term; Table 2.3).

In order to better illustrate how populations diverged in fitness, we performed a Principal Component Analysis (PCA) of the measures of male and female total fitness. The two axes generated by this analysis provide an intuitive interpretation of fitness variation (Fig. 3). The major axis, capturing 61% of the variation, expresses the position of target genomes along an antagonistic continuum between high male/low female fitness and high female/low male fitness. The minor axis, capturing the remaining 39% of variation, expresses the overall, sexually concordant, quality of genomes (Fig. 3). Separate ANOVAs on the principal component scores of the genomes on the two axes showed that populations differed significantly in their score on the first, sexually antagonistic, axis (Table 2.4). In contrast, populations did not differ significantly in their scores for the second, sexually concordant, axis (Table 2.4).

We also performed a PCA on the four individual fitness measures, male and female larval and adult fitness. Although the interpretation of the PC axes in this case is less intuitive than in the case of male and female total fitness, the axes are informative about the qualitative patterns in fitness of each sex at each life stage. Specifically, the first axis captures a net negative effect of male fitness and a net positive effect of female adult fitness (see Table 2.5, PC1 for loadings). This axis can therefore be interpreted as sexually antagonistic and similarly to the outcomes of the prior analysis of total fitness, we found significant differences between populations along this axis ($F_{3,73}$=14.1, P<0.0001). The third axis expresses a similar effect, with a net positive effect of male fitness and a net negative effect of female fitness (Table 2.5, PC3). However, populations did not differ in their position along this axis ($F_{3,73}$=0.18, P=0.91). The fourth axis is sexually concordant with positive loadings for all four fitness measures (Table 2.5, PC4) and, as in the previous analysis of total fitness, populations did not differ in their score along this axis ($F_{3,73}$=1.7, P=0.18). Finally, the second PC axis revealed effects that were not visible in the analysis of total fitness. This axis expresses a negative correlation between larval and adult fitness. The association between low larval and high adult fitness was pronounced in females and weak in males (Table 2.5, PC2) and populations differed significantly in their position along this axis ($F_{3,73}$=6.0, P=0.001).

2.4 Discussion

The results we present here demonstrate that small and independently evolving populations can diverge significantly in their sex-specific fitness. Importantly, our data suggest that this divergence does not arise because some populations fix more deleterious mutations than others and hence suffer from an overall decrease in fitness across both sexes (PC2 in Fig. 3). Rather, divergence occurs along a sexually antagonistic fitness continuum. Thus, populations that increase in the fitness of one sex tend to decrease in the fitness of the other, with minimal change in the average fitness across both sexes.

The way in which populations diverge on a continuum between high male/low female and low male/high female fitness suggests that population differentiation mainly occurs through changes in the frequency of sexually antagonistic alleles. Allele frequency changes have most likely occurred through random genetic drift because the populations have evolved under tightly controlled and standardized conditions. Drift could have occurred at different stages in the history of the populations used in this study. The initial establishment of the small experimental populations from the large and genetically diverse Dahomey stock will have induced founder events, but stochastic changes can also have taken place subsequently, during the many generations that the populations were maintained at a small effective population size. Whatever the underlying cause of genetic drift, be it initial sampling or reproductive stochasticity, random changes in the frequency of antagonistic alleles caused the population mean fitnesses for each sex to diverge in an antagonistic pattern, whereby some populations increased in the fitness of males but decreased in that of females while others underwent changes in the opposite direction.

Although it is clear that genetic drift will affect the evolution of phenotypic traits in small populations, fitness is by definition under strong selection and the large differences in sex-specific performance we observe between populations (Fig. 2) may seem surprising. The rapid divergence in sex-specific fitness is, however, in line with theory predicting that sexually antagonistic variation should be highly sensitive to genetic drift (Connallon and Clark 2012). One reason for this is that opposing fitness effects in males and females can result in weak net selection across the sexes, meaning that mutations can be almost neutral despite having strong effects on the fitness of each

sex. Quasi-neutral variation of this type can persist for long periods of time in populations with large effective sizes, but will erode rapidly when subjected to more intense genetic drift. The loss of genetic variation will then reveal the fitness effects of the segregating alleles, leading to potentially large differences in male and female fitness between populations, such as those observed here. This situation contrasts with polymorphism under sexually concordant selection. Here, classical theory predicts that genetic variation will only persist for appreciable amounts of time (rather than being eliminated rapidly by selection) if the product of effective population size and selection coefficient is smaller than one ($N_e s < 1$). This means that the level of fitness variation that can be maintained under mutation-selection balance in large populations is small, and certainly not large enough to generate a divergence in fitness between small sub-populations comparable to that observed here.

The antagonistic fitness divergence that we infer between our populations is also interesting because it provides indirect evidence of sexually antagonistic variation segregating within their population of origin, Dahomey. This is an outbred stock population that has been maintained in the laboratory for over thirty years at large population size with overlapping generations. Sexually antagonistic variation has previously been shown to occur in two other independent laboratory populations. Antagonism was revealed in the LHm population by the imposition of male-limited experimental evolution (Rice 1996) and by quantitative genetic analysis of standing variation (Chippindale et al. 2001). Genetic variation with sexually antagonistic effects was also shown to segregate in the IV population, from a comparison of paternal and offspring fitness (Connallon and Jakubowski 2009). Taken together, our findings and the previous results suggest that sexual antagonism is widespread in laboratory populations of *D. melanogaster*. Whether this variation is stably maintained over long periods of time is currently unknown. On the one hand, all three populations, LHm, IV and Dahomey, had been maintained in the laboratory for several decades before being assessed for antagonism, indicating that variation can be maintained over long periods of time. On the other hand, there is some evidence that domestication can result in a loss of additive genetic variation with sex-specific fitness effects. Jiang et al. (2011) repeated a male-limited evolution experiment first conducted approximately 15 years earlier on the same base population, LHm, (Rice 1996) but were unable to replicate the changes in sperm competitiveness documented in the earlier studies. It thus appears that over the course of domestication, additive genetic variation with sexually antagonistic

effects can be lost. Artificial selection experiments in plants have also shown that genetic correlations between male and female floral traits that would be expected to maintain sexually antagonistic variation can be easily broken down (Delph et al. 2011b). The conflicting messages of these datasets illustrates that we continue to lack a full understanding of how sexually antagonistic variation is maintained and of the identity of those factors that affect the rate at which sexual antagonism is resolved (Stewart et al. 2010).

We can use our results to draw some inferences about the genetics of fitness. In particular, we can compare the patterns of fitness divergence between lines and across different life stages. Similar to Chippindale et al. (2001), we found that divergence between populations in larval fitness was sexually concordant (Fig. 2A), indicating that populations accumulated alleles that were either generally beneficial or generally deleterious to larval performance, independently of sex. This is in line with the view that juveniles do not have differentiated sex roles and accordingly mutations will impact the fitness of males and females to a similar extent by either increasing or decreasing larval performance. In adults, in contrast, significant sexually antagonistic effects were observed in the adult stage, reflecting the difference between male and female reproductive roles (a significant sex-by-population interaction; Fig. 2B). While population fitness diverged in a sex-specific manner, small differences in overall fitness were also apparent (a significant population effect). Interestingly, these disappeared when comparing total fitness between populations (Fig. 2C), indicating that differences in overall adult fitness and differences in overall larval fitness cancelled each other out. This was also supported by the Principal Component Analysis of individual fitness components, where one axis of comprised negative loadings for larval fitness components, but positive loadings for adult fitness (axis 2). The association between increased larval and decreased adult fitness suggests that sexual antagonism over male and female adult phenotypes is overlaid by adaptive conflict over optimal larval and adult phenotypes. Our data suggest that some genotypes increase larval fitness at the expense of adult performance, while others have the opposite effect.

In order to obtain more insights into the nature and dynamics of fitness evolution in our populations, we compared our current fitness data to the results of earlier analyses of the larger set of selection lines from which the populations here were drawn (Reuter et al. 2008). The earlier study was conducted after about 30 generations of evolution and

assayed male testis and accessory gland sizes in the experimentally evolved populations. The average measures of accessory gland size obtained by (Reuter et al. 2008) and the average male adult fitness measured here across the four populations are not significantly correlated (Pearson's Product Moment Correlation, r=0. 46, t_2= 0.74, P=0.54). However, average testis size does correlate significantly with average male adult fitness (r=0.98, t_2=6.35, P=0.024), despite the small number of datapoints (N=4). This striking result suggests that male reproductive morphology contributes to sexually antagonistic fitness effects. This is surprising, given that these traits are male-specific and antagonism is thought to arise in general from divergent selection on characters that are shared between the sexes. Unless spurious, the correlation we have found implies that the size of some sex-specific morphological characters is influenced by developmental mechanisms that are shared between the sexes and thus subject to antagonistic selection.

2.4.1 Conclusions

In our study, we have observed divergence in sex-specific fitness between laboratory populations of small size. It is conceivable, and indeed likely, that similar processes occur in natural populations that are subdivided into small groups of reproductive individuals, linked by low levels of migration. Genetic drift in sexually antagonistic genetic variation is of potential significance since it could lead to subgroups differing in their sex-specific fitness, with some showing increased female and decreased male fitness while other subgroups show the opposite pattern. Interestingly, this kind of fitness divergence could have an impact on the evolutionary dynamics of sexually antagonistic loci. This would be the case, for example, in populations that are composed of local demes occupied by groups of breeding individuals. Such 'meta-populations' are usually characterized by some turnover in demes, where breeding groups can go extinct and be re-founded by offspring emigrating from other demes. If the persistence of demes were to vary with female fitness, then increased extinction rates of demes with low female fecundity would add a selective pressure against female-detrimental alleles at the level of the meta-population. This additional force would act even in the absence of within-deme competition and be expected to shift the conditions for the invasion and maintenance of sexually antagonistic alleles. Accordingly, it would become harder for new male-beneficial/female-detrimental mutations to establish themselves in the meta-population and those that did so would be expected to segregate at lower frequencies compared to those in the absence of meta-population dynamics. Similar effects of multi-

level selection have been proposed to influence the evolution of male traits that increase male fitness at the expense of that of their mating partners. While positively selected within mating groups, such male traits can be selected at the level of the population because groups with less harmful males have a higher overall productivity (Eldakar et al. 2009). In the future, it would be interesting to develop models that generate predictions for the dynamics of antagonistic alleles in a meta-population context.

2.5 References

Abbott, J. K., and E. H. Morrow. 2011. Obtaining snapshots of genetic variation using hemiclonal analysis. Trends Ecol Evol 26:359-368.

Andersson, M. 1994. Sexual selection. Princeton University Press, Princeton, NJ.

Balloux, F., and L. Lehmann. 2003. Random mating with a finite number of matings. Genetics 165:2313-2315.

Bateman, A. J. 1948. Intra-sexual selection in *Drosophila*. Heredity 2:349-368.

Bonduriansky, R., and S. F. Chenoweth. 2009. Intralocus sexual conflict. Trends Ecol. Evol. 24:280-288.

Chippindale, A. K., J. R. Gibson, and W. R. Rice. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. Proc. Natl. Acad. Sci. U. S. A. 98:1671-1675.

Connallon, T., and A. G. Clark. 2011. The resolution of sexual antagonism by gene duplication. Genetics 187:919-937.

Connallon, T., and A. G. Clark. 2012. A general population genetic framework for antagonistic selection that accounts for demography and recurrent mutation. Genetics 190:1477-1489.

Connallon, T., and E. Jakubowski. 2009. Association between sex ratio distortion and sexually antagonistic fitness consequences of female choice. Evolution 63:2179-2183.

Crawley, M. J. 2002. Statistical computing: an introduction to data analysis using S-PLUS. John Wiley and Sons, Chichester.

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

Delph, L. F., J. Andicoechea, J. C. Steven, C. R. Herlihy, S. V. Scarpino, and D. L. Bell. 2011a. Environment-dependent intralocus sexual conflict in a dioecious plant. New Phytol. 192:542-552.

Delph, L. F., J. C. Steven, I. A. Anderson, C. R. Herlihy, and E. D. Brodie III. 2011b. Elimination of a genetic correlation between the sexes via artificial correlational selection. Evolution 65:2872-2880.

Eldakar, O. T., M. J. Dlugos, J. W. Pepper, and D. S. Wilson. 2009. Population structure mediates sexual conflict in water striders. Science 326:816-816.

Foerster, K., T. Coulson, B. C. Sheldon, J. M. Pemberton, T. H. Clutton-Brock, and L. E. B. Kruuk. 2007. Sexually antagonistic genetic variation in red deer. Nature 447:1107-1111.

Fry, J. D. 2010. The genomic location of sexually antagonistic variation: some cautionary comments. Evolution 64:1510-1516.

Gavrilets, S., and W. R. Rice. 2006. Genetic models of homosexuality: generating testable predictions. Proc. R. Soc. Lond. B 273:3031-3038.

Jiang, P. P., S. Bedhomme, N. G. Prasad, and A. Chippindale. 2011. Sperm competition and mate harm unresponsive to male-limited selection in *Drosophila*: an evolving genetic architecture under domestication. Evolution 65:2448-2460.

Kidwell, J. F., M. T. Clegg, F. M. Stewart, and T. Prout. 1977. Regions of stable equilibria for models of differential selection in 2 sexes under random mating. Genetics 85:171-183.

Lande, R. 1980. Sexual dimorphism, sexual selection and adaptation in polygenic characters. Evolution 34:292-305.

Mainguy, J., S. D. Cote, M. Festa-Bianchet, and D. W. Coltman. 2009. Father-offspring phenotypic correlations suggest intralocus sexual conflict for a fitness-linked trait in a wild sexually dimorphic mammal. Proc. R. Soc. Lond. B 276:4067-4075.

Mank, J. E., L. Hultin-Rosenberg, M. Zwahlen, and H. Ellegren. 2008. Pleiotropic constraint hampers the resolution of sexual antagonism in vertebrate gene expression. Am. Nat. 171:35-43.

Patten, M. M., D. Haig, and F. Ubeda. 2010. Fitness variation due to sexual antagonism and linkage disequilibrium. Evolution 64:3638-3642.

R Development Core Team. 2006. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reuter, M., J. R. Linklater, L. Lehmann, K. Fowler, T. Chapman, and G. D. D. Hurst. 2008. Adaptation to experimental alterations of the operational sex ratio in populations of *Drosophila melanogaster*. Evolution 62:401-412.

Rice, W. R. 1984. Sex-chromosomes and the evolution of sexual dimorphism. Evolution 38:735-742.

Rice, W. R. 1996. Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. Nature 381:232-234.

Rice, W. R., and A. K. Chippindale. 2001. Intersexual ontogenetic conflict. J. Evol. Biol. 14:685-693.

Stewart, A. D., A. Pischedda, and W. R. Rice. 2010. Resolving intralocus sexual conflict: genetic mechanisms and time frame. J. Hered. 101:S94-S99.

Svensson, E. I., A. G. McAdam, and B. Sinervo. 2009. Intralocus sexual conflict over immune defense, gender load, and sex-specific signaling in a natural lizard population. Evolution 63:3124-3135.

2.6 Tables and figures

Table 2.1 Larval fitness between the four populations and between the sexes. Table showing the results of a two-way ANOVA on the dependent variable, (log-transformed) larval fitness, with the independent variables, sex, selection line, and the interaction term, sex-by-selection line.

Larval fitness

	DF	*Sum Sq*	*F-value*	*P-value*
Sex	1	0.0017	0.0321	0.858
Line	3	0.7676	4.7745	0.003
Sex*Line	3	0.2404	1.4955	0.218
Residuals	146	7.8238		

Table 2.2 Adult fitness between the four populations, and between the sexes. Table showing the results of a two-way ANOVA on the dependent variable, adult fitness, with the independent variables, sex, selection line, and the interaction term sex-by-selection line.

Adult fitness

	DF	*Sum Sq*	*F-value*	*P-value*
Sex	1	0	0	1
Line	3	0.8937	5.0318	0.002
Sex*Line	3	2.1976	12.3728	<0.001
Residuals	146	8.6438		

Table 2.3 Total fitness between the four populations and between the sexes. Table showing the results of a two-way ANOVA on the dependent variable, (log-transformed) total fitness, with the independent variables, sex, selection line, and the interaction term, sex-by-selection line.

Total fitness

	DF	*Sum Sq*	*F-value*	*P-value*
Sex	1	0.0001	0.0011	0.973
Line	3	0.4185	1.1995	0.312
Sex*Line	3	3.0703	8.8005	<0.001
Residuals	146	16.9789		

Table 2.4 Results of two, one-way ANOVAs using the principal component scores of genomes on the sexually antagonistic and sexually concordant axes as the dependent variables, with selection line as the independent variable.

PCA Scores

Antagonistic axis

	DF	*Sum Sq*	*F-value*	*P-value*
Line	3	2.8626	7.2709	<0.001
Residuals	73	9.5801		

Concordant axis

	DF	*Sum Sq*	*F-value*	*P-value*
Line	3	0.6281	2.1161	0.105
Residuals	73	7.2227		

Table 2.5 Axis loadings of a Principal Component Analysis of larval and adult relative fitness in males and females. The data entries in rows 1 to 4 of the table specify the weighting of each of the four fitness components in each of the four PC axes. The values in rows 5 and 6 provide the percentage of variance in the data captured by each of the axes and the P-values of one-way ANOVAs testing the difference between populations in scores on each of the four axes, respectively.

	PC1	*PC2*	*PC3*	*PC4*
Male larval fitness	+0.19	-0.05	+0.47	+0.86
Male adult fitness	-0.92	+0.14	-0.18	+0.31
Female larval fitness	0.16	-0.60	-0.72	0.32
Female adult fitness	0.29	0.79	-0.48	0.25
Variance captured	35%	27%	23%	15%
P-value	<0.0001	0.001	0.91	0.18

Figure 1. Crossing scheme for hemiclonal analysis. The figure illustrates the crosses needed to (A) sample, (B) amplify and (C) express individual haploid genomes (chromosomes X, II and III) from the four experimental populations. Chromosomes X/Y, II and III are shown left to right. The co-segregating translocation of chromosomes II and III carried by the clone-generator (CG) stock is shown as horizontally linked chromosomes II and III; compound X chromosomes are shown as linked in a horizontal V-shape. Chromosomes forming a target genome are denoted by a T. The males and females shown at the bottom of panel C are those that enter the fitness assays.

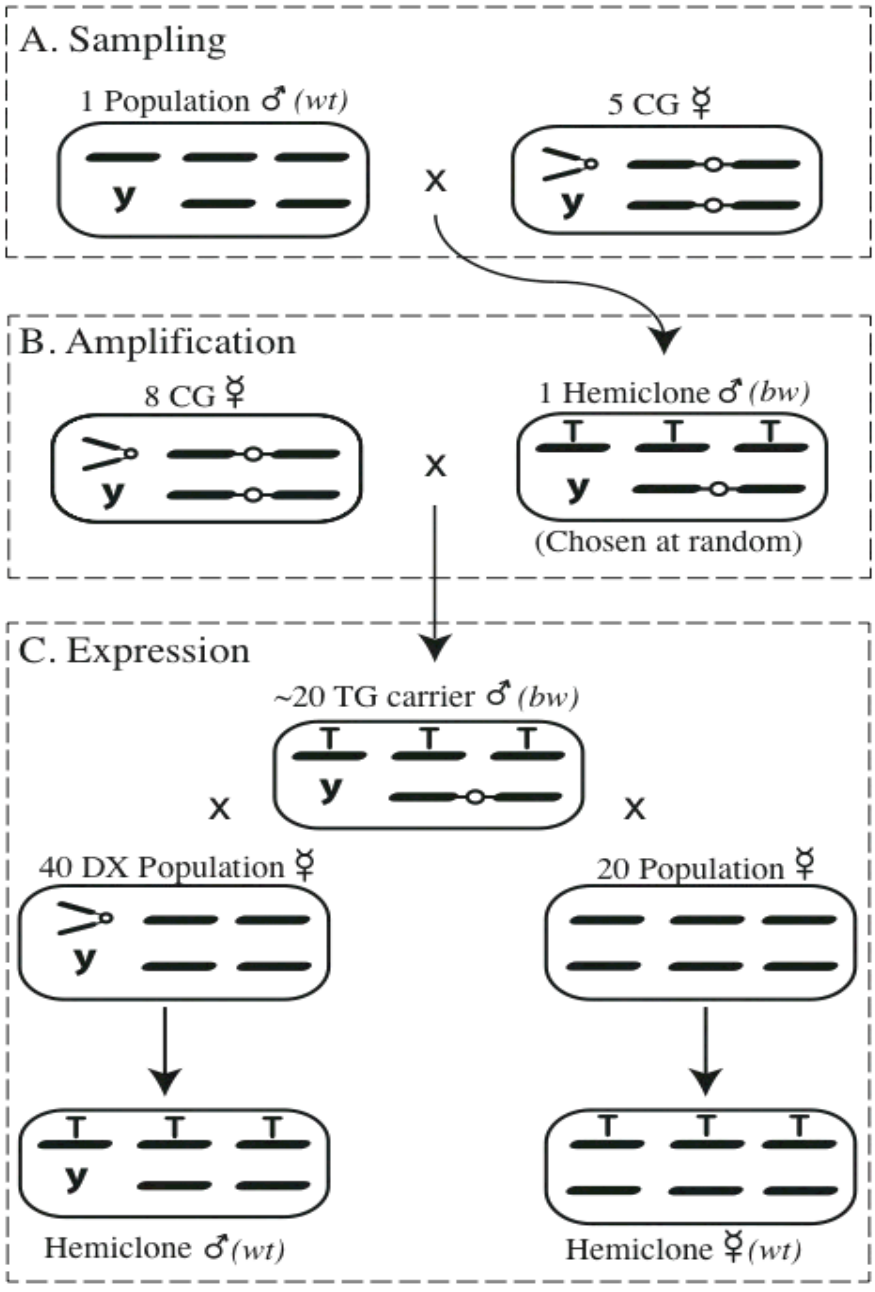
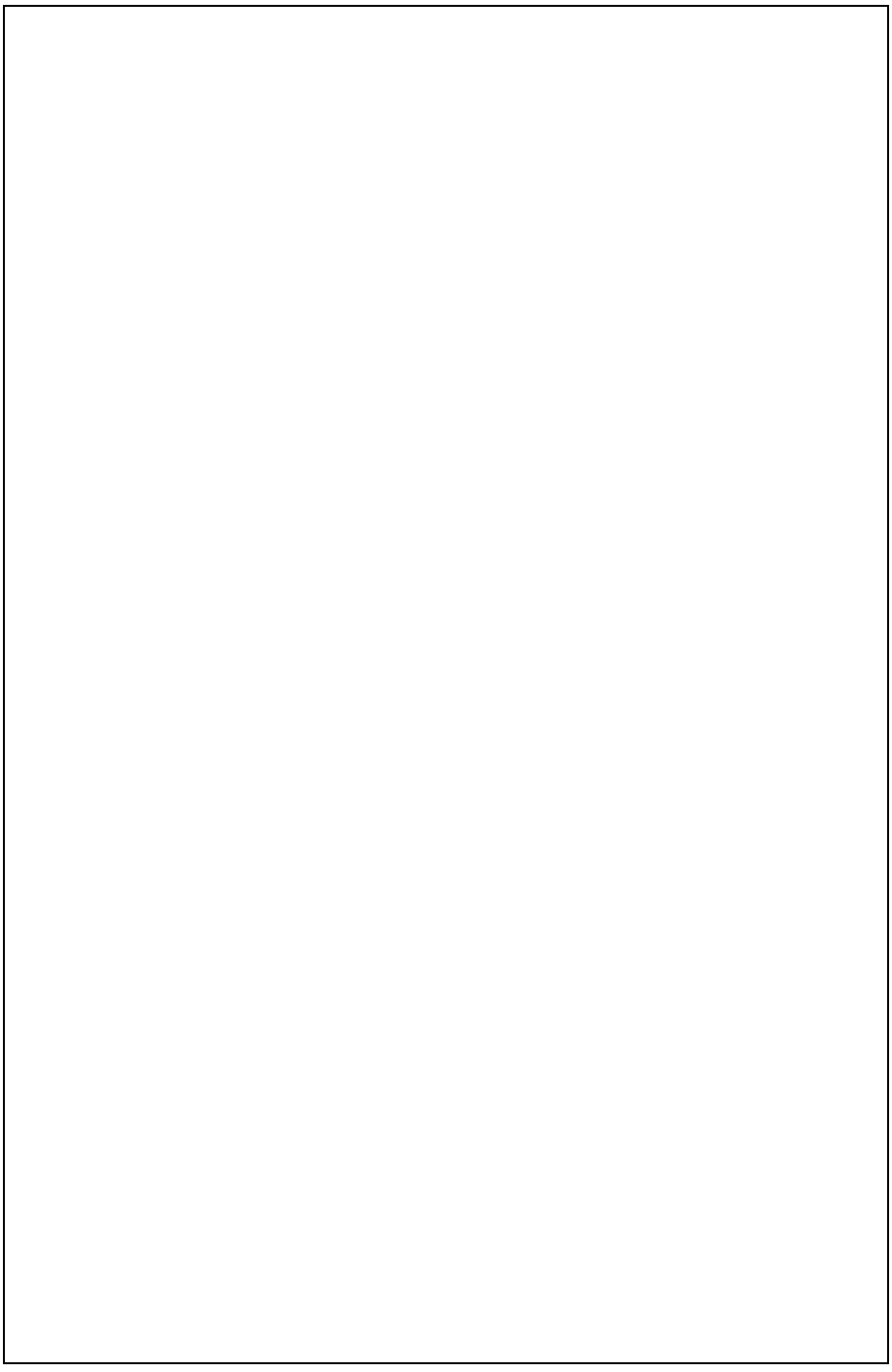


Figure 2. Mean male and female fitness in the four experimental populations. The figure shows larval (A), adult (B) and total fitness (C) in males and females of each population. Lines connect male and female fitness values from the same population, vertical bars indicate standard errors.



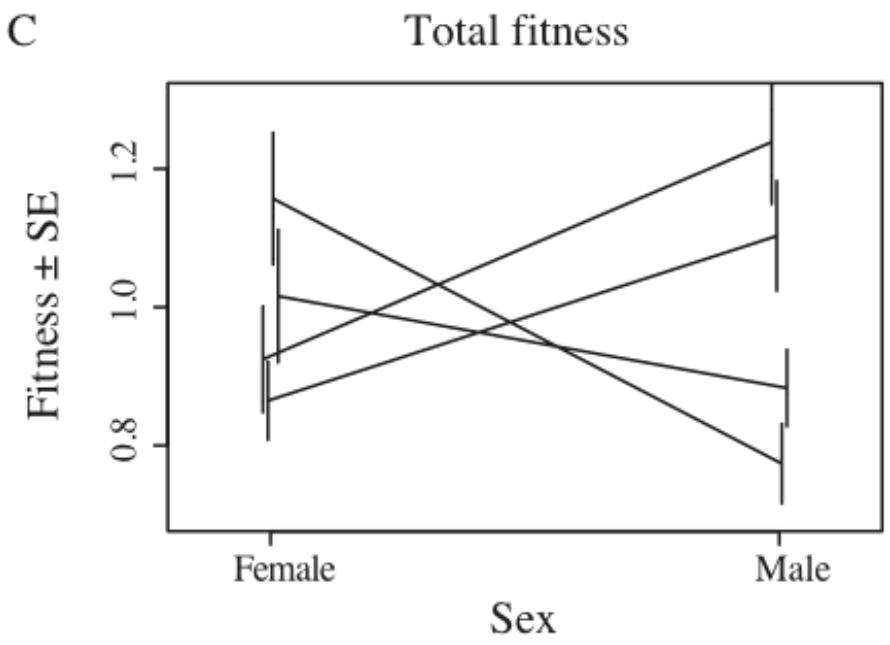
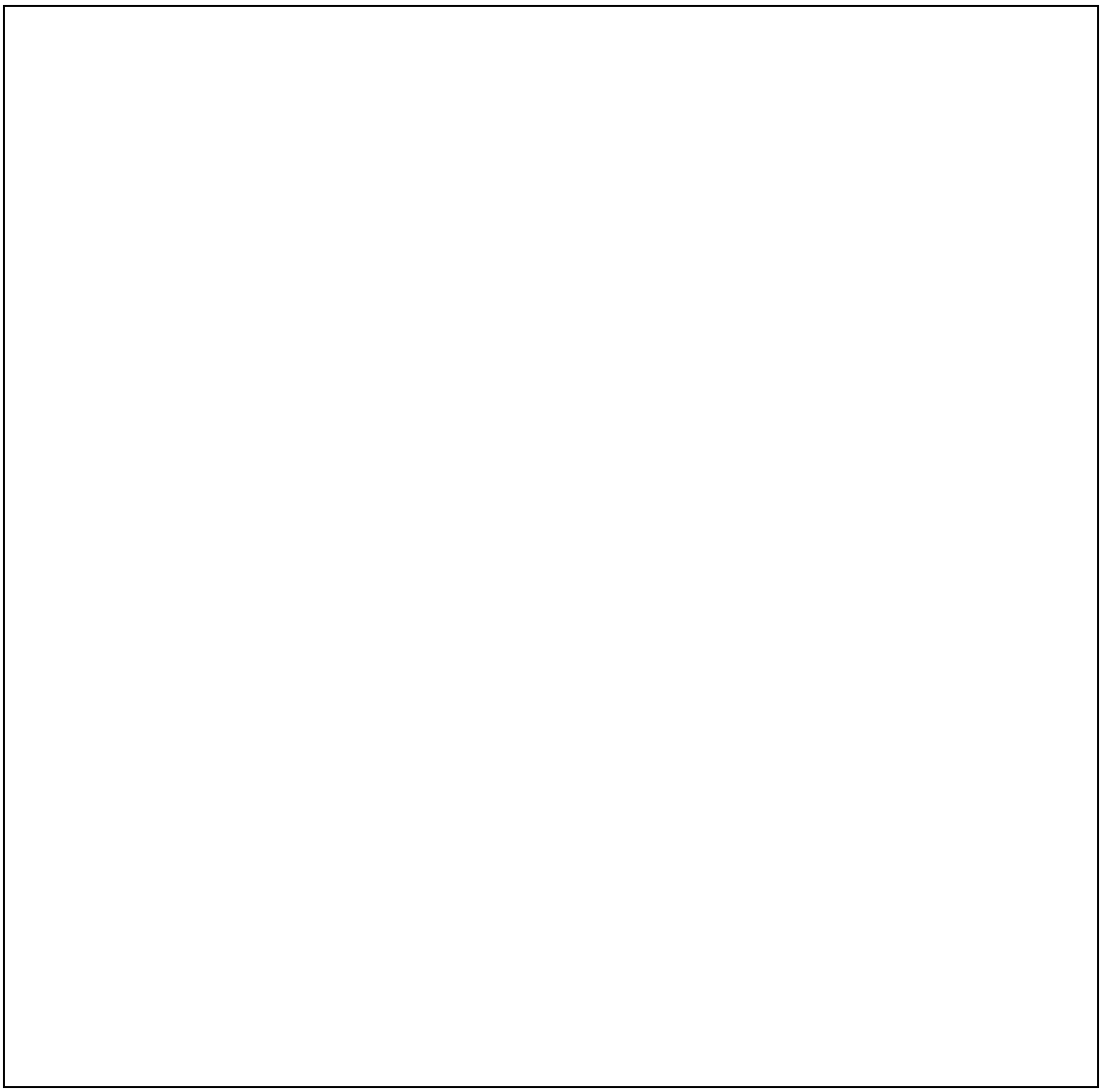

C
Total fitness
Fitness ± SE
0.8
1.0
1.2
Female
Male
Sex

Figure 3. Male and female total fitness of individual target genomes. The four populations of origin are indicated by varying shades of grey. The direction of the two arrows indicate the orientation of the two principal component axes PC1 and PC2 and the relative length of the two arrows is proportional to the proportion of variance captured.



3

Morphological determinants of male fitness in *Drosophila melanogaster*

3.1 Introduction

Sexual dimorphism is the result of selection that favors different trait values in males and females. Evidence suggests that a wide range of traits differ in their optimal value between the sexes, including growth rate, locomotion, metabolic rate, thermoregulation, and size (Glucksmann 1981), to name but a few. This phenotypic separation between males and females could suggest that each sex is subject to differential selection, and therefore pursues different habits in their fight for survival. However, such differences in survival strategy are a rare occurrence, especially for higher animals (Darwin 1874), and this does not explain the extent, and elaborate nature, of some sexually dimorphic traits. On the contrary, male-specific traits frequently reduce the survival of the individuals carrying them. Many of these adaptations are driven by sexual selection, competition between members of the same sex to maximize reproduction. This intrasexual competition is most intense in males where parental investment is low, and reproductive success is positively correlated with the number of successful copulations each male can acquire. As a consequence males typically gain a reproductive advantage from numerous matings. On the contrary females often gain relatively small reproductive benefits from mating repeatedly, and can even pay a cost through increased mortality or reduced fertility (Fowler and Partridge 1989). For example, in polygynous insects female fitness is often maximized by one, or just a few, matings. For females in these species their optimal mating rate reflects a trade-off between the benefits and costs of mating. For example, female *D. melanogaster* experience a short-term benefit of increased egg production from multiple matings, but pay a long-term cost of reduced longevity (Partridge et al. 1987; Fowler and Partridge 1989; Chapman et al. 1998), whilst male fitness is only limited by the number of mating opportunities (Arnqvist and Nilsson 2000). Males suffer no obvious costs from multiple matings. Within many species, intense intrasexual competition among males can drive the evolution of highly variable phenotypes and complex reproductive behaviors. Furthermore, the rapid and exaggerated divergence of these sexually selected traits is actually thought to facilitate speciation in some instances (Westeberhard 1983). The variability and importance (e.g. Burkhardt et al. 1994; Bouteiller and Perrin 2000) of sexually selected traits for male reproductive success results in males frequently experiencing greater reproductive variance than females within a given population (e.g. Bouteiller and Perrin 2000).

In highly polygynous species, where males face intense competition for mating opportunities we frequently observe exaggeration of traits for use as weapons and/or ornaments during intrasexual competition (Harvey et al. 1978; Cluttonbrock et al. 1980; Houle and Rowe 2003). Body size is a trait that affects reproductive success for many species. In polygynous mammals there is often a large difference in body size between the sexes, with males usually larger than females. This sexual size dimorphism (SSD) evolves when the relationship between reproductive success and body size, differs between the sexes, causing selection to favor different adult body sizes for males and females. For many species body size is an essential determinant of reproductive success. A textbook example is elephant seals, a species that shows extreme sexual size dimorphism (SSD) with males typically 2-7 times heavier than females. For each male, reproductive success is principally controlled by dominance rank, which correlates strongly with size relative to other males, largely because it confers an advantage during fights. A male's ability to dominate other males to gain access to females is essential, as less than one third of males copulate each season. The reproductive potential of male elephant seals can reach 17 times that of a female (Le Boeuf 1974), providing a significant reproductive payoff, and hence a large incentive for intense male-male competition. Whilst males benefit from very active behaviour to gain mating opportunities, females generally gain reproductive advantages through variables such as age and experience (Le Boeuf 1980). In general this makes a more passive female role suitable for optimal reproductive success, exemplifying the divergent selective pressures experienced between the sexes.

Whilst we typically observe larger males than females in mammals that exhibit SSD (Ralls 1977), the opposite pattern is found in insects, for which females are usually the larger sex (Fairbairn 1997). Fecundity is positively correlated with adult body size, giving females a direct reproductive advantage from increased size (Mueller 1985; Zwaan et al. 1995; Houle and Rowe 2003). This size advantage for females is thought to out-weigh the possible size advantage for males during intrasexual competition (Darwin 1874). *D. melanogaster* is a polygynous insect species with pronounced female biased SSD, and is frequently used in the study of size-related fitness effects. The size of *D. melanogaster* is typically measured by proxy from traits that are highly correlated with overall body size, such as thorax length, wing length, dry weight, and wing area (e.g. Cavicchi et al. 1985; David et al. 2003). For example, measurements of thorax and wing lengths, under standardized conditions, have been used to show the

scale of SSD in *D. melanogaster*. On average, female wing and thorax length are 15-16% longer than those of males (David et al. 2003). Although males are smaller than females, the exact effect of size on male reproductive success is unclear. Most empirical evidence suggests that larger males have greater mating success, and longevity relative to smaller males (Partridge and Farquhar 1983), suggesting that directional selection in favour of increasing size is typical of both sexes. However, much of this advantage is not attributed to size itself, but instead to size-associated traits. Larger males deliver more frequent and louder courtship song to virgin females, and move more frequently when not courting, increasing their chances of finding females (Ewing 1964). Larger males also move faster when tracking females, and are more likely to induce a female to re-mate for a second time (Partridge et al. 1987; Pitnick 1991). However, much of this evidence for a large-male advantage comes from experiments on phenotypic variation in outbred populations for which size has been environmentally manipulated. This ignores both the genetic contribution to male size, and the effects of traits that are genetically correlated with male size. There is plenty of evidence to support the hypothesis that the genetic component of variation in *D. melanogaster* body size is highly polygenic (e.g. Partridge and Fowler 1992), and that size is genetically correlated between the sexes. However, the genetic component of size, and the effect of this on male fitness is poorly understood for *D. melanogaster*.

One can investigate the theoretical optimal phenotype for males by manipulating evolution such that it is male-specific. Recent work using *D. melanogaster* has limited whole genomic haplotypes (haploid complement of one of each pair of chromosomes X, II and III) to expression in males for over 80 generations. This male-limited evolution is facilitated by the absence of molecular recombination in male *D. melanogaster,* and the use of clone-generator females with chromosomal constructs that eliminate recombination in females. Male-limited evolution totally removes selection towards trait values that exclusively benefit females, allowing the evolution of an extreme male-specific optimal phenotype (Prasad et al. 2007; Abbott et al. 2010). Measurements on evolved lines show that this masculinised phenotype has smaller wing size and shows an increase in developmental stability relative to a control population. Furthermore, Abbott et al. (2010) showed that male-limited lines, which had evolved to be smaller, also had greater reproductive performance relative to this control population. These small high fitness males evolved under optimal growth conditions, suggesting that these males acquired some fitness advantage from reduced size, even when food for growth

was abundant. These findings suggest that in this experiment, the overall direction of selection on males favored a reduction of wing size relative to the mean wing size of the ancestral population (Chippindale et al. 2001). The results imply that when subject to normal selection pressures males are displaced from their optimal wing size.

To obtain a full understanding of male size and its fitness consequences for insect species, one must consider how the fitness effects associated with size are affected by environmental variation. Experiments that show a large-male advantage through environmental manipulation of size fail to account for the genetic component of phenotypic size (e.g. Partridge 1983; Partridge et al. 1987; Pitnick 1991). The size generated by a particular genotype may have different fitness consequences under different environmental conditions. Previous experiments have considered how the genetic component of size is affected by environmental context, and how this impacts on male fitness. Under specific environmental conditions some studies of insect species have reported a small-male advantage. For example, studies across a range of spider taxa reveal that smaller male size is beneficial in species that must climb up to high habitats to reach females (Moya-Larano et al. 2007b). In other species, different environmental conditions, such as temperature, have been shown to mediate the degree of selection on male size. For example, experiments on the seed beetle (*S. limbatus*), show an increase in small male advantage at lower temperatures (Moya-Larano et al. 2007a). The theory behind a small-male advantage relies on the presence of intense scramble competition. Under these conditions it is predicted that that small males can gain a reproductive advantage from being more mobile and requiring less food, thus maximizing the efficiency and time available for mate finding. However, it is important to recognize that the weight of current empirical evidence suggests that these examples of small-male advantage are the exception rather than the rule, even under archetypal scramble competition.

A large proportion of experiments on insect species that measure how fitness is affected by size, have used wing size as a proxy for overall size. As above, wing size serves as a very accurate predictor of overall size (Cavicchi et al. 1985). Variation in wing size is intrinsically linked to variation in wing shape. By proxy, experiments that test the fitness affects of wing size, are also measuring the fitness affects associated with changes in wing shape. In *D. melanogaster,* when wing size is varied by environmental manipulation, we find strong allometric variation with wing shape (Weber 1990).

However, experiments on the genetic basis of variation in wing size and shape, suggest that wing shape is less sensitive than wing size to small environmental alterations, such as a small change in food availability (Breuker et al. 2006). Within the genus *Drosophila*, wing shape is known for its evolutionary conservation across species (e.g. Houle and Rowe 2003). Speculative arguments suggest that the relative stability, and conservation of wing shape is likely to be the result of strong stabilizing selection and/or developmental constraints (Debat et al. 2009). However, there is a significant amount of extant genetic variation for wing shape as studies have shown a strong wing shape response to artificial selection (Weber 1990; Houle and Rowe 2003). Plus, experiments on the dynamics of potential wing shape variation have found little evidence of any absolute developmental constraints on wing shape (Mezey and Houle 2005). Collectively, these experiments suggest that variation for wing shape is highly stabilized within populations, but also unconstrained by development, with substantial genetic variability. One way these contradictory results could be reconciled is if wing shape were subject to divergent selection between the sexes. Specifically, this would require particular variations of wing shape to have opposite fitness effects when expressed in males and females. Then as one sex is moved closer to its optimal wing shape, the other may get shifted further from its optimum shape, creating an intragenomic tug of war. This sexual antagonism (SA) can inhibit the adaptation of each sex towards their optimal phenotype. In addition to this, the allometric relationship between wing size and shape, means that divergent selection on wing size between the sexes, can also directly vary the selective pressures on wing shape, and vice versa. The interaction between directional selection on wing size and shape could lead to an evolutionary deadlock, which would explain the stagnant variation for wing shape, and its high conservation across species. In parallel, the persistence of substantial genetic variation for wing shape would be explained by the capacity of SA to maintain genetic variation (Gavrilets and Rice 2006).

Abbott et al.'s (2010) experiments on male-limited evolution reported changes for wing shape as a direct result of male-limited evolution, suggesting that male wing shape is under divergent selection between the sexes, and that variation for male wing shape cannot reach its optimal trait value when under selection in both sexes. Based on functionality arguments, we can speculate that male wing shape might be important for mating, due to the role of wings in generating the male mating song (Ewing and Bennet

1968). However, direct evidence for the effects of wing shape variation on male fitness is very limited, and additional quantitative information is needed.

In *D. melanogaster*, the morphological traits wing size and wing shape have been shown to significantly affect male fitness (Partridge 1983; Abbott et al. 2010). To our knowledge, the relative contribution of wing shape and wing size towards male fitness remains untested. This interaction is especially important given the strong allometric variation between these traits. The potential for divergent selection on variation for wing size and wing shape, both within and between the sexes, provides scope for a complex array of interacting selection gradients. In addition, the interaction of genetic and environmental factors contributing to phenotypic variation of wing morphology is poorly understood. Previous experiments suggest that larger male wing size is advantageous when size is determined by environmental manipulation. However, genetic effects on wing size are often disregarded. Recent evidence suggests that males can benefit from genetically small wing size and a more masculinized wing shape (Abbott et al. 2010). As we observe significant variation, both within and between these wing traits, across different environments (Webber 1990), they serve as obvious candidates for the maintenance of fitness variation.

In this study, we examine the relationship between male wing size, wing shape, and reproductive success. Our approach is to measure wing size and shape for male flies reared under different levels of nutritional stress, so as to environmentally manipulate adult size. This allows us to estimate how the interaction between wing size and wing shape might affect overall selection on these traits under a range of environments (see Appendix 1). To maximize the range of genetic fitness variation captured by our experiment, we used specific genomic haplotypes chosen on the basis of their pattern of fitness variation in males from a previous study. Experimental genomes (herein referred to as target genomes (TG)) were derived from a population-wide study of SA fitness variation in *D. melanogaster* (Innocenti and Morrow 2010). We used nine of these genomes, spanning a range of antagonistic fitness relationships between the sexes; 3 with high male / low female fitness, 3 with low male / high female fitness, and 3 with intermediate fitness for both sexes. This spectrum of fitness classes is useful because it allows us to measure wing morphology that is associated with extreme high and low male fitness. Given our focus on the relationship between male fitness and wing morphology, using these extreme male fitness classes will allow us to capture the most

extreme ends of variation for male wing morphology that are associated with fitness. Each TG is reared at three larval densities (high, intermediate, and low), where intermediate denotes the usual larval density (LD) of the rearing conditions for each generation of the experimental population from which the TGs have been derived. We measured the fitness of male flies carrying TGs when reared at each of the three LDs to provide data on the fitness consequences of LD change for each TG. The measure of relative fitness used for TG males is their ability to acquire a single first mating with a test female when in competition with a standard competitor male, that has been reared at intermediate LD. These mating trials constituted direct estimates of traits contributing to a male's ability to acquire matings. The fitness measures also provide binary (win / lose) outcomes for each individual male, providing greater analytical power than if fitness were grouped across many individuals. As TG males from each mating trial were subsequently measured for wing size and wing shape, we obtain a unique data set of individual male fitness scores across three LDs, and evaluate their association with two specific measures of wing morphology, size and shape.

We aim to identify the relative contribution of wing size and shape to male fitness across a range of LDs designed to manipulate adult size. We expect the allometric variation between wing size and shape, caused by the environmental manipulation of wing size (Weber 1990), will cause the relative proportion of fitness variation, captured by these traits, to vary between LDs. We also estimate the direction of selection on wing size and shape in males. We anticipate some selection against larger male wings for *D. melanogaster,* because of the clear SSD for smaller males, and recent evidence of fitness gains associated with smaller male wing size (Abbott et al. 2010). We predict that wing shape changes that are associated with changes in wing size may contribute to this selection against bigger wings, as males may be unable to achieve their optimal wing shape, and optimal wing size, simultaneously.

3.2 Methods and Materials

3.2.1 Experimental populations

The base population of *D. melanogaster* (LHm) has been maintained as a large (N > 1750 adults per generation), outbred population, with a 1:1 sex ratio for more than 450 non-overlapping generations. Each generation is sired by 56 groups of 16 males and 16 females. Within these groups, individuals compete for matings during a 48hr period in vials containing 10mg of dry yeast powder sprinkled onto 10ml of agar-cornmeal-molasses media (competition phase). Subsequently, all flies are transferred to new vials with the same culture media, but now without dry yeast powder, for 18hrs (oviposition phase). The flies are then discarded, and the eggs laid during the oviposition phase are reared at a standardised density of ~175 eggs per vial. This density is achieved by scraping excess eggs, using a spatula, off the media surface in each oviposition vial. The remaining eggs complete their development (growth phase) to produce the next generation of adults, from which the 56 groups are collected to repeat the adult competition phase for the next generation. Sorting of the flies to form the 56 groups is conducted using CO_2 anesthesia, and all phases are temperature controlled at 25^{o}C.

The competitor strain (LHm-bw) that is used to measure relative fitness of specific genomes has been maintained as a population with an identical rearing regime to that (described above) applied to the LHm base population. LHm-bw is homozygous for a recessive brown eye mutation (bw), located on chromosome 2, that has been placed into an LHm genetic background.

Each sampling unit, for the nine genomes we tested, consisted of a complete genomic haplotype of chromosomes X, II, and III. To sample, store, and amplify these complete genomic haplotypes, we used a clone generator (CG) stock (Thattai and van Oudenaarden 2004). The CG genotype (compound X, Y chromosome, homozygous viable translocation of chromosomes II and III) ensured transmission of the paternal X chromosome to sons and co-segregation of the paternal 2nd and 3rd chromosomes, allowing the production of numerous males carrying identical combinations of chromosomes X, II, and III. Each complete haplotype of these three chromosomes is referred to as a target genome (TG). We note that the tiny 4th chromosome is ignored. Due to the absence of recombination in male *D. melanogaster*, the CG stock can also be

used to maintain and amplify TG's by crossing male TG carriers to CG females. When TGs are expressed in a CG background they will be denoted TG-CG.

To express TGs in a male LHm background, we used an LHm stock (DxLHm) bearing a compound X chromosome [C(1)DX], a Y chromosome and autosomes of LHm origin. The compound X allowed paternal transmission of an entire TG (chromosomes X, II, and III), from father to son. Prior to use in the current study, we reared a sample of ~ 1000 individuals from the DxLHm stock at a fixed low larval-density (50 larvae per vial), for two generations. This aimed to minimize variation due to any maternally transmitted effects of larval density, which could be passed onto our experimental generation of TGs.

3.2.2 Origin and selection of experimental genomes

For this study we used 9 hemiclonal lines with distinct, sexually antagonistic, fitness patterns. These TGs were derived from a much larger study that tested the fitness of 100, randomly sampled, LHm haploid genomes in males and females (Innocenti and Morrow, 2010). Male fitness was measured under conditions that simulated the competition phase of the LHm rearing regime and calculated as the proportion of progeny sired by each genome, following 48hrs of scramble competition at a 1:1 sex ratio. Their results supported an analogous study using the LHm population (Chippindale et al. 2001), in showing a negative relationship between male and female relative fitness (Fig. 3.S1, Supporting material), intersexual genetic correlation = -0.52, 95% C.I. = -0.86; -0.1). Our TGs included each extreme and the middle of this negative cline between male and female adult fitness. Three genomes exhibited high female / low male fitness (Female Benefit - FB), three exhibited high male / low female fitness (Male Benefit - MB) and three exhibited an intermediate level of fitness for both sexes (Neutral – N). Each trio of genomes is collectively referred to as a fitness-class (MB, FB, and N).

3.2.3 Amplification and expression of TGs

Amplification of our 9 TG's was required to meet the logistical demands of the experiment and to generate the required number of male flies carrying each of the TGs. Female numbers in amplification and TG expression crosses were estimated based on each female laying a minimum of 15 eggs in 18hrs after at least one mating. Mating

groups of ~ 5 TG-CG males with 8-10 CG females provided sufficient TG-CG males (> 40) for the expression phase.

To measure the effect of specific fitness components, we first had to express our TGs in an LHm genetic background, randomly sampled from the base population. For male expression we mated TG-CG males with DxLHm females. 50% of the emerging male offspring from this cross inherit their father's full TG, with a random combination of paternal LHm genome, giving a wild-type eye phenotype. The remaining 50% of the emerging males carried the TG X, and the translocated autosomes II-III, giving a brown eyed phenotype.

3.2.4 Larval density manipulation

For each genome, larval density (LD) was controlled by manipulating the number of larvae that grew up in a fixed volume of media (10ml of agar-cornmeal-molasses media). Matings for the expression of TG's in a male LHm genetic background were carried out in chambers supplied with plates of grape juice media. The eggs from these crosses were incubated until first instar larvae hatched. Groups of these first instar larvae were transferred into vials to create treatments of low, intermediate, and high LD. For this male expression cross, we observed a high proportion of death between the first instar stage and eclosion (~ 30% - see Appendix 1), because the compound X chromosome creates an XXX chromosomal haplotype when complemented with a normal X chromosome. To compensate for this loss, we transferred batches of 60, 240 and 400 first instar larvae, to create an appropriate range of low, intermediate, and high LD treatments for male expression.

For each mating competition, TG males reared at all three LDs were matched with a standardized LHm-bw competitor male. All competitors were reared at intermediate LD, which best matches the conditions of the LHm rearing regime. In synchrony with the TG expression crosses, we set up the LHm-bw competitors, to provide virgin males and females for the subsequent mating competitions. 60 vials of 16 males and 16 females from the LHm-bw population were tossed on every day, for 5 days, to ensure virgin flies were available to compete against all density treatments. Each day the eggs which had been laid up were density controlled by removing excess eggs by eye, to provide ~ 175 emerging adults per vial, as per the LHm rearing regime. This ensured that the competitors were of a standardized size, so as not to bias mating competitions.

3.2.5 Mating Competitions

We measured male fitness in one-on-one competitions, where two males competed for the first mating with a virgin female. For each competition, one male carrying a TG (LHm-TG) was pitched against a standard LHm-bw male to compete for mating with a LHm-bw female. All flies were virgin and ~24hrs old at the time of the experiment. Flies were allowed to interact for 3 hrs. This usually gave sufficient time for one mating to occur, but not enough for a second mating. After 3 hours of scramble competition between the TG male and competitor male, the female was separated into an individual vial and allowed to lay eggs for 48 hours, before being discarded. Paternity was assessed by scoring the phenotype of these offspring, upon eclosion. The recessive bw marker allows for quick scoring of eye colour. Broods with wild-type eyes were sired by the LHm-TG male, and broods with brown eyes were sired by the LHm-bw competitor. The males from each competition were kept together and frozen for subsequent wing mounting. A small proportion of competitions produced no offspring (null = 6.8%), and a surprisingly high number produced mixed broods (mixed = 20%). These were all excluded from fitness analysis and subsequent wing mounting.

3.2.6 Wing morphology

The two males (one TG-LHm, and one LHm-bw) from every mating competition that produced a unanimous winner were scored for wing size and shape. Wings were mounted by hand onto glass microscope slides, and sealed with glued cover slips. All wings were visualised using a compound microscope (objective lens 5x), attached to a video camera. The program Velocity Acquisition was used to capture and label photographs of the wings. Individuals with damaged wings were discarded from the analysis. Subsequent digitization of landmarks (LM) was carried out using the programs tpsUtil and tpsDig (Rohlf 2010). Eleven LMs were marked on each wing, in the order and location depicted in figure 3.S2. The digitization of LMs was done in a randomized order with respect to genome and density treatment, to avoid systematic bias caused by improved or variable user technique over time.

The morphometric analyses of wings were performed in R - version 1.40-devel, 64-bit, 2011 (R Development Core Team 2006). The final matrix of LM coordinate values was produced by flipping the orientation of the raw left wing LM values onto the raw right wing LM values, whilst maintaining the relative location of each LM. This ensures that

subsequent processing of LMs recognizes each of the 11 LMs from left and right wings relative to one another. Generalized procrustes analysis (GPA) was used to remove the effects of variation in translation, rotation, and scale of the LM locations between different male wings. GPA uses least squares to align wings and effectively minimise the distance between the corresponding LMs of different wings while maintaining their shape unaltered. The point of alignment was the centroid, the theoretical centre of the wing that minimises centroid size, the sum of the squared distances between the centroid and the landmarks. Scaling during GPA standardised all wings to a centroid size of one. The scaling factor, i.e., the original centroid size, was a measure of wing size, (Bookstein FL 1991; Rohlf 1999). In our analyses, we used the average centroid size of left and right wing as a measure of a male's wing size.

The data that provided the foundation of the wing shape analysis consisted of the full matrix of left and right superimposed LM coordinates. Mean LM scores were subsequently calculated from these superimposed left and right wing coordinates. This matrix of mean superimposed LM scores constituted our estimates of wing shape and were subjected to further statistical analysis. Given that changes in shape will cause correlated changes in the coordinates of landmarks, we used principal component analysis (PCA) to transform superimposed landmark coordinates into a series of linearly uncorrelated variables. With 11 landmarks and two dimensions, PCA produced 22 axes, each of which captured an independent aspect of shape variation among the wings analysed. The wing PC scores could then be analysed in order to elucidate how shape varied across genomes and density treatments as well as how it affected a male's fitness.

3.2.7 Statistical analysis

As above, to generate informative fitness data, we first discarded all repeats that produced ‘mixed’ and ‘null’ results. These were vials in which a double mating or no mating had occurred, respectively. Each of the remaining 847 mating trials corresponded to a binary response of ‘TG win’, or ‘TG lose’, plus a particular genome and LD. For each genome, a mean fitness value was also calculated at each LD, as the proportion of TG win relative to TG lose.

Analysis of wing morphology used a filtered data set for which any mating trials without a full complement of the measures, male fitness, TG wing size, competitor wing

size, and TG shape (22 PC axes), were removed. This data set consists of 570 repeats; 153 low LD, 218 intermediate LD, and 199 high LD.

The fitness effects of wing size and shape were analysed using a binomial generalised linear model (GLM) with logit link function. Statistical significance of model terms was analysed using chi-squared tests. To examine the fitness effects of wing size the dependent variable 'fitness' (win / lose) was tested for variation with the covariate competitor wing size, and the terms TG wing size, genome, LD, and the interaction genome-by-LD. The inclusion of competitor wing size as the first covariate in this model allowed us to effectively measure how relative TG size affects male fitness. The effect of wing shape was analysed by adding PC axes as linear predictors to the same model, after the competitor and TG size terms (see Table 3.5). For testing fitness variation with wing shape, across all LDs, we ran the model both with and without the TG wing size term (Table 3.5). Theoretically, when the TG wing size term was included before the PC axes, any fitness variation that was explained by wing size was accounted for within the model, before the effects of each PC axis were measured within the same model. This accounted for fitness differences generated by wing shape change, beyond shape variation that was the result of allometric variation with changes in wing size.

The binomial GLMs for analysis of wing shape included a large number of linear predictors, the 22 PC axes. In order to avoid over-fitting, these models were simplified using the 'step' function in R. This function uses the Akaike Information Criterion (AIC) as a measure of model fit and explores alternative models by adding and removing terms (combining forward and backward procedures).

Overall variation in TG wing shape was investigated using multivariate analysis of variance (MANOVA) on the whole matrix of PCs derived from the superimposed LM data. We tested for variation in shape with the covariate TG wing size, and independent variables LD, genome and their interaction. Overall statistical significance of each term in the MANOVA was analysed using a Pillai test. In addition, we ran 22 separate ANOVA tests with the same model but using each PC axis as the dependent variable. This established exactly which PC axis showed significant variation with each term.

Throughout our analysis we modelled the genome term as a fixed effect because we were working with a distinctly non-random sample of genomes. Each genome was chosen because it represented a particular position across the range of male fitness variation (high, intermediate, or low) for the LHm population. Hence we did not expect the phenotypic effects of each genome to follow a normal distribution, as would usually be assumed.

3.3 Results

3.3.1 Male Fitness

We first tested how LD and genome affected male fitness using a binomial GLM on data from all mating trials that produced a clear winner. This model used the binary dependent variable male fitness (win / lose), with three variables: LD, genome, and their interaction. This revealed highly significant effects of LD, genome, and the interaction term (Table 3.1A). To establish which of the three LDs differed significantly in their fitness effects we sub-divided this same analysis into pair-wise comparisons between LDs: low versus intermediate, intermediate versus high, and low versus high (Table 3.1B). This revealed that both LD and genome significantly affect male fitness between all three levels of LD treatment. Mean male fitness decreased as LD increased (low, 0.706; intermediate, 0.601; high, 0.404). We also found very highly significant genome effects on male fitness across all pair-wise models, and very highly significant genome-by-density interactions between the comparisons of low versus intermediate LDs, and low versus high LDs. However, we found that this interaction term showed non-significant fitness affects between the intermediate versus high LDs (Fig. 3.3). This lack of genetic fitness variation in the step-up from the intermediate to high LD suggested that genetic determinants of fitness were dominated by the environmental affects of the high LD, creating this synchronous response between genomes as LD rose from intermediate to high treatments.

We examined how our measures of male pre-copulatory fitness compared with overall fitness measures on the same nine genomes reported previously by Innocenti and Morrow (2010). We correlated our mean male fitness values for each genome at each LD with Morrow's mean net fitness measures, using a Spearman's rank correlation. Performing this analysis showed significant positive correlations at the intermediate LD (rho = 0.766, n = 9, P = 0.021) and high LD (rho = 0.883, n = 9, P = 0.003), and an almost significant positive correlation at the low LD (low, rho = 0.65, n = 9, P = 0.066). These results were evidence that our one-on-one competition results, at each larval density, exhibited a similar ranking of genomes by their relative fitness to that reported by Innocenti and Morrow (2010). This suggested that our measures of pre-copulatory male fitness were reliable indicators of overall male fitness across all LDs.

All subsequent analysis included morphological measures of wing size, wing shape, or both. Herein, all analysis (sections 3.3.2 and 3.3.3) used a filtered data set, for which any mating trials without a full complement of fitness, wing morphology measures were removed. The filtered data set had a sample size of 570 (153 low LD, 218 intermediate LD, and 199 high LD). We first assess components of wing morphology alone, looking at effects of genome and LD on wing size and wing shape. These analyses were then combined to investigate the allometric relationship between these traits. Finally, we analysed the effects of wing morphology on male fitness across LDs, and estimated the direction of selection on male wing shape under normal rearing conditions.

3.3.2 Wing Morphology

To test the effects of LD and genome on male wing size, we used a two-way ANOVA of the dependent variable TG wing size with the variables LD, genome, and their interaction. We found that wing size varied very significantly between the three LDs (Table 3.2), with wing size decreasing from low to high LDs; corresponding to mean centroid scores (± SEM); low = 1014 ± 27.9, intermediate = 936 ± 30.1, and high = 895 ± 35.7 (Fig. 3.1). *Post hoc* Tukey tests revealed that these wing size differences described by the ANOVA were significant between all pair-wise combinations of the three LDs (Table 3.2). The ANOVA also revealed very significant variation in male wing size between genomes, and in the interaction of larval density-by-genome (Table 3.2). This demonstrated that the genomes varied in the degree to which LD affected male adult wing size (Fig. 3.2).

The wing size response of genomes to changes in LD appeared to be related to the fitness effect of each genome (Fig. 3.2). This pattern was especially pronounced between the MB and FB fitness classes. FB genomes seemed to be more sensitive to changes in LD and showed a decrease in wing size between low and high LDs, which was almost 1.5-fold that observed in MB genomes (a reduction of 13.03% for FB and of 9.09% for MB). To investigate the relationship between fitness class and this phenotypic robustness for size, we used the mean wing size of each genome at each LD (3 LDs x 9 genomes = 27). We performed a two-way ANOVA of the dependent variable mean wing size, with the variables LD, fitness class, and their interaction. This analysis revealed very highly significant variation for wing size between LDs, but no significant variation was observed with either fitness class or the interaction of LD-by-fitness class (Table 3.3A). For all genomes mean wing size decreased as LD was

increased across the three LD treatments. We subsequently sub-divided this analysis into three pair-wise ANOVAs between LDs (low versus intermediate, intermediate versus high, and low versus high). This only revealed significant wing size variation associated with fitness class between the comparison of intermediate and high LDs (Table 3.3B). All three pair-wise analyses also revealed very highly significant effects of LD on mean wing size, but non-significant affects for the interaction term LD-by-fitness class.

We further investigated the relationship between the phenotypic robustness of male wing size and male fitness, by correlating the size differential of each genome (mean wing size at low LD, minus mean wing size at high LD) with the mean relative male fitness, across the three LDs (Fig. 3.4). This revealed a significant negative correlation (rho = -0.70, n = 9, P = 0.043). The plot of these two correlates showed the points neatly aligned along a negative axis, except for one outlier, suggesting that there was a substantial degree of dependence between phenotypic robustness for size, and male fitness.

We investigated the relationship between overall wing shape and wing size using MANOVA. Here, the term 'overall wing shape' refers to the whole matrix of PC axes derived from the superimposed LM data (see Methods and Materials). The basic model tested each of the 22 PC axes as the dependent variable with the variables, TG wing size, genome, LD and all their respective interactions. This analysis revealed significant overall shape variation with all terms in the model (Table 3.4A). To establish the patterning of these effects across each level of LD, we sub-divided the analysis into three pair-wise MANOVA comparing the group of 22 PC axes between pairs of LDs (low versus intermediate, intermediate versus high, and low versus high). This revealed significant variation for all terms in all three pair-wise models (Table 3.4B). Collectively, these results showed that wing shape was highly variable between genomes and between LDs, and there was a significant interaction between genome and LDs on wing shape. Importantly, we also observed very significant allometric variation between wing size and wing shape, so when wing size varied between LDs we saw a corresponding change in wing shape. Much of the wing shape variation attributed to LD was likely the result of allometry with the size variation created by different LDs. We further investigated overall wing shape variation within each fitness class using three separate MANOVA tests for MB, FB, and N genomes, respectively (Table 3.4C).

Again we found highly significant wing shape variation with all terms for all three fitness classes, except for one non-significant interaction of genome-by-LD for the MB fitness class. This showed that there was a degree of consistency between the wing shape change of MB genomes across LDs. This suggested that MB genomes had a relatively synchronous pattern of allometric wing shape variation, whereas males expressing FB and N genomes were relatively non-synchronous in their allometry.

To increase the resolution of our analysis of wing shape variation, we ran 22 separate ANOVA tests, each using one of the PC axes as the dependent variable, with the variables, TG wing size, LD, and their interaction. The combinations of variables that showed significant variation with each of the PC axes are laid out in Table 3.5. Here we found that significant wing shape variation is most frequent with the variables LD and genome, which showed significant variation for 16 and 13 of the 22 PC axes respectively. Wing shape varied with the interaction term, genome-by-LD for 8 PC axes, and with wing size for 7 PC axes. These results suggested that components of wing shape variation, which result from the effects of LD and genome were frequently independent of wing size effects. However, for the PC axes that we observed significant effects of wing size, we always observed a corresponding effect of LD, suggesting that allometric variation was intrinsically linked to the effects of LD treatment.

3.3.3 Fitness effects of wing morphology

We investigated the effect of male wing size on reproductive success, using a binomial GLM. We modeled the binary dependent variable male fitness (win / lose) with the covariate competitor wing size, and the variables TG wing size, genome and LD. The initial analysis (Table 3.6A) included data from all three LDs, and revealed significant fitness variation with both TG wing size and competitor wing size. The opposing signs of the coefficient values associated with TG wing size (+ 0.021) and competitor wing size (- 0.007) demonstrated that overall relatively larger TG males had a significant mating advantage in one-on-one competitions. In addition to this size effect, we observed significant fitness effects of both LD and genome, but a non-significant effect of the interaction of genome-by-LD. This suggested that LD and genome still affected TG fitness, even after accounting for the fitness effects of both competitor and TG wing size. Specifically the fitness effects of LD, after accounting for the effect of TG wing

size, showed that size independent affects of LD manipulation also affected male fitness.

To establish how the transition between each LD affected fitness, we divided the same binomial GLM into a pair-wise comparison of low versus intermediate and intermediate versus high LDs (Table 3.6B). There was no effect of TG wing size on male fitness across the low and intermediate LDs. However, there were significant effects of LD and genome. The results across the intermediate and high LDs revealed an almost opposite pattern. Here, TG wing size had a significant effect on fitness, while the effect of the LD was non-significant. These results suggested that fitness differences between the low and intermediate LD were independent of TG wing size, but still affected by LD. Conversely, fitness differences between intermediate and high LDs were a product of wing size effects, which appeared independent of LD effects.

To investigate the effect of wing shape on male fitness, we extended the binomial GLM used above to measure the fitness effects of wing size, to incorporate all PC axes. We used them as independent variables along with competitor wing size, TG wing size, LD and genome. This full model was simplified using sequential AIC analysis to find the most appropriate final model for subsequent statistical analysis (see Methods and Materials, section 3.2.7). These simplified final models were used for every binomial GLM that includes PC axes as independent variables. Each final model is shown in the results tables of the corresponding binomial GLM (Table 3.7/3.8). The first binomial GLM tested the dependent variable Male fitness, with the covariate competitor size, and the variables TG wing size, LD and genome. This revealed significant fitness variation with three PC axes (PC6, PC8, and PC19) (Table 3.7A). When the TG wing size term was removed from this model, so that the effect of wing size was ignored, there was significant variation for male fitness with two of the same PC axes (PC6, and PC19), but non-significant variation with PC8, which was actually excluded from the final model by the AIC analysis (Table 3.7B). Here we observed substantial similarity between the fitness effect of wing shape, both with (Table 3.7A) and without (Table 3.7B) accounting for TG wing size effects. By showing that there are no extra effects of wing shape on male fitness, in the absence of wing size effects, these results demonstrated that allometric variation of shape with changes in wing size, did not contribute substantially to the fitness effects of wing shape. Therefore, although we did observe substantial allometric variation in wing shape with changes in wing size across

LDs, we did not find that this wing size effect impacted on the overall fitness effects of wing shape.

However, the inclusion of all three LDs in our model could have distorted the true relationship of wing size and wing shape with male fitness. The three LDs generated significantly different non-overlapping wing size groups (Table 3.2), which were largely determined by environmental, as opposed to genetic effects. As we observed significant wing shape allometry with environmentally induced variation of wing size, the observed fitness effects of wing shape could have been dominated by effects of LD on TG wing size, and/or other effects of LD manipulation. Furthermore, across the three LDs we observed distorted fitness effects of size, where the majority of size related fitness effects occurred between just the intermediate and high LDs.

To ameliorate these effects, caused by the LD treatments, and to specifically account for genetic effects, we ran three separate binomial GLM models, within each LD. Within each LD, variation for wing morphology came from genetic differences between genomes, and developmental/micro-environmental variation and was not artificially exaggerated by experimental manipulation of larval competition for food. Our three separate binomial GLMs tested for male fitness variation with the covariate competitor wing size, and the variables TG wing size, multiple PC axes, and genome. Again each model used for statistical analysis was the 'final' model, generated by sequential simplification.

Within the low LD, male fitness was unaffected by either TG wing size or competitor wing size (Table 3.8A). In generating the final model we excluded the TG wing size term completely, because TG wing size did not significantly affect fitness within this LD. However, there were significant wing shape effects on male fitness for two of the ten PC axes that were retained in the simplified final model (PC17 and PC19), which collectively accounted for 16.02% of the fitness variation. At the intermediate LD, male fitness was significantly affected by TG wing size but the effect of competitor wing size remained non-significant (Table 3.8B). Also wing shape significantly affected fitness for three of the five PC axes in the final model (PC6, PC8, and PC11). At this intermediate LD the proportion of fitness variation explained by these shape components (43.56%) was more than double the proportion of fitness variation explained by wing size (19.62%). Finally, within the high LD TG wing size was

expected to be substantially smaller, relative to wing sizes generated under the normal rearing conditions of the ancestral population (LHm). Here we found significant fitness effects associated with competitor wing size, TG wing size, and just one shape component, PC19 (Table 3.8C). However, at the high LD the proportion of fitness variation explained by wing size (32.94%) exceeded that explained by wing shape (10.10%), the opposite pattern to that observed within the intermediate larval density. The influence of TG wing size for male fitness appeared to increase as larval density rose, whereas wing shape appeared to influence male fitness the most within the intermediate LD.

Beyond these morphological effects, each model was also associated with genomic differences, which significantly contributed to fitness variation within all LDs (Table 3.8 A/B/C), demonstrating that there were traits which were independent of wing size or shape that were important for male reproductive success. Importantly, we found that the proportion of explained fitness variation was roughly constant across the three LDs. This suggested that the mating success was not more random at any particular LD. However, between LDs the relative proportion of fitness variation that was explained by wing shape, wing size, and other genomic effects varied substantially. These results suggested that the relative importance of wing size and wing shape for male fitness varied with changes in the amount of larval competition for food.

We have detected fitness variation with male wing shape. To understand the mechanisms of where and how wing shape changes affect male fitness we estimated the direction of selection on each of the 11 wing LMs, using the PC axes that significantly affected male fitness within the intermediate LD. Here we focussed on the intermediate LD to specifically measure selection under the rearing conditions which best match, those that the LHm population was adapted for. Furthermore, using just one LD minimized the effect of wing shape distortions that occurred across larval LDs. First, a new PCA was applied to the subset of raw LM data from the intermediate LD. A simplified binomial GLM, that included the variable TG wing size as the first independent variable, was used to establish which PC axes significantly affected male fitness within this new matrix of PC axes. For each of the 11 wing LMs we estimated the scale and direction of LM movement. This was derived from the loadings values that corresponded to the PC axes with significant effects on fitness (PC5, 8 10, and 12). The 22 loadings values for each of these four PC axes were weighted by the their

respective linear coefficients, estimated by the binomial GLM. These new transformed loadings values, were then summed for each of the 22 LM coordinate axes to generate an estimate of the scale and direction of selection for each LM. We plotted arrows for each LM representing the direction and relative scale of selection based on these values (Fig. 3.5). Here we found the majority of LMs that showed a large degree of movement were located at the most distal parts of the wing (ie. LM 8, 9, and 11). The location of LMs 8 and 9 were strongly selected to move further apart, whilst LM 11 was selected to move inwards, in the direction of LM 8. The scale of selection on the more central and proximal LMs was on average substantially smaller (i.e. LM 1, 2, 3, 4, and 5).

We subsequently looked at how shape changes caused by allometric effects of wing size compared with the direction of selection on male wing shape. We estimated the change in wing shape that resulted from wing size variation across all LDs. We estimated this shape change using the loadings values that corresponded to PC axes that showed significant variation with wing size (Table 3.5) (PC axes 1, 2, 4, 6, 7, 13, and 17). The loadings scores for each PC axis were weighted by their respective linear coefficients as estimated in each of the ANOVA models of shape on size (Table 3.5), to account for the directionality of shape change. These values were then summed to estimate the overall direction of wing shape variation across our three LDs, and plotted on figure 3.5. We observed that the direction of selection on male wing shape for fitness frequently differed from the direction of wing shape change that resulted from increasing wing size. The shape change associated with increased male wing size and increased male fitness opposed each other for 8 out of the 22 X/Y coordinate axes.

3.4 Discussion

Our study was focused on male wing morphology, with the aim of understanding the affects of variation in wing morphology on reproductive success. We investigated the fitness and wing morphology of a panel of genotypes that span a wide range of male fitness values. We used experimental alteration of larval densities to manipulate the size of adults at eclosion. Here, we examine our results in the context of male-specific fitness and selection on male morphology. From our analyses we make inferences about the selection operating on male wing morphology and identify possible reasons for the maintenance of SSD for smaller males, especially in the context of environmental variation.

3.4.1 Genetic and environmental effects on wing morphology

We observe a significant decrease in male wing size as LD is increased, which is congruent with the pattern observed for the vast majority of other insect species (reviewed in Peters and Barbosa 1977). We also find that the effect of LD manipulation on adult wing size, varies between genomes, demonstrating that genetic effects play an important role in the phenotypic size response to changes in LD. These latter results contradict findings from previous studies, which show no significant genotype-by-environment interaction for body size (Prout and Barker 1989); (Santos et al. 1994). For example Santos et al. (1994) found no genome-by-LD interaction in *D. melanogaster* for thorax length between lines artificially selected for large thorax length and control lines. In particular, Santos et al. (1994) observed no variation in the rank order of genomes when they are subjected to different LDs. However, these contradictions with our findings come from experimentation on a comparison of artificially selected lines, relative to control lines. In contrast, rather than being artificially selected for larger size, our TGs are selected on the basis of their male fitness. As a result, we expect greater extant variation for size across our genomic sample than in Santos et al. (1994). In turn, we expect greater variation between the gene-environment interactions that generate the phenotypic size of our nine TGs, and hence we observe a strong genome-by-LD interaction. Interestingly, through artificially selecting larger individuals Santos et al. (1994) also provide evidence suggesting that the gene-environment interactions in relation to body size form an important component of total fitness for both sexes. Correspondingly, our finding of variation in the gene-environment interactions for the size of genomes with highly distinct fitness effects is

actually, in part, supported by this research (see below, *3.4.6 Phenotypic robustness for male size*).

Overall we find that male wing shape also varies between genomes and between different LDs. In addition, similar to the pattern for variation in wing size, we find that different genomes vary for their response in wing shape across LDs. However, it is unlikely that this variation for wing shape with LD is driven solely by effects of LD alteration. Importantly, we observe strong allometry between wing shape and wing size, as found by (Weber 1990). This relationship suggests that the variation in wing shape across different LDs is at least partially a product of the observed variation in wing size. Evidence for the contribution of LD to variation in wing shape supports our findings that it is an important determinant of wing shape variation. Even under relatively small changes in LD, which do not even exceed our intermediate LD, Bitner-Mathe & Klaczko (1999) record significant shape variation with LD, whilst other environmental variants, such as temperature, appear to have little affect on wing shape. However, even in the absence of LD alterations we still observe highly significant wing shape variation. Our analysis of shape variation within each of the larval densities reveals that genomic and allometric variation for shape persists in the absence of extreme environmental variation. This is interesting as wing shape has been shown to be relatively less variable than wing size (Breuker et al. 2006). Conversely, our results show wing shape to be highly variable between our nine TGs, both between and within different environmental treatments.

Collectively, our results show variation for wing size and wing shape between genomes, where other experiments testing the same traits have found non-significant effects. One possible explanation for these differing results is that we use a non-random complement of genomes, chosen on the basis of their extreme high and extreme low male fitness. We therefore expect that the trait values of characteristics correlated with fitness, such as wing size and wing shape, may occupy the extreme ends of population-wide standing variation, adding greater power to our model.

3.4.2 Fitness effects of wing morphology

In agreement with the majority of the published literature (e.g. Partridge et al. 1987; Pitnick 1991), we find evidence that larger males are more likely to win mating opportunities when in direct competition with a smaller competitor. However, we find

that this is not universally true across all of the environmental conditions we impose; other factors can dominate male fitness under certain conditions. Specifically, larger TG males have an overall advantage within the intermediate and high LDs, where on average the TG males are either of similar or smaller size relative to their standardized competitors. Conversely, within the low LD, where TG males are on average larger than their competitors, we observe no reproductive advantage from any increase in actual male size relative to their competitor. In other words, once a male is larger than its immediate competitor, further size increases reap no further increase in fitness. The reproductive advantage from larger male size is usually attributed to traits that are associated with large, rather than small, males such as the ability to deliver more vigorous courtship or increased speed of movement for chasing females. Furthermore, in natural populations males found *in copula* are, on average, larger than single males (Markow 1988). This suggests that our observation of large male advantage is not an artefact driven by laboratory conditions. However, if large size itself were an essential determinant of male fitness, we would not expect there to be such extensive variation for male body size within populations of *D. melanogaster* (Turner et al. 2011). Specifically, larger males may gain advantages in mating, but could pay some other costs, which affect overall fitness. For example genetically large size has been shown to reduce relative fitness when flies are forced to develop under intense larval competition (Santos et al. 1994). Or, another example relates to the fact that large size correlates positively with larval development time in holometabolous insects, i.e. it takes time to get big (Roff 1980; Fairbairn 1990). In turn this could result in larger males acquiring fewer mating opportunities because they take longer to develop.

Our evidence shows that when male size is manipulated by environmental factors then there is a range over which a male's ability to acquire mating opportunities does not vary with size, but instead with other variables such as competition for food during development. This effect is revealed when we compare fitness variation with size, between adjacent LDs (low versus intermediate, and intermediate versus high). Between low and intermediate LDs we find non-significant variation for fitness with size, even though we see significant size differences between these LDs (Table 3.1). However, we do observe significant fitness variation with both LD and genome. Importantly, between intermediate and high LDs we observe an almost opposite pattern of variation with male fitness. Here we find highly significant variation for fitness with the covariate size. But although genetic fitness differences persist between the

intermediate and high LDs, we lose the variation with LD that was found between low and intermediate LDs. These results show that fitness differences between low and intermediate LDs reflect size-independent effects of rearing conditions on the flies' performance, whereas the fitness decline from intermediate to high LD can be explained simply by the smaller size of flies emerging from more crowded cultures. Looking across all three LDs, it appears that for male *D. melanogaster* their ability to compete for matings increases as a function of their size relative to the size of their immediate competition for females. However, once males reach a certain positive size ratio with their competitors, any further size increases will not further increase reproductive success. This suggests there is a theoretical upper level size threshold for males, which varies depending on the size of their local competition for females. Here we show that relative male size, as opposed to absolute male size, is most important for reproductive success. In theory this should allow males to maximize their own fitness by being relatively larger, whilst minimizing the potential costs associated with larger size, as mentioned above. As there appears to be no selective pressure for infinitely bigger male size we expect that, within a given population, male size may result from a trade-off between the fitness benefits of being bigger than competitors and the fitness costs associated with extremely large male size.

3.4.3 Relative importance of traits affecting male fitness: wing size vs wing shape

Across all LDs we find that the overall proportion of fitness variation explained by wing shape is comparable to that explained by wing size. This suggests that, for males, wing shape constitutes an important trait for reproductive success. However, the allometry we observe between wing size and shape (Table 3.5), demonstrate that these two traits are intrinsically linked. In estimating the relative contribution of wing shape and wing size to male fitness, we minimized the effects of this allometric variation by confining our analysis to each LD separately, where the range of wing sizes is considerably smaller than across different LDs.

Within the low LD, we find size is irrelevant to male fitness whereas shape explains a small percentage of variation in mating success. At this density we expect all TG carrying flies to reach a size that is considerably bigger than that of their competitor. Under these conditions we show that absolute size has no bearing on male fitness. As above, this does not imply that male size is not important for fitness when all flies are relatively large. Instead, it reinforces our previous argument, showing that once a male

is considerably larger than its competitor, it gains very little from further increased size. However, the presence of wing shape effects on male fitness, even when most males are relatively bigger than their opponents, suggests that particular wing shape traits can still provide added value, beyond benefits associated with size. Within the low LD the effects of genomic differences explains relatively more fitness variation than at any other LD, suggesting that other factors become relatively more important, when components of wing morphology have a small effect on male fitness. Here, in the total absence of any size effects, and only small effects of wing shape, we expect other components of male fitness, such as the rate of adult locomotion (Long and Rice 2007), to determine the reproductive success of male *D. melanogaster*.

Within the intermediate LD, TG males are exposed to rearing conditions which almost exactly match those which they been evolving under for more than 350 generations. We therefore expect males to develop wing morphologies comparable to those found under their normal rearing conditions. This makes the intermediate LD the most informative picture of the extant variation of wing morphology within our experimental design. Here, male fitness does show variation with wing size, but we find the effect of wing shape explains more than double the proportion fitness variation than that of wing size. This is surprising given the perceived importance of size for male mating success (Partridge and Farquhar 1983). However, as is the case with male size, the fitness variation we observe with wing shape might, in part, be the result of correlated traits. Deciphering the exact role of wing shape in altering male fitness would require a demonstration of how particular wing shape variants can affect male fitness. One possibility is that wing shape significantly affects female preference by altering the male mating song. Certainly it is known that mating song is important for male fitness (e.g. Greenacre et al. 1993), and that certain song characteristics can significantly improve reproductive success (Ritchie et al. 1998). For example, *D. montana* females show a preference for males that produce short sound pulses with a high carrier frequency (Hoikkala et al. 1998). We can therefore theorize that particular wing shapes may generate more attractive signals, perhaps by determining the pitch of the sound generated. An investigation into the effect of wing shape on *D. melanogaster* song and the associated fitness effects, would provide an interesting foundation for future research.

Within the high LD, TG males, are on average, smaller than their respective competitor. Here, we find that the wing size of TG males and the wing size of their respective competitor, significantly affects male fitness. Within this LD, male relative size becomes an important fitness component explaining more than three times the proportion of fitness variation than is the case for wing shape. The relative importance of male size within the high LD, fits with the observation that male size dominates the effects of LD in the step-up between intermediate and high LDs (Table 3.6).

Overall, as LD changes from low to high LD we observe a progressive increase in the relative proportion of fitness variation explained by wing size. Collectively, our results show that size becomes a crucial determinant of male fitness, when a male's size is similar, or smaller than that of their immediate competition. Interestingly, Pitnick (1991) suggested that a lower level size threshold exists for males, which may be created by female discrimination between males of different sizes. Indeed some experimental evidence supports a threshold hypothesis of this type, suggesting that males over a particular size are able to induce females to re-mate more quickly than smaller males. However, there is still no direct evidence that this is an effect of female choice, as opposed to male coercion (Pitnick 1991). Our evidence certainly supports the existence of a crucial size ratio with local competition, where males suffer significant fitness costs associated with large negative size differential. Conversely, we also show that males gain significant advantages from a positive size differential with competitors, but gain nothing extra from an extremely large size differential.

In contrast to the effects of wing size, we find that the proportion of fitness variation explained by wing shape peaks at the intermediate LD, and is reduced at low and high LDs. One possible explanation for this pattern of wing shape effects comes from the allometric variation of wing shape with changes in wing size (Table 3.5). As in other studies (Weber 1990), we observe significant variation between wing shape and size across LDs. The relative contribution of wing shape towards male fitness at low and high LDs, suggests that the extreme wing size variation generated by our manipulation of larval competition may distort wing shape to a degree that ameliorates any possible benefits associated with wing shape variation, which is generated under normal rearing conditions.

3.4.4 Selection on male wing morphology

Overall, the selection we observe on wing shape suggests that males with a specific pattern of LMs 8, 9, and 11 are more successful in our one-on-one competitions (Fig. 3.5). Overall, these results are quite similar to the general pattern of shape change documented for experiments on male-limited evolution, which sampled genomes from the same base population. Abbott et al. (2010) reported the evolution of shorter and wider wings under male-limited evolution, where there is a total absence of any female-specific selection. In our study we find that the LMs subject to the greatest degree of directional selection (LMs 8, 9, and 11), show a similar pattern of movement to that which results from male-limited evolution. Specifically, LMs 8 and 9 show a strong tendency to move further apart in both studies. However, we find that LM 11 is selected to move further towards the centre of the wing, which specifically contradicts the change at this LM, which is observed in response to male-limited evolution. However, where our results do differ from those of Abbott et al. (2010), we do not necessarily show direct contradictions to their findings. The direction of selection, which we measure, and the evolved direction of the response to selection, which Abbott et al. (2010) measure do not need to be fully aligned. In particular, the response to selection will be affected by the extent and direction of genetic correlations between traits. For example, whether a wing becomes larger is dependent on which shape is selected for, and how this shape variant is genetically correlated with size. Whilst we measure the direction of selection on male wing shape, Abbott et al. (2010) measure the response of wing shape to male-specific selection. The missing link is the genetic architecture which dictates exactly how wing shape changes with variation in wing size. The measurement of this genetic architecture that dictates why and how wing morphology varies would provide an interesting foundation for a future study.

As above, we observe strong allometry across LDs (Table 3.5), showing that variation in wing size causes corresponding changes in wing shape. To estimate how male wing shape changes in response to changes in wing size, we plotted the mean direction of LM movement as wing size increases (Fig. 3.5). These results suggest that increases in male size, which result from environmental manipulation of LD, are predominantly caused by the enlargement of the more proximal area of the wing. Importantly, we observe that the wing shape change, which results from increased wing size, frequently opposes the direction of selection on wing shape (Fig. 3.5). This suggests that some beneficial shape characteristics are not achievable with relatively large wings. Specifically we see

almost opposite effects of allometry on increased size vs shape selection, on LMs 8 and 9 where we observe the greatest degree of selection on wing shape. As a result, we expect some selection against large male wings, because selection on wing shape is not fully aligned with allometric shape changes that are caused by increased male wing size. This hypothesis is supported by unpublished data collected in the Reuter laboratory using these same TGs, which showed a positive fitness effect for wing shape characters that are associated with small wing size. Furthermore, Abbott et al.'s (2010) male-limited evolution experiment, also found an overall reduction in male wing size, and a corresponding increase in fitness. We hypothesize that males face a trade-off between optimizing either wing size, or wing shape. Furthermore, our results suggest that the coefficient of this trade-off between optimal wing size and optimal wing shape should vary with environmental conditions that affect adult size. This complex relationship between the selective forces operating on wing size, wing shape, and male fitness may also explain the high level of extant genetic variation for size and shape in *D. melanogaster*.

3.4.5 Genomic effects on male size

Firstly, we note that our data set is not appropriate for quantitative genetic analysis, because of the very small (n = 9), and distinctly non-random sample of genomes. Nevertheless, because our TGs cover a wide range of male fitness variation, we can make some inferences about the genetic relationships between male fitness, and male size. Our results across all LDs show that environmental determinants of male size, which generate a size range of 283.29 centroid units, significantly affect male fitness. However, we also find that genetic determinants of male size within each LD have, at best, a weak relationship with male fitness. Although we observe size variation between genomes, this genetically determined size does not appear closely linked with male fitness. Within each LD we find non-significant genetic correlations between mean TG size and mean TG fitness. However, we do observe positive correlation coefficients within the intermediate and high LDs, where we also find male fitness does vary with TG wing size (Table 3.6). Because of the small sample size, and the non-random nature of our genomic fitness scores it is difficult to draw conclusions as to whether our lack of genetic size effects on male fitness represents the true relationship between these two variables. An analogous study (Reuter et al. unpublished data) which measured mean size for 15 non-random TG's at low larval density, found the same non-significant relationship between male size and fitness. However, more data

from a much larger sample size is required to establish how genetic determinants of wing size, affect male fitness. Here we provide evidence which suggests that size of individual males, relative to that of their competitors, is an important determinant of reproductive success. We also provide some evidence that genetic effects may play a role in maintaining a favorable size differential relative to competitors that develop under the same environmental conditions.

3.4.6 Phenotypic robustness for male size

Our results suggest that high male fitness may be associated with environmental robustness of size. Firstly, our statistical analysis shows that variation in wing size across LDs varies between genomes. In other words the phenotypic response of size to changes in LD has a variable genetic component. This result is best visualized in Fig. 3.2, where we observe that the trajectory of size change for MB genomes is less steep than that of FB genomes. This suggests that when MB genomes are expressed in males, they show greater phenotypic robustness for size, to increases in LD. Although we do not find significant differences between fitness classes for their size response to changes in LD, we do observe a significant negative correlation between size differentials (calculated as mean wing size at low LD minus mean wing size at high LD) and overall mean male fitness (Fig. 3.4). But for one outlier, we observe the fitness classes grouped along this negative cline, with MB genomes showing the lowest size differentials, then N genomes, and finally FB genomes with the highest size differentials. This significant negative relationship reinforces the argument that MB genomes, when expressed in males, have a greater ability to resist size reductions caused by intense larval competition.

If we view this result in the context of male fitness alone, our findings suggest that for *D. melanogaster* environmental robustness for size constitutes an important component of male fitness. This association between phenotypic consistency and high male fitness supports the 'selection for perfection' hypothesis, which suggests that some sexual species can incur severe fitness costs from developmental inaccuracies, which translate into a disadvantage in intrasexual competition. This theory has been cited as an explanation for the slower development time of male *D. melanogaster* relative to females, as they have to develop more complex sex organs (Miller and Pitnick 2003). Here we suggest that the developmental stability of size is greater for high fitness males. We hypothesize that this robustness for size may provide a means through which males

can maintain a positive size differential relative to their local competition, when LDs are high. This means that, upon eclosion from development in the same larval environment, a given male is at least relatively larger than their immediate competitors.

However, it is important to also view our results in the context of sex-specific fitness, because we use males expressing FB genomes as our reference for low male fitness. From this perspective the lower phenotypic robustness of size for the FB genomes is consistent with the idea that phenotypic masculinization of females is stressful (Prasad et al. 2007; Abbott et al. 2010), and may therefore reduce resistance to stressful conditions during development, such as intense larval competition. In addition, quantitative genetic evidence suggests that polymorphisms responsible for buffering environmental variation are generally sex specific (Fraser and Schadt 2010). As a result, we might expect that expressing FB genomes in males will generate a lower level of environmental robustness, than if the same genomes were expressed in females.

3.4.7 Measurements of male fitness

We chose to measure male fitness in a one-on-one competition against standardized competitors. This methodology encompasses all male traits that contribute to fitness prior to copulation, and differs from the more commonly used measure of overall fitness, derived from the relative number of offspring sired by each TG. The latter overall fitness measure encompasses traits contributing to fitness before and after copulation. Here we demonstrate that our one-on-one fitness measures produce results that closely match the overall measures of male fitness. This could suggest that variation generated by traits that contribute to male fitness after copulation, such as sperm competition may have a relatively small effect on overall male fitness, and that access to females largely determines male fitness. Alternatively, any traits that affect male fitness after copulation may be positively correlated with the traits that are important before copulation. The traits that contribute to our fitness measures include any that are involved in the acquisition of mating opportunities, such as male mating song. Given the prevalence of sperm competition in *D. melanogaster,* and its importance for male reproductive success (Bretman et al. 2009), it is most likely that fitness related traits that operate before and after copulation are in fact highly correlated. We can therefore infer from our data that the male mating signals provide a reliable indication of overall male fitness.

Another, more robust explanation for the consistency between measures of overall male fitness, and our one-on-one competition measures comes from a recent study that measured pre and post copulatory reproductive success in the same stock population as that which our sample of genomes was derived (LHm) (Pischedda and Rice 2012). This study showed that under the conditions that the LHm population has adapted for, it is male mating order that dictates which male sires the most offspring, whereby the last male to mate sires the majority of progeny from each female. Specifically they demonstrated that, after adjusting for mating order, only 2% of residual variation in male mating success was attributable to differences in fertilization success. If we assume that all LHm populations have adapted to this selective pressure for high mating success, the similarity we observe between pre and post copulatory fitness can be explained by the absence of fitness variation that is attributable to factors other than mating success. In other words our measure of mating success (via one-on-one competitions) captures almost exactly the same fitness variation as the overall measures used in previous studies.

Finally, we also find that the correlation between pre-copulatory fitness and overall fitness, for males, remains even as LD differs from the normal conditions of the rearing regime (intermediate LD), to which our sampled population has become adapted. These findings are in line with the theory that males expressing these genomes experience proportionally similar levels of fitness change with increases and decreases in larval density.

3.4.8 Conclusions

In this chapter we measured fitness as a function of mating success, and tested for associations with aspects of wing morphology. An important finding was that the size of a given male (measured here by proxy from wing size) is important relative to his immediate competition for mating opportunities. Further increases in size when a male is larger than his competitors do not substantially increase a male's ability to acquire matings. Specifically, our results suggest that male size only becomes a critical determinant of male fitness when a given male is smaller than his immediate competition for females. Where effects of size contribute relatively less to the fitness of a male, i.e. when a male is larger than his competitor, wing shape accounts for proportionately more fitness variation than wing size. We also find that overall, across

a range of sizes, wing shape significantly affects male fitness to a similar degree to wing size.

We find some evidence of divergent selection on male wings. Specifically, we show that the allometric change in wing shape associated with increased wing size is, in part, divergent with the direction of selection on male wings for increased fitness. These results suggest that males may not be able to acquire their optimal shape and optimal wing size simultaneously. We hypothesize that strong selection on females for increased size (due to fecundity selection), combined with the high intersexual correlation for the genetic component of size, could force males away from their optimal wing shape.

Finally, we provide evidence that high male fitness is associated with a high phenotypic robustness for male size when subject to increased levels of larval competition for food during development. Specifically, we find a significant negative correlation between overall male fitness and a proxy estimate of robustness for size across the LDs we imposed. Further study is required to establish when this is truly a product of greater phenotypic robustness, or just a by-product of MB genomes carrying a relatively low SA load compared with other genomic fitness classes tested in this chapter.

3.5 References

Abbott JK, Bedhomme S, Chippindale AK (2010) Sexual conflict in wing size and shape in *Drosophila melanogaster*. J Evolution Biol 23(9): 1989-1997.

Arnqvist G, Nilsson T (2000) The evolution of polyandry: multiple mating and female fitness in insects. Anim Behav 60: 145-164.

Bitner-Mathe BC, Klaczko LB (1999) Plasticity of *Drosophila melanogaster* wing morphology: effects of sex, temperature and density. Genetica 105(2): 203-210.

Bookstein FL (1991) Morphometric tools for landmark data. New York: Cambridge University Press.

Bouteiller C, Perrin N (2000) Individual reproductive success and effective population size in the greater white-toothed shrew *Crocidura russula*. P Roy Soc B-Biol Sci 267(1444): 701-705.

Bretman A, Fricke C, Chapman T (2009) Plastic responses of male *Drosophila melanogaster* to the level of sperm competition increase male reproductive fitness. P Roy Soc B-Biol Sci 276(1662): 1705-1711.

Breuker CJ, Patterson JS, Klingenberg CP (2006) A single basis for developmental buffering of *Drosophila* Wing Shape. Plos One 1(1).

Burkhardt D, Delamotte I, Lunau K (1994) Signaling Fitness - Larger males sire more offspring - Studies of the stalk-eyed fly *Cyrtodiopsis whitei* (Diopsidae, Diptera). J Comp Physiol A 174(1): 61-64.

Cavicchi S, Guerra D, Giorgi G, Pezzoli C (1985) Temperature-related divergence in experimental populations of *Drosophila melanogaster* .1. Genetic and developmental basis of wing size and shape variation. Genetics 109(4): 665-689.

Chapman T, Miyatake T, Smith HK, Partridge L (1998) Interactions of mating, egg production and death rates in females of the Mediterranean fruit fly, *Ceratitis capitata.* P Roy Soc B-Biol Sci 265(1408): 1879-1894.

Chippindale AK, Gibson JR, Rice WR (2001) Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. P Natl Acad Sci USA 98(4): 1671-1675.

Cluttonbrock TH, Albon SD, Harvey PH (1980) Antlers, body size and breeding group-size in the *Cervidae*. Nature 285(5766): 565-567.

Darwin C (1874) The descent of man and selection in relation to sex (2nd edn.), London (1874); Murray J, editor.

David JR, Gibert P, Mignon-Grasteau S, Legout H, Petavy G et al. (2003) Genetic variability of sexual size dimorphism in a natural population of *Drosophila* melanogaster: an isofemale-line approach. J Genet 82(3): 79-88.

Debat V, Debelle A, Dworkin I (2009) Plasticity, canalization, and developmental stability of the *Drosophila* wing: Joint Effects of Mutations and Developmental Temperature. Evolution 63(11): 2864-2876.

Ewing AW (1964) Influence of wing area of courtship behaviour of *Drosphila melanogaster*. Anim Behav 12(2-3): 316-&.

Ewing AW, Bennetcl.Hc (1968) Courtship songs of *Drosophila*. Behaviour 31: 288-+.

Fairbairn DJ (1990) Factors influencing sexual size dimorphism in temperate waterstriders. Am Nat 136(1): 61-86.

Fairbairn DJ (1997) Allometry for sexual size dimorphism: Pattern and process in the coevolution of body size in males and females. Annu Rev Ecol Syst 28: 659-687.

Fowler K, Partridge L (1989) A cost of mating in female fruit-flies. Nature 338(6218): 760-761.

Fraser HB, Schadt EE (2010) The quantitative genetics of phenotypic robustness. Plos One 5(1).

Gavrilets S, Rice WR (2006) Genetic models of homosexuality: generating testable predictions. P Roy Soc B-Biol Sci 273(1605): 3031-3038.

Glucksmann A (1981) Sexual dimorphism in human and mammalian pathology: Academic Press, N.Y.

Greenacre ML, Ritchie MG, Byrne BC, Kyriacou CP (1993) Female song preference and the period gene in *Drosophila.* Behav Genet 23(1): 85-90.

Harvey PH, Cluttonbrock TH, Kavanagh M (1978) Sexual dimorphism in primate teeth. J Zool 186(Dec): 475-485.

Hoikkala A, Aspi J, Suvanto L (1998) Male courtship song frequency as an indicator of male genetic quality in an insect species, *Drosophila montana*. P Roy Soc B-Biol Sci 265(1395): 503-508.

Houle D, Rowe L (2003) Natural selection in a bottle. Am Nat 161(1): 50-67.

Innocenti P, Morrow EH (2010) The sexually antagonistic genes of *Drosophila melanogaster*. Plos Biol 8(3).

Le Boeuf BJ, and Michael L. Bonnell (1980) Pinnipeds of the california islands: abundance and distribution. p 475 Santa Barbara Museum of Natural History.

Le Boeuf J (1974) Male-male competition and reproductive success in elephant seals. American Zoologists 14(1).

Long TAF, Rice WR (2007) Adult locomotory activity mediates intralocus sexual conflict in a laboratory-adapted population of *Drosophila melanogaster*. P Roy Soc B-Biol Sci 274(1629): 3105-3112.

Markow TA (1988) Reproductive-behavior of *Drosophila melanogaster* and *Drosophila nigrospiracula* in the field and in the laboratory. J Comp Psychol 102(2): 169-173.

Mezey JG, Houle D (2005) The dimensionality of genetic variation for wing shape in *Drosophila melanogaster.* Evolution 59(5): 1027-1038.

Miller GT, Pitnick S (2003) Functional significance of seminal receptacle length in *Drosophila melanogaster*. J Evolution Biol 16(1): 114-126.

Moya-Larano J, El-Sayyid MET, Fox CW (2007a) Smaller beetles are better scramble competitors at cooler temperatures. Biology Lett 3(5): 475-478.

Moya-Larano J, Vinkovic D, Allard C, Foellmer M (2007b) Gravity still matters. Funct Ecol 21(6): 1178-1181.

Mueller HC, and K. Meyer (1985) The evolution of reversed sexual dimorphism in size: a comparative analysis of the falconiformes of the western palearctic." 2: 65-101. Curr Ornithol.

Partridge L, Farquhar M (1983) Lifetime mating success of male fruitflies (*Drosophila melanogaster*) is related to their size. Anim Behav 31(Aug): 871-877.

Partridge L, Fowler K (1992) Direct and correlated responses to selection on age at reproduction in *Drosophila melanogaster*. Evolution 46(1): 76-91.

Partridge L, Ewing A, Chandler A (1987) male size and mating success in *Drosophila melanogaster* - the roles of male and female behavior. Anim Behav 35: 555-562.

Partridge L, and Farquhar, M (1983) Lifetime mating success of male fruitflies (*Drosophila melanogaster*) is related to their size. Animal Behaviour 31, 871-877.

Peters TM, Barbosa P (1977) Influence of population-density on size, fecundity, and developmental rate of insects in culture. Annu Rev Entomol 22: 431-450.

Pischedda A, Rice WR (2012) Partitioning sexual selection into its mating success and fertilization success components. P Natl Acad Sci USA 109(6): 2049-2053.

Pitnick S (1991) Male size influences mate fecundity and remating interval in *Drosophila melanogaster.* Animal Behaviour 41: 735-745. .

Prasad NG, Bedhomme S, Day T, Chippindale AK (2007) An evolutionary cost of separate genders revealed by male-limited evolution. Am Nat 169(1): 29-37.

Prout T, Barker JSF (1989) Ecological aspects of the heritability of body size in *Drosophila buzzatii*. genetics 123(4): 803-813.

R Development Core Team (2006) R: a language and environment for statistical computing. Vienna, Austria, R Foundation for Statistical Computing.

Ralls K (1977) Sexual Dimorphism in Mammals - Avian models and unanswered questions. Am Nat 111(981): 917-938.

Ritchie MG, Townhill RM, Hoikkala A (1998) Female preference for fly song: playback experiments confirm the targets of sexual selection. Anim Behav 56: 713-717.

Roff D (1980) Optimizing development time in a seasonal environment - the ups and downs of clinal variation. Oecologia 45(2): 202-208.

Rohlf FJ (1999) Shape statistics: procrustes superimpositions and tangent spaces. J Classif 16(2): 197-223.

Rohlf J (2010) TpsUtil, TpsDig (http://life.bio.sunysb.edu/morph/).

Santos M, Fowler K, Partridge L (1994) Gene-environment interaction for body-size and larval density in *Drosophila melanogaster* - an investigation of effects on development time, thorax length and adult sex-ratio. Heredity 72: 515-521.

Thattai M, van Oudenaarden A (2004) Stochastic gene expression in fluctuating environments. Genetics 167(1): 523-530.

Turner TL, Stewart AD, Fields AT, Rice WR, Tarone AM (2011) Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. Plos Genet 7(3).

Weber KE (1990) Selection on wing allometry in *Drosophila melanogaster*. Genetics 126(4): 975-989.

Westeberhard MJ (1983) Sexual selection, social competition, and speciation. Q Rev Biol 58(2): 155-183.

Zwaan B, Bijlsma R, Hoekstra RF (1995) Artificial selection for developmental time in *Drosophila melanogaster* in relation to the evolution of aging - direct and correlated responses. Evolution 49(4): 635-648.

3.6 Tables and Figures

Table 3.1 Genetic variation for fitness across LDs. Results of binomial GLMs to test the dependent variable male fitness with the variables, genome, LD, and their interaction (A). This analysis is then divided into three pair-wise comparisons of low-intermediate, intermediate-high, and low-high LDs (B).

A. All LD treatments

	Df	*Deviance*	*Residual Dev*	*P-value*
LD	2	42.552	875.7	<0.001
Genome	8	68.039	807.66	<0.001
LD*Genome	16	58.595	749.06	<0.001

B. Pair-wise analysis between LD treatments

Low - Intermediate

	Df	*Deviance*	*Residual Dev*	*P-value*
LD	1	5.689	595.09	<0.001
Genome	8	63.962	531.13	<0.001
LD*Genome	8	31.652	499.48	<0.001

Intermediate - High

	Df	*Deviance*	*Residual Dev*	*P-value*
LD	1	17.003	587.34	<0.001
Genome	8	53.09	534.25	<0.001
LD*Genome	8	11.201	523.05	0.19

Low - High

	Df	*Deviance*	*Residual Dev*	*P-value*
LD	1	41.755	568.97	<0.001
Genome	8	52.851	516.12	<0.001
LD*Genome	8	40.512	475.61	<0.001

Table 3.2 Wing size differences between LDs and genomes. Results of a two-way ANOVA of the dependent variable wing size with LD, genome, and their interaction. Subsequently a post-hoc Tukey test was run on the linear model for TG size with the independent variable, LD.

A. Wing size variation between LDs

	Df	*Sum Sq*	*F-value*	*P-value*
LD	2	1230269	615.69	< 0.001
Residuals	552	566492		

B. Post-hoc Tukey test

	diff	*lwr*	*upr*	*P-value*
Intermediate-High	41.506	34.224	48.789	< 0.001
Low-High	118.601	110.614	126.587	< 0.001
Low-Intermediate	77.094	69.26	84.927	< 0.001

Table 3.3 Wing size variation between fitness classes, across LDs. Results of two-way ANOVA on the dependent variable mean wing size with the variables fitness class LD, and their interaction (A). This analysis is divided into three pair-wise comparisons, using mean wing size data from low-intermediate, intermediate-high, and low-high LDs (B).

A. All larval densities

	Df	*Sum sq*	*F-value*	*P-value*
Fitness class	2	1632	2.98	0.076
LD	2	61250	111.85	<0.001
LD*Fitness class	1	1829	1.67	0.201
Residuals	18	4928		

B. Pair-wise analysis between adjacent larval densities

Low - Intermediate

	Df	*Sum sq*	*F-value*	*P-value*
Fitness-class	2	496.8	1.06	0.376
LD	1	26338.7	112.41	<0.001
LD*Fitness class	2	1211.5	2.59	0.116
Residuals	12	2811.7		

Intermediate - High

	Df	*Sum sq*	*F-value*	*P-value*
Fitness-class	2	2190.3	4.08	0.044
LD	1	6507.2	24.24	<0.001
LD*Fitness class	2	266.7	0.49	0.621
Residuals	12	3221.7		

Low - High

	Df	*Sum sq*	*F-value*	*P-value*
Fitness-class	2	1491	2.3401	0.139
LD	1	59029	185.2544	<0.001
LD*Fitness class	2	1266	1.9861	0.179
Residuals	12	3824		

Table 3.4 Overall variation of male wing shape. Results of a multivariate analysis of variance (MANOVA) test using the complete matrix of PC axis scores as the dependent variable, with the variables TG wing size, LD, and genome. Initial analysis used data from all genomes, across all LDs (A). This analysis was divided into a pair-wise comparisons between adjacent LDs (B) and into separate MANOVA tests, applied to different subsets of fitness class (C).

A. All larval densites

	Df	*Pillai Sc.*	*Approx F*	*P-value*
TG size	1	0.675	49.112	<0.001
Density	2	3.983	23.757	<0.001
Genome	8	0.713	13.126	<0.001
Genome*Density	16	1.015	1.648	<0.001

B. Pair-wise analysis between adjacent larval densities

Low - Intermediate

	Df	*Pillai.Sc*	*Approx F*	*P-value*
TG size	1	0.771	50.673	<0.001
Density	1	0.429	11.308	<0.001
Genome	8	3.979	15.203	<0.001
Genome*Density	8	0.771	1.639	<0.001

Intermediate - High

	Df	*Pillai.Sc*	*Approx F*	*P-value*
TG size	1	0.599	25.6169	<0.001
Density	1	0.594	25.1015	<0.001
Genome	8	3.921	16.7751	<0.001
Genome*Density	8	0.648	1.539	<0.001

C. Within fitness class

Male benefit genomes

	Df	*Pillai Sc.*	*Approx F*	*P-value*
TG size	1	0.641	10.465	<0.001
Density	2	0.955	5.402	<0.001
Genome	2	1.436	15.066	<0.001
Genome*Density	4	0.603	1.065	0.334

Female benefit genomes

	Df	*Pillai Sc.*	*Approx F*	*P-value*
TG size	1	0.801	30.72	<0.001
Density	2	1.117	9.668	<0.001
Genome	2	1.621	32.68	<0.001
Genome*Density	4	0.786	1.89	<0.001

Neutral genomes

	Df	*Pillai Sc.*	*Approx F*	*P-value*
TG size	1	0.649	15.242	<0.001
Density	2	1.201	12.452	<0.001
Genome	2	1.501	24.92	<0.001
Genome*Density	4	0.568	1.384	0.015

Table 3.5 Variation of individual PC axes. Graphical representation of the results from 22 two-way ANOVAs, each using one of the PC axes as the dependent variable. Each PC axes represents a component of shape variation, and is tested for variation with the variables, TG wing size, LD, genome, and the interaction of LD-by-genome. For simplicity we only show which terms vary significantly with each PC axis. Dark grey squares represent a significant relationship with the corresponding PC axis ($P < 0.05$). +/- signs in the 'TG wing size' column represent the coefficient of allometric variation between each PC axis and TG wing size.

Multiple ANOVA results
Axis score ~ TG size + LD * Genome

PC Axis ~	*TG wing Size +*	*LD +*	*Genome+*	*LD* Genome*
PC1	+			
PC2	+			
PC3				
PC4	+			
PC5				
PC6	+			
PC7	+			
PC8				
PC9				
PC10				
PC11				
PC12				
PC13	-			
PC14				
PC15				
PC16				
PC17	+			
PC18				
PC19				
PC20				
PC21				
PC22				

Table 3.6 Fitness variation across LDs. Results of binomial GLM for the dependent variable male fitness with the covariate competitor wing size, and variables, TG wing size, LD, and genome. This model was first applied to the complete data set, covering all LDs (A), and then subdivided into pair-wise analyses between adjacent LDs (B).

A. All larval densities

	Df	*Deviance*	*Residual Dev*	*P-value*
Competitor wing size	1	9.373	779.63	0.002
TG wing size	1	24.417	755.21	<0.001
LD	2	20.583	734.63	<0.001
Genome	8	68.105	666.52	<0.001
LD*Genome	16	24.212	642.26	0.085

B. Pair-wise analysis between adjacent larval densities

Low-Intermediate

	Df	*Deviance*	*Residual Dev*	*P-value*
Competitor wing size	1	2.141	500.73	0.143
TG wing size	1	1.439	499.29	0.23
LD	1	9.767	489.52	0.002
Genome	8	52.288	437.23	<0.001
LD*Genome	8	13.876	422.93	0.085

Intermediate-High

	Df	*Deviance*	*Residual Dev*	*P-value*
Competitor wing size	1	8.825	569.07	0.003
TG wing size	1	48.113	520.95	<0.001
LD	1	0.057	520.89	0.812
Genome	8	42.282	478.61	<0.001
LD*Genome	8	11.079	466.79	0.197

Table 3.7 Male fitness variation with wing morphology, within each LD treatment. Results of two binomial GLMs testing the dependent variable male fitness with the covariates, competitor size and wing shape (PC axes), and the independent variables LD and genome. The covariate TG wing size is included (A) and excluded (B) to account for fitness effects of allometry between wing size and shape. Models were based on the minimized model generated by AIC analysis. Where applicable the proportion of explained fitness variation for the covarites size and shape is calculated from the corresponding deviance values.

A. All larval densities

Male fitness ~ Comp size + TG size + PC6 + PC8 + PC9 + PC19 + PC20 + LD + Genome

	Df	*Deviance*	*Residual Dev*	*P-value*
Competitor wing size	1	9.373	779.63	**0.002**
TG wing size	1	24.417	755.21	**<0.001**
PC6	1	8.651	746.56	**0.003**
PC8	1	7.428	739.13	**0.006**
PC9	1	0.839	738.29	0.359
PC19	1	6.973	731.32	**0.008**
PC20	1	2.32	729	0.127
LD	2	21.73	707.27	**<0.001**
Genome	8	57.778	649.49	**<0.001**
Explained variance: size		**0.175020966**		
Explained variance: Shape		**0.165236651**		

B. All larval densities, excluding TG size term

Male fitness ~ Comp size + PC1 + PC2 + PC6 + PC19 + PC20 + LD + Genome

	Df	*Deviance*	*Residual Dev*	*P-value*
Competitor wing size	1	9.373	779.63	0.002
PC1	1	1.73	777.9	0.188
PC2	1	3.473	774.42	0.062
PC6	1	11.682	762.74	**<0.001**
PC19	1	6.012	756.73	**0.014**
PC20	1	2.125	754.6	**0.144**
LD	2	16.469	738.14	**<0.001**
Genome	8	72.66	665.48	**<0.001**
Explained variance: Shape		**0.143243418**		

Table 3.8 Fitness variation with wing morphology with each LD. Results of three binomial GLMs testing the dependent variable male fitness with the covariate, competitor wing size, and variables TG wing size, wing shape (PC axes), LD, and genome. Each final model was generated by AIC analysis (see Methods and Materials). Where applicable the proportion of explained fitness variation for the covariates size and shape is calculated from the corresponding deviance values.

A. Low larval density

Male fitness ~ Comp size + PC1 + PC2 + PC3 + PC4 + PC10 + PC17 + PC18 + PC19 + PC21 + PC 22 + Genome

	Df	*Deviance*	*Residual Dev*	*P-value*
Competitor wing size	1	0.965	209.25	0.325
PC1	1	0.279	208.97	0.597
PC2	1	0.123	208.84	0.725
PC3	1	1.62	207.22	0.203
PC4	1	0.341	206.88	0.559
PC10	1	0.001	206.88	0.977
PC17	1	5.278	201.6	**0.021**
PC18	1	3.105	198.5	0.078
PC19	1	6.973	191.53	**0.008**
PC21	1	2.118	189.41	0.145
PC22	1	0.378	189.03	0.538
Genome	8	55.264	133.77	**<0.001**
Explained variance: TG wing size		**N/A**		
Explained variance: TG wing shape		**0.16025901**		

B. Intermediate larval density

Male fitness ~ Comp size + TG size + PC6 + PC8 + PC11 + PC19 + PC22 + Genome

	Df	*Deviance*	*Residual Dev*	*P-value*
Competitor wing size	1	0.984	290.57	0.321
TG wing size	1	13.141	277.43	**<0.001**
PC6	1	8.543	268.89	**0.003**
PC8	1	10.89	258	**<0.001**
PC11	1	9.742	248.25	**0.001**
PC19	1	1.627	246.63	0.202
PC22	1	1.387	245.24	0.238
Genome	8	20.648	224.59	**0.008**
Explained variance: TG wing size		**0.196245632**		
Explained variance: TG wing shape		**0.435694872**		

C. High larval density

Male fitness ~ Comp size + TG size + PC19 + Genome

	Df	*Deviance*	*Residual Dev*	*P-value*
Competitor size	1	5.915	262.26	**0.015**
TG size	1	19.375	242.89	**<0.001**
PC19	1	5.944	236.94	**0.014**
Genome	8	27.569	209.38	**<0.001**
Explained variance: TG wing size		**0.329489992**		
Explained variance: TG wing shape		**0.101083278**		

Figure 3.1 Wing size at each LD. Box plot of wing size scores across all nine genomes for each LD. Thick black horizontal lines represent the mean size score for each density, and each box represents the upper and lower quartiles, each containing 25% of values above and below the mean. Whiskers represent the most extreme values, excluding outliers.

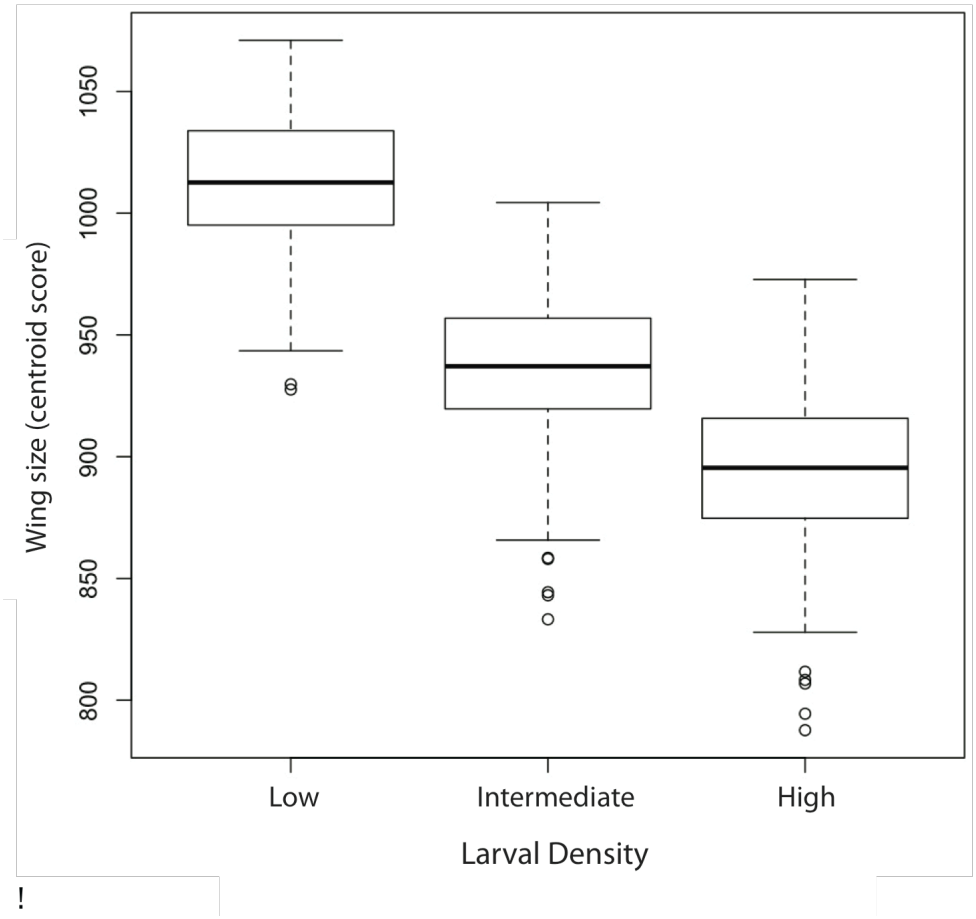


Figure 3.2 Wing size variation between LDs. Plot of mean wing size, measured as centroid score, for each LD. Here we show genomes of male benefit (MB - blue) and female benefit (FB – red), when expressed in males. Neutral genomes (N) are excluded here to improve the clarity of contrasting trajectories of MB and FB genomes. Error bars represent the standard deviation.

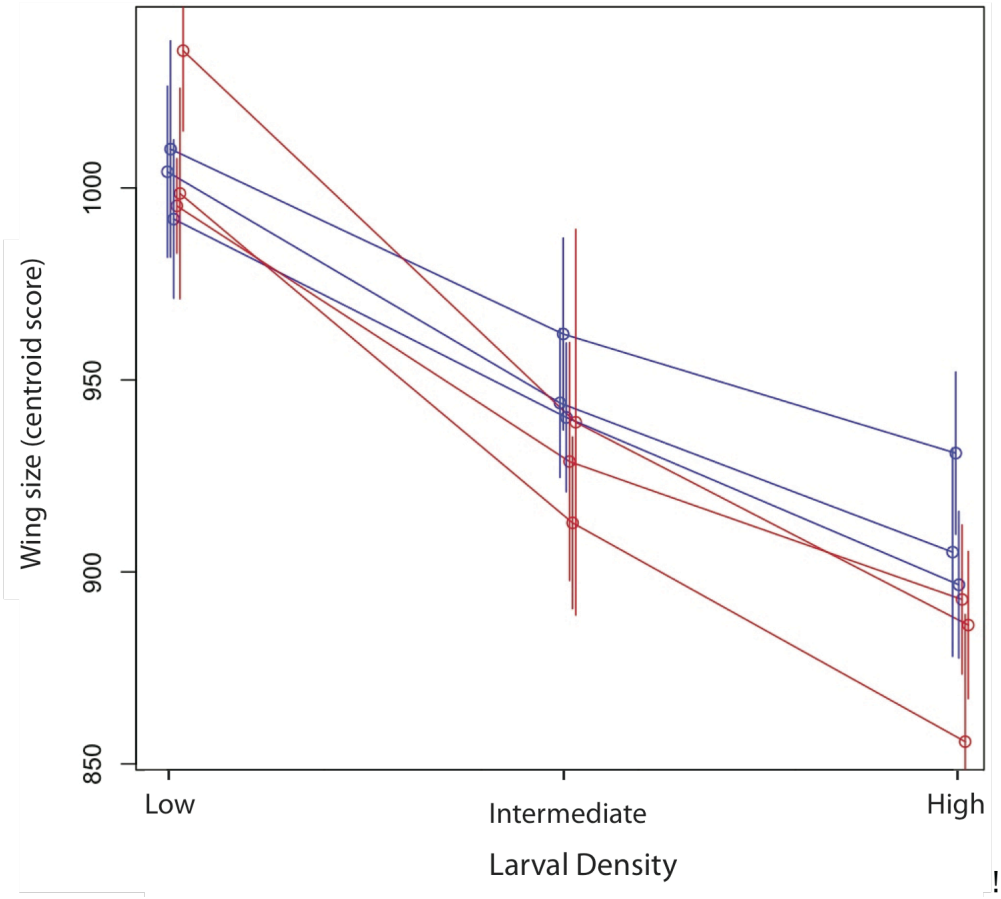


Figure 3.3 Mean fitness of TGs across LDs. Plot representing the mean male fitness of each TG for each LD. Genomes are color coded by their fitness-class; MB (blue), N (black), and FB (red).

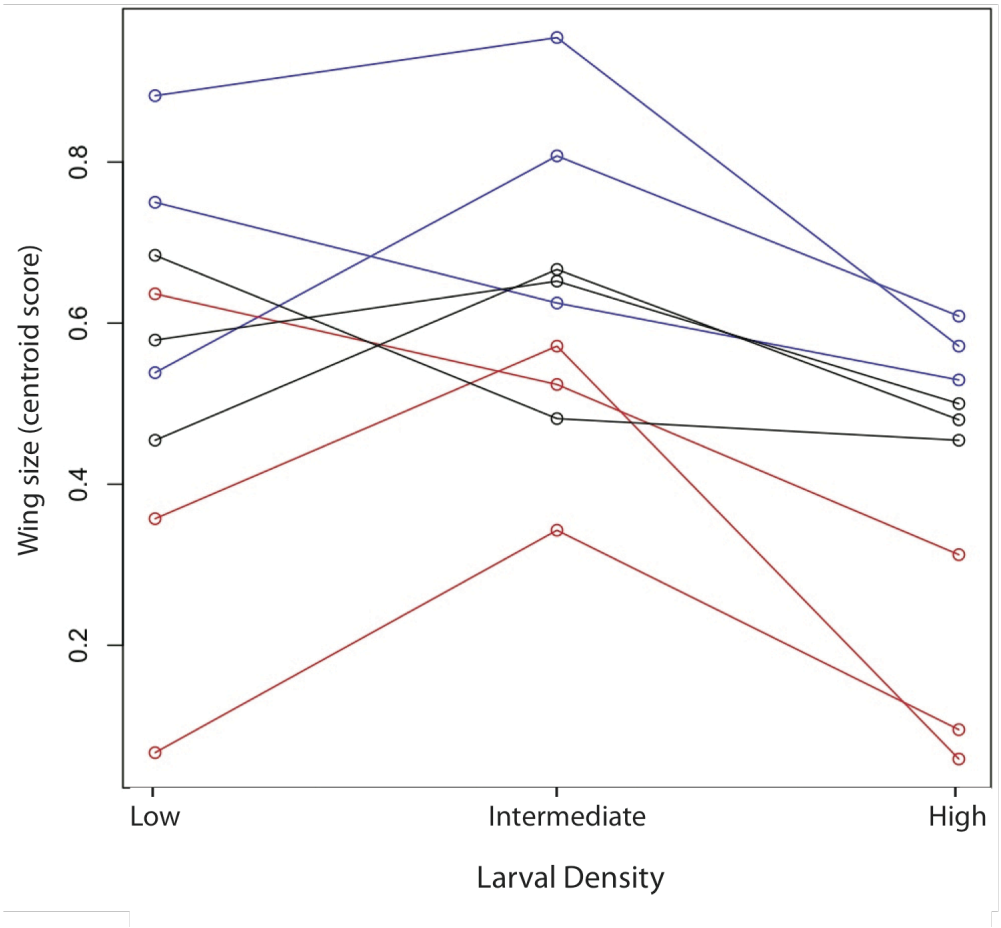


Figure 3.4 Male size differential vs mean male relative fitness. Correlation of male size differential across LD (Low – High), with mean male relative fitness (averaged across all LDs) (r = -0.488, n = 9, P = 0.182). Colours indicate the fitness-class of each genome (MB = blue, N = red, and IB = black).

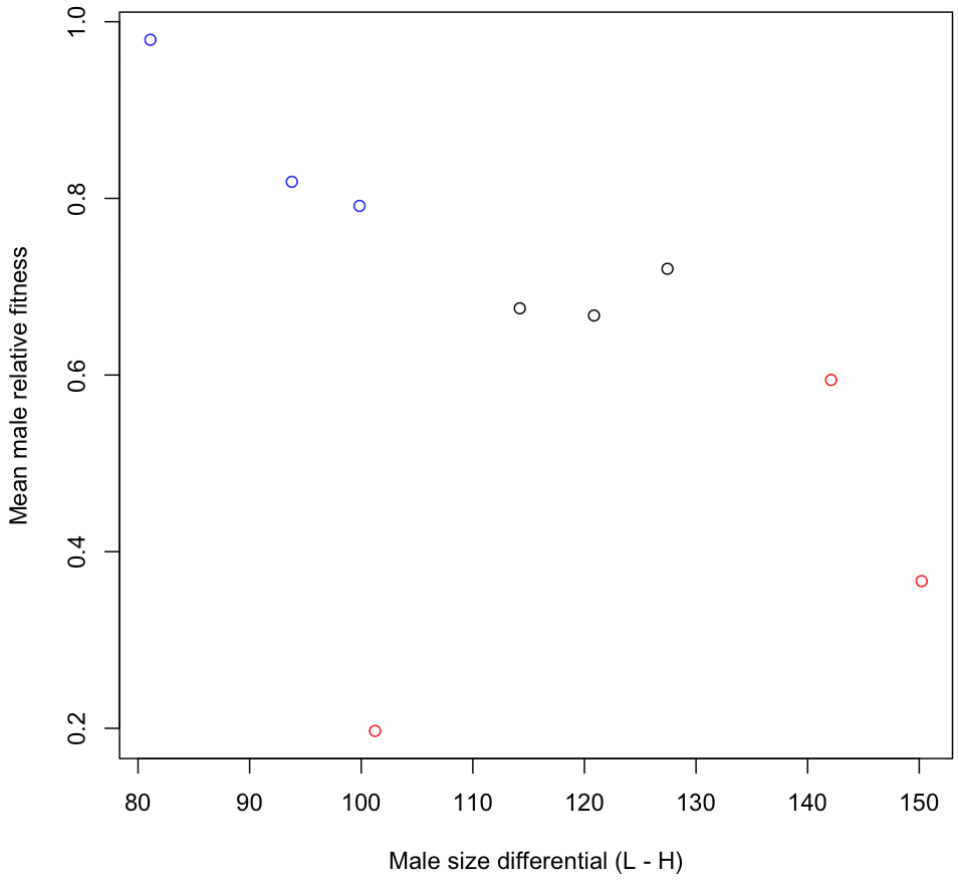
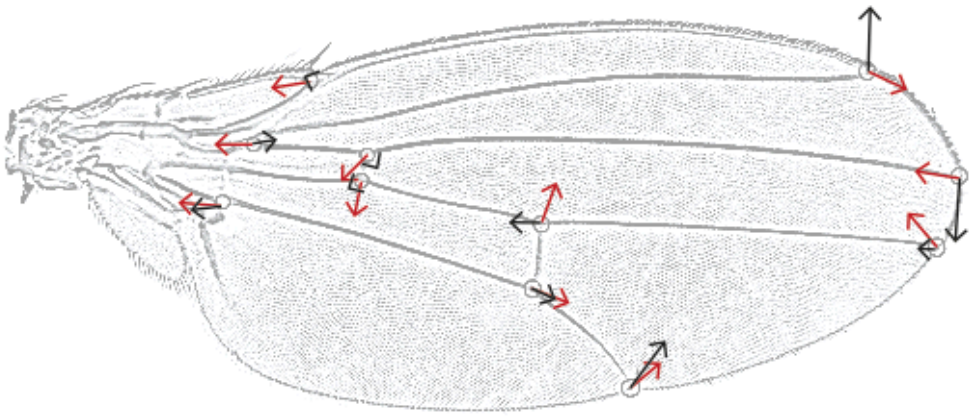


Figure 3.5 Selection on male wing shape. *D. melanogaster* wing showing the direction of selection for increased male fitness on LMs (black arrows). Here directionality is deduced from the sum of the loadings values from the PC axes that significantly effect male fitness (PC6, PC8, and PC19). Arrows are scaled relative to each other to show the degree of selection at each LM. For greater clarity the scale of selection is multiplied by a factor of 2. We also show the direction of wing shape variation that results from allometry with wing size, across LDs (red arrows). This allometric variation is not scaled, and just represents the direction of selection.



Supporting material

Figure 3.S1 Negative intersexual genetic correlation between mean male and mean female relative fitness (-0.52, 95% C.I. = -0.86; -0.1), extracted directly from, Innocenti and Morrow (2010). The fitness scores were measured for 100 hemiclonal lines (equivalent to TGs) expressed in each sex. Colored points indicate the TGs from which we chose our nine TGs used for experimentation in chapter 3 and 4 (red = Female Benefit – FB, blue = Male Benefit – MB, and black = Neutral – N).

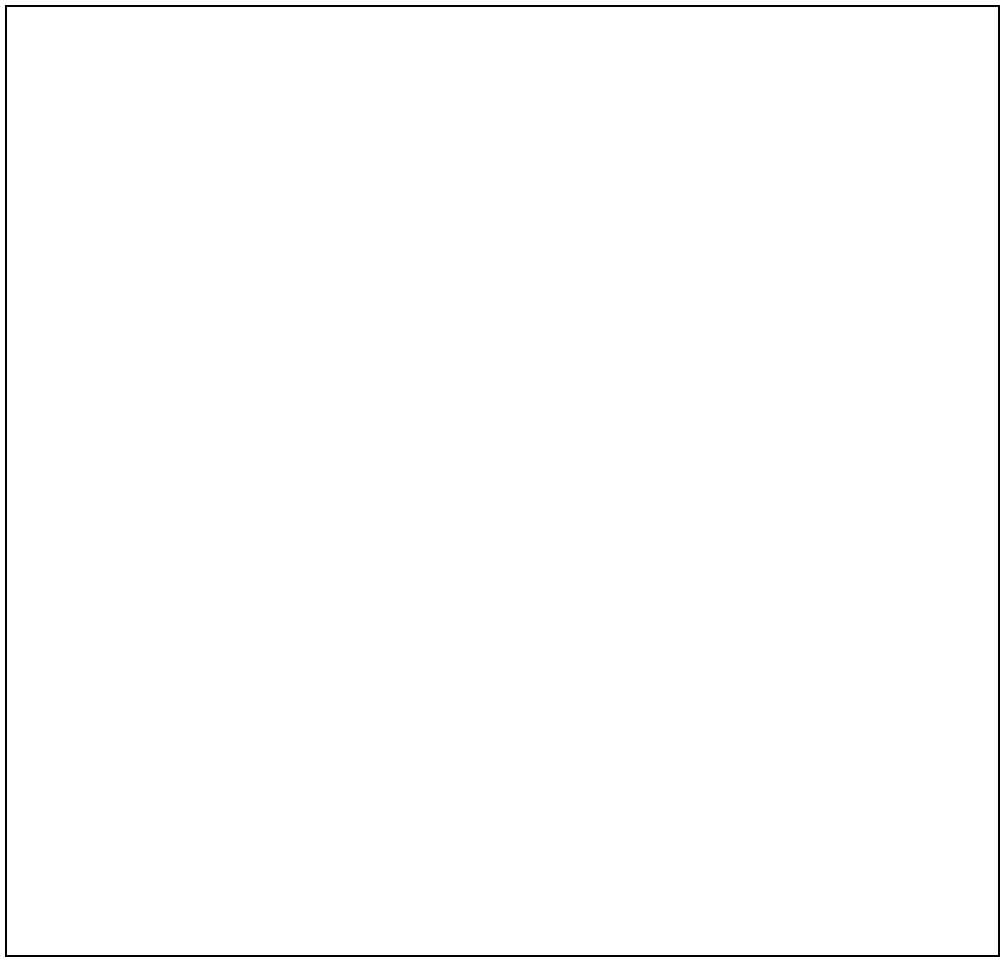


Figure 3.S2 Digitization of *D. melanogaster* wing LMs. Figure shows the order and location of the wing vein intersections where each of the 11 LMs are located for digitization using tpsDig (Rohlf 2010).

4

The contribution of wing morphology to sex-specific fitness and sexual antagonism

4.1 Introduction

Sexually antagonistic (SA) variation for fitness is caused by sexually divergent selection on traits shared by males and females. Because many male and female phenotypic traits have a common genetic basis, alleles frequently have opposing fitness effects when they are expressed in the two sexes. In other words, alleles that benefit the fitness of one sex can cause deleterious effects in the other sex (see sections 1.3 and 2.1.1 for further details). New mutations with SA effects on fitness can invade a population so long as the benefit to one sex outweighs cost in the other (Rice 1984). When SA alleles do invade they can often be maintained in stable polymorphism (Gavrilets and Rice 2006) due to divergent selection between the sexes, especially if located on the X or Z chromosomes in heterogametic systems (see section 1.4).

4.1.1 Trait values that contribute to sexual antagonism

Much of the earliest empirical evidence that laid the foundations for our current understanding of SA variation for fitness took a quantitative genetic approach. A textbook example using the model organism *D. melanogaster* examined 40 randomly chosen genomic haplotypes (a haploid complement of chromosomes X, II and III). When expressed in males and females, the fitness effects of these haplotypes showed a negative intrasexual genetic correlation for adult fitness (Chippindale et al. 2001), and hence demonstrated the existence of genome-wide SA fitness variation in *D. melanogaster*. Other studies have demonstrated the presence of SA variation for fitness in a range of species, including red deer (Foerster et al. 2007), fly catchers (Brommer et al. 2007), mountain goats (Mainguy et al. 2009), and lizards (Svensson et al. 2009). All of these studies detected sexual antagonism for organisms in their natural environment, showing that SA is not just an oddity of laboratory-adapted populations, as has been suggested previously (Chapman et al. 2003). As an alternative to using quantitative genetics, some studies have taken a trait focused approach to investigate sexual antagonism. A recent example comes from research on bank voles (*M. glareolus*), where it was demonstrated that the sex hormone testosterone has sexually antagonistic fitness effects. In mammals testosterone typically shows sexually dimorphic expression, has sexually divergent fitness effects, and is specifically known to affect mating behavior. By artificially selecting for high levels of testosterone Mokkonen et al. (2011) demonstrated that raised testosterone levels increased male reproductive success by increasing their mating rate. However, raised testosterone markedly

decreased female reproductive success by reducing the number of matings a female would partake in. These findings show that testosterone, through its effect on mating behavior comes under sexually antagonistic selection in this population of bank voles.

A particular trait will contribute to SA fitness variation if a) the trait is positively genetically correlated between the sexes, and b) the trait values have opposing fitness effects between the sexes. The higher the intersexual genetic correlation for a given trait, the less sex-independent variation there is for selection to act upon. As a result intersexual genetic correlations restrict the evolution of sexual dimorphism in response to divergent selection. When traits are correlated between the sexes, and subject to sexually divergent selection, then the phenotype of males or females can be displaced from their optimal trait value. Whilst a particular trait value may have negative fitness effects in one sex, the same trait value can have positive fitness effects in the other sex.

An interesting experiment has used these principles to demonstrate the SA fitness effects of a particular behavioural trait in *D. melanogaster* (Long and Rice 2007). This study exploited the extant SA fitness variation known to exist in the LHm population, and measured the sexually dimorphic trait of adult locomotory activity (defined as the frequency of movement over multiple 8 second intervals). First it was demonstrated that locomotory activity was highly genetically correlated between the sexes. Genomes that generated relatively high movement rates when expressed in males also generated relatively high movement rates in females, and vice versa. Importantly, it was also shown that the selection gradients (fitness change with increased locomotion) for adult locomotion were in opposite directions for the two sexes. Male locomotion correlated positively with fitness, whereas females suffered a fitness cost with increased rates of locomotion. This trade-off was what would be predicted from what is known of the reproductive roles of male and female *D. melanogaster*. Males are usually required to find and chase females to acquire matings (Greenspan and Ferveur 2000), making high locomotion rate an essential trait for optimizing reproductive success. Conversely, females pay a reproductive cost from excessive movement, as their fecundity is largely mediated by food consumption and the conversion of this energy into egg production (Knight and Robertson 1957). As a result, energy and time spent on movement reduces their lifetime fecundity. Importantly, Long and Rice's (2007) experiment also demonstrates that traits for which there is extant sexual dimorphism, such as locomotory rate, can contribute significantly to the overall SA variation for fitness.

Consequently, the existence of sexual dimorphism for a particular trait does not necessarily imply that all sexual antagonism has been resolved.

Another trait that could contribute to SA variation in *D. melanogaster* is wing morphology. Recent research, also using the LHm population, has shown that wing morphology may be subject to divergent selection between the sexes (Prasad et al. 2007; Abbott et al. 2010). This study used male-limited evolution, where genomic haplotypes were restricted to expression in males only. This allowed evolution towards the optimal male phenotype in the absence of selection due to the phenotype they express in females. Following more than 80 generations of male-limited evolution, male wings evolved into an extreme masculinised phenotype that was consistent with the direction of sexual dimorphism. These evolved morphological changes were associated with increased male fitness. These findings suggested that, under normal conditions where genomes were expressed in both sexes across generations, there must have been selection on females that restricted males from achieving their optimal wing morphology.

In *Drosophila*, the extant sexual dimorphism for wing morphology is for male wings to be smaller in size than those of females, with a shape that is relatively shorter and broader (Gilchrist et al. 2000; Gidaszewski et al. 2009; Abbott et al. 2010). The shape of many morphological structures is known to vary with their overall size and such allometric relationships are, in part, responsible for the shape differences observed between *Drosophila* males and their larger female conspecifics (Gidaszewski et al. 2009). However there is not yet a broad consensus on the extent to which allometric variation with size determines sexual shape dimorphism (SShD) in *Drosophila*. Gilchrist et al. (2000) measured wing shape in males and females across a range of environmental size clines in independent populations of *D. melanogaster*. Their results suggested that within each of the three tested populations, the overall variation in SShD was predominantly attributable to sexual size dimorphism (SSD). They suggested that allometry serves as a developmental constraint on SShD, and as such SSD explains almost all wing shape differences between the sexes. However, subsequent research by Gidaszewski et al. (2009), estimated that, on average, allometry and size dimorphism explain just half the total variation in shape dimorphism across 9 separate *Drosophila* species within the *D. melanogaster* subgroup. This latter study shows that the combination of allometry and SSD is not the only contributing factor towards SShD.

On balance it is likely that other factors must also contribute to SShD, as has been shown in other species (e.g. O'Higgins and Collard 2002; Bruner et al. 2005).

In *D. melanogaster* the two most likely factors that could contribute to selection on wing morphology are flight and courtship song. As courtship song is only performed by males, there are alternative functions which provide a foundation for divergent selection towards different optimal wing morphologies for each sex (see chapter 3 section 3.4.4). In the wild, females need to fly to find oviposition sites, and because of their larger body size it is logical to predict that they could be more dependent on wings with good aerodynamic properties. Conversely, males predominantly spend their time searching out mating opportunities, and the courtship song that they generate with their wings has been shown to be an important component of intrasexual competition among males (Snook et al. 2005).

In Chapter 3 (section 3.4.2) the majority of our results on the fitness effects of male wing size suggested that male reproductive success increased with larger wing size, which by proxy shows that a larger body size was good for males, as is also the case for females (e.g. Knight and Robertson 1957). Specifically, we found that a positive size ratio for a given male, relative to their immediate competitors, had the greatest bearing on reproductive success. However, our findings also suggested that the fitness of male *D. melanogaster* in our study was more directly a consequence of variation in wing shape than in wing size. We also provided some evidence suggesting that variation in wing shape with increased wing size could conflict with directional selection on wing shape, which constituted indirect evidence that males could be subject to some conflicting selection between optimal wing size and optimal wing shape.

Contrary to these findings for selection on male wing morphology, previous experiments on female *D. melanogaster* suggest that their fecundity is strongly correlated with phenotypic size (e.g. Sang 1950; Tantawy and Vetukhiv 1960; Partridge et al. 1986), largely because the number of eggs they can produce directly correlates with body size (Robertson 1953). Collectively, these studies suggest that the overall pattern of the fitness effects of size in *D. melanogaster* are consistent with the notion that females gain more than males from increased size (Charnov et al. 1981). Importantly, body size is highly genetically correlated between male and female *D. melanogaster*. If the consequence of this allometric relationship is that wing size

variation explains a substantial proportion of the variation we observe in wing shape then there is substantial scope for sexually divergent selection on overall wing morphology in *D. melanogaster* when selective pressures on wing size and/or shape differ between males and females. In Chapter 3 we found evidence that the direction of selection on male wing shape that is associated with increased mating success in part opposes allometric shape change that is caused by increased size. As a consequence we expect males may be subject to some selective pressure towards smaller wing size due to sexual selection towards wing shapes that, through allometry, are associated with a smaller wing size. On the contrary, we expect that females should be selected for larger wings at the expense of optimal shape due to the strength of fecundity selection for increased body size, and by proxy increased wing size. If real, these sexually divergent selection pressures, in combination with strong intrasexual genetic correlations for components of wing morphology provide a foundation for variation in wing morphology contributing towards SA fitness variation in *D. melanogaster*. There is substantial evidence for strong intersexual genetic correlations for size between the sexes in *D. melanogaster* (e.g. Karan et al. 1999), but limited evidence as to how the non-allometric components of wing shape correlate between the sexes. Furthermore, whilst there is an abundance of information on the sex-specific effects of body size, by proxy of wing size, on the fitness of each sex, the contribution of wing shape towards fitness variation has been largely untested.

4.1.2 Sex-specific effects of development time on morphology and fitness

One trait that is intrinsically linked to the body size of insect species is development time. Differences in development time between males and females are cited as a key mechanism in generating SSD, as eventual body size and development time are positively correlated. In other words, growing to be larger takes more time for both sexes (Roff 1980; Fairbairn 1990). Therefore, in the majority of insect species, where we observe a female biased SSD, we expect females to develop for longer than males. In *D. melanogaster* the duration of development typically takes around 8.5 - 9.5 days under optimal conditions although it can vary substantially within and between populations. *D. melanogaster* males, the smaller of the two sexes, usually take longer to develop than females. This is an inversion of the normal pattern of SSD with development time, whereby the larger sex takes longer to develop. This is somewhat paradoxical. Given that size and development time are correlated between the sexes in *D. melanogaster*, we would not expect males to require longer development than

females. In part this contradiction could be explained by the extended pupation period in males, but this is thought to explain no more than 1% of the SSD at eclosion (Nunney 1983). A better established reason for the longer development time of males, is an explanation based on the relative growth rates of each sex. Specifically, female larvae have been shown to increase their dry mass at a rate 25% to 33% faster than that of male larvae (Chippindale et al. 2003). However, this poses a second question: why then do males grow more slowly than females?

A key theory for explaining the extended development of male *D. melanogaster* comes from the 'selection for perfection hypothesis' (Chippindale et al. 2003). This proposes that males may grow more slowly because they benefit from more precise development, which is consequently slower. Particular emphasis has been placed on the importance of gonad development in males (L Partridge personal communication, Nunney 1996). This theory suggests that physiological imperfections can result in big fitness costs for males, because such imperfections can reduce a male's competitiveness in intrasexual selection. Consequently, we expect that where intrasexual competition among males is high then those males should be under greater selective pressure to develop more perfectly. In Section 3.4.6 we suggested that high male fitness was potentially associated with a high phenotypic robustness for size in response to changes in the level of larval competition for resources. This is consistent with the selection for perfection hypothesis, in demonstrating that male fitness was associated with phenotypic stability in response to environmental stress. Consequently, we predict that higher fitness in males, should be associated with a slower growth rate during development. For females, we expect a similar pattern for development time, as they gain direct fecundity benefits from increased size. Overall the relationship between development time and fitness in males remains poorly understood. In particular it is essential to establish the relationship between growth rate and male fitness in order to understand whether males are truly under directional selection for slower development. If males do benefit from a lower growth rate, this would set the scene for further studies to see whether increased morphological perfection is also associated with a slower growth rate.

4.1.3 Environmental lability

In Section 3.4.4 we showed that variation of environmental conditions, in this case larval density, can significantly alter the relative importance of wing shape and wing size for the fitness of male *D. melanogaster*. Much of this variation may be attributable

to genetic effects that were linked with high male fitness and so affected the phenotypic response to the manipulation larval density. In the laboratory we often go to great lengths to maintain a constant environment, with the intention of reducing environmental variation when measuring a given trait. However, interactions between genes and their environment are important for the maintenance of variation in quantitative characters. In the most simple terms, when the phenotypic effect of a particular allele is favored in environment A, but costly in environment B, and if there is environmental oscillation between environments A and B, then we expect divergent selection pressures to help maintain stable polymorphism for the allele in question (Felsenstein 1976). This basic process impacts upon the heritability of many traits, as they are expressed across different environments in the wild and vary in their fitness effects across those environmental conditions. Theoretical models suggest that estimates of genetic parameters in a single environment, as is the case in many laboratory based studies, may be of limited value to understanding trait evolution in natural environments (Gillespie and Turelli 1989). Therefore, to fully understand the fitness effect of a given trait, we should look to understand how phenotypes respond to environmental change. It is well established that phenotypic correlations between life-history traits are likely to change significantly upon exposure to novel rearing conditions (e.g. Matos et al. 2000; Sgro and Partridge 2000), often disappearing or reversing over short periods of evolutionary time. Chippindale et al. (2003) took the argument for the importance of gene-environment interactions to an extreme, suggesting that "the notion of a tractable genetic architecture underlying life-history traits is an intellectual mistake". Certainly we cannot underestimate the importance of environmental effects on life-history traits. Experiments designed to address environmental variation should provide greater insight as to the selective pressures that may operate in ever changing natural environments.

4.1.4 Aims of this study

Here we aimed to compare the sexes in terms of the relationships between three sexually dimorphic life history traits of *D. melanogaster*; wing size, wing shape and development time. We built on the data acquired in chapter 3 on male fitness and wing morphology (size and shape) for a panel of 9 TGs, by measuring the fitness and wing traits in females when expressing the same complement of target genomes (TGs). Development time was also measured for both sexes at low and intermediate LDs. LD was again standardised at densities classified relative to standard rearing conditions of

the LHm population, from which our nine TG's were sampled; intermediate (normal LHm rearing conditions), low and high. These densities are equivalent to those applied to TG males in chapter 3 (see section 3.2). In particular, these three rearing treatments provide two novel larval densities, one higher, and one lower than the conditions which our sample of genomes has adapted for.

We directly compared the fitness effects of wing morphology (size and shape) between the sexes. In addition, we investigated how wing shape varies for both sexes across a range of different wing sizes, that is, we estimate how allometry between wing size and wing shape varies between the sexes. If the allometry of wing morphology is aligned between males and females, then the selective forces that operate on wing size, and wing shape, must always operate in tandem across the sexes. By measuring wing size, wing shape, and fitness of genomes in both sexes, we aimed to produce an appropriate data set to estimate the selective forces that operate on wing morphology within and between the sexes. In particular, we directly tested for evidence of sexual conflict over wing morphology, and/or evidence for the genetic architecture and selective forces that could cause such sexual conflict. For example, if males are selected towards specific wing shapes, which, through allometry require a smaller wing size, then we might expect this to conflict with selection on females for greater wing size.

In this chapter, the environmental manipulation of adult body size through varying LD allows us to further investigate the robustness of the patterns between the life history traits that we assay. In part we aim to establish whether the patterns in females complement the evidence reported in chapter 3 suggesting that high male fitness is associated with the environmental robustness of wing size. The inclusion of measurements on a third trait, development time, allows the investigation of another trait known to be both sexually dimorphic and intrinsically linked to the body size of both sexes.

4.2 Methods and materials

For chapter 4 we measured female fitness and wing morphology to generate a data set analogous to that acquired for males in chapter 3. To allow direct comparison with the male data, we used the same 9 TGs, but now expressed in a female background. Much of the detailed protocol corresponding to the measurement of wing size and shape in females was identical to that described for males in chapter 3. Where there were differences in procedure, they are outlined below. In addition to measuring wing morphology in females, we measured development time in both males and females.

4.2.1 Experimental populations

The base population of *D. melanogaster* (LHm) has been maintained as a large (n > 1750 adults per generation), outbred population, for more than 450 non-overlapping generations. The attributes and maintenance of this stock, along with those of the competitor (LHm-bw), clone generator (CG), and male expression (Dx-LHm) stocks were all identical to those described in the chapter 3, section 3.2.

As in chapter 3, the nine TGs genomes were chosen because they represented a particular category of sex-specific fitness in both males and females, sampled from across the spectrum of SA fitness variation reported by Innocenti and Morrow (2010) for a random sample of the LHm base population (see supporting material, figure 3.S1). Whereas in chapter three male fitness effects were of interest, here our focus is on fitness effects in both sexes. The notation used to describe each genome is also identical to that used throughout chapter 3. Briefly, the nine genomes fell into three distinct fitness classes based on their fitness effects in males and females. Three genomes exhibited high female / low male fitness (Female Benefit - FB), three exhibited high male / low female fitness (Male Benefit - MB) and three exhibited an intermediate level of fitness for both sexes (Neutral – N). Each trio of genomes is collectively referred to as a fitness-class (MB, FB, and N).

4.2.2 Amplification and expression of TGs

Amplification of TGs and expression of TGs in a male LHm background, was carried out as described in section 3.2.3, by crossing to CG and Dx-LHm stocks, respectively. To express TG's in a female LHm background we mated TG-CG males to virgin LHm females. The females of interest constituted half of the female progeny from this cross

and received the TG from their father with a random haploid complement of their maternal LHm genome (TG-LHm). The remaining half of the female progeny carried the X chromosome from the TG together with translocated autosomes, and could be identified by the eye colour marker on their translocated autosomes.

4.2.3 Larval density manipulation

For each genome, larval density (LD) was controlled by manipulating the number of larvae that grow up in a fixed volume of media (10ml of agar-cornmeal-molasses media). The matings for the expression of TGs in male and female LHm genetic backgrounds were carried out in chambers supplied with plates of grape juice media. The eggs from these crosses were incubated until first instar larvae hatched. Groups of these first instar larvae were transferred into vials to create treatments of low, intermediate, and high LD. For females this corresponded to densities of 50 for low LD, 175 for intermediate LD and 300 for high LD. For the male cross in pilot experiments we observed a high proportion of death between the first instar stage and eclosion (~ 30%, see Appendix 1). To compensate for this increased mortality, we transferred batches of 60, 240 and 400 first instar larvae from male expression crosses, to create a range of low, intermediate, and high LDs equivalent to the LD treatments applied to females. For both sexes, we set up sufficient replicates of each treatment to provide enough offspring for accurate measurement of wing morphology and fitness. For females this corresponded to 7 low, 4 intermediate, and 3 high replicates; for males this corresponded to 6 low, 3 intermediate, and 3 high replicates.

To provide standardized competitors for the female fitness assay we set up a large stock of LHm-bw in synchrony with the expression crosses described above (details given in section 3.2.5). A total of 80 vials were set up, and tossed on for 5 consecutive days. LD was controlled at 175 eggs per vial to ensure that the size range of competitors was matched to that resulting from the similar density conditions that were maintained for the LHm rearing regime.

4.2.4 Measurement of development time

For the low and intermediate LD treatments we measured the development time of the male and female TG-LHm from egg laying to eclosion. Once the first adult flies began to eclose from the density-controlled vials (described above in section 4.2.3) we transferred all the newly eclosed flies into separate vials every 4 hours, until the end of

the period of eclosion. For each time interval, the transferred flies were frozen, subsequently scored for phenotype, and counted. This yielded counts of exactly how many TG-LHm flies eclosed during each 4 hour interval for TGs expressed in males and females. In measuring development time we also collected data on the total number of flies that eclosed from the low and intermediate LD treatments for both sexes. This enabled us to quantitatively compare the rearing conditions experienced by males and females at these LDs.

4.2.5 Fitness assays

Female fitness was measured as the rate of egg laying under competitive conditions that were identical to those imposed during each generation of the LHm rearing regime. We conducted a total of 81 fitness assays. That was 9 TGs, each assayed 3 times at each of the 3 LDs. To set up each assay we collected 6 LHm-TG females, 10 LHm-bw females (competitors), and 16 LHm-bw males. All flies were collected as virgins under cold anesthesia, during peak eclosion times for each LD (see Appendix 1). 24 hours after collection these groups were combined, using CO_2 anaesthesia, and placed in vials containing agar-cornmeal-molasses media and 10mg of dry yeast powder. These flies are allowed to interact for 48hrs (competition phase), after which the females were isolated into individual vials, and allowed to lay eggs for 18hrs (oviposition phase). Females were then discarded, and the progeny allowed to develop into adults. Once the period of eclosion had ended, the offspring of each LHm-TG female were counted. The total number of offspring generated by the 6 LHm-TG females in each assay was used as our measure of female fitness. For each genome we conducted nine replicate fitness assays, three for each of the three LDs. Mean fitness per TG for each LD, was calculated as the average of the 3 replicate assays within each LD.

4.2.6 Wing morphology

The flies used for measuring female wing morphology were the same LHm-TG individuals that were extracted and frozen to measure development time. As with the measurement of male wing morphology, the time frame over which wings were mounted for morphological measurement spanned the 'peak eclosion' phase for each LD (see Appendix 1). The techniques for mounting of wings, processing of landmark (LM) data, and morphometric analysis, was conducted using an identical methodology to that applied to male wings (see section 3.2.6). As we were primarily interested in the comparison of male and female wings, we carried out the generalized procrustes

analysis (GPA) on a full matrix of LM coordinates that included all male and female results. This matrix was composed of LM coordinates derived from the right wings of males and females, and consisted of 2000 replicates across 3 LDs and 2 sexes (570 male, and 1430 female). From this GPA we used the centroid size as a measure of wing size (see section 3.2.6). Principal component analysis (PCA) was applied to the full matrix of superimposed LM coordinates. This transformed the superimposed LM coordinates into a series of linearly uncorrelated variables, composed of 22 axes. This matrix of PC axes across both sexes was then analysed to reveal how shape varied between the genomes and LD treatments, across both sexes.

4.2.7 Statistical Analysis

Analysis of female wing morphology used a filtered data set for which any mating assay without a full complement of the measures, female fitness, TG wing size, and TG shape (22 PC axes) were removed from the analysis. This data set consisted of 1430 replicates; 323 low, 485 intermediate, and 622 high. As above, to analyse wing morphology between the sexes, this female data set was complemented with the analogous male data set (570 replicates) acquired for chapter 3, making the whole sample a total of 2000 replicates, across the two sexes.

When comparing the effects of aspects of wing morphology on the fitness of males and females we used a significantly smaller matrix of PC axes and centroid scores. This was because the only common unit of replication linking the wing morphology of each sex to a corresponding fitness value was that of the individual genomes at each LD. To generate this reduced matrix we calculated mean scores for all centroid sizes and PC axes, that corresponded to each genome at each of the three LDs. This data-set consisted of 54 data points, two sexes in each of nine genomes, at each of three LDs, with 27 data points per sex and 18 per LD. Likewise, the data set used for analysis of development time with fitness and wing morphology between the sexes, also used a data set where genome was the only unit of replication, except it also excluded any values from the high LD, for which we did not collect any development time data. We specify in the results where we used particular data sets, and where we transformed particular variables to aid analysis.

As in chapter 3, all statistical analyses were performed in R - version 1.40-devel, 64-bit, 2011 (R development core team 2006). Variation in wing size, fitness, and

development time was analysed by fitting linear ANOVA models with various combinations of covariates and variables, depending on the particular questions we wanted to answer. Each model is described within the results section below. Overall variation in wing shape across all LDs was analysed using multivariate analysis of variance (MANOVA) on the complete matrix of PC scores, and the statistical significance of each term was calculated with a Pillai test. To estimate how much of the overall shape variation is explained by each term in this MANOVA we calculated an approximation of the sum of squares (SS) for each term, derived from 22 separate univariate ANOVA's (Table 4.4). Theoretically, the PC axes were all independent of each other, so by separating them into univariate tests we did not fail to detect variation that would otherwise be detected in the original multivariate test. Each of the univariate models fitted one PC axis as the dependent variable, with the same combination of terms used in the original global analysis. We added the SS across these 22 analyses for each term, to give an overall estimate of the SS for each term. From these estimated SS values we calculated the total shape variation across all the univariate models, and the proportion of shape variation explained by each term.

As in chapter 3, our analysis of the fitness effects of wing size and wing shape consisted of a large number of linear predictors. In particular, there were 22 PC axes that represent variation in wing shape. Again to avoid over-fitting we generated the final models for statistical analysis using the 'step' function in R, which generated a model of best fit, based on Akaike Information Criterion (AIC). To correlate coefficients derived from linear models on wing shape we used the 'wtd.cor' function in R, which is part of the 'weights' add-on package. This allowed coefficients to be weighted by a given vector. Here we specifically weighted correlations of coefficients from linear models by $(1-P_{\text{model A}})(1-P_{\text{model B}})$. This metric gave the most weight to the coefficients that were significant in both models, less to those that were significant in only one model, and least to those that were non-significant in both models.

As in chapter 3, throughout our analysis we modelled the genome term as a fixed effect because we were working with a distinctly non-random sample of genomes. Consequently we did not expect the phenotypic effects of each genome to follow a normal distribution, as would usually be assumed.

4.3 Results

We first analysed the data on female fitness, specifically testing how female fitness varied between genomes across the different LDs we imposed. This gave us a picture of the variation for female fitness which was comparable to that which we had previously established in chapter 3 for males. We subsequently analysed the data on female wing morphology obtained here in conjunction with the male morphology data from chapter 3, to investigate how wing shape and size varied between the sexes. We then specifically tested for SA effects of wing morphology by comparing the fitness effects of shape and size changes between the sexes. Finally, we analysed development time across the sexes, focussing specifically on its relation to fitness and wing size.

4.3.1 Female fitness

For each individual fitness assay, female fitness was calculated as the sum of the total number of live offspring that emerged from each of the 6 TG oviposition vials for each replicate (see Methods and Materials). We had 9 of these of these replicates for each genome, 3 for each of the three LDs. We first used this data set to assess the genetic and LD effects on the fitness of TG females. We fitted a two-way ANOVA on the dependent variable female fitness with the variables LD, genome and the interaction term LD-by-genome. This revealed significant fitness variation between LDs, genomes, and the interaction term LD-by-genome (Table 4.1A). The effect of LD corresponded to a reduction in mean fecundity as LD increased, with an average of 251.85 ± 17.98 (mean ± SE) offspring per replicate at low LD, 182.63, ± 7.67 offspring per replicate at intermediate LD and 133.00 ± 8.38 offspring per replicate at high LD. The relatively small standard error of these means demonstrated that the LD effect on female fitness was significant for the step-up between each of the LDs that we imposed. This contrasted with the pattern observed in males, where overall fitness did not differ between the low and intermediate LD treatments. The significant genome effect on female fitness replicated our prior knowledge of these nine genomes, which were selected on the basis that they showed distinct differences for sex-specific fitness. Finally, the significant LD-by-genome interaction demonstrated that the genetic effects on fitness of the target genomes were not constant across environments but that genomes showed differential fitness responses to changes in LD.

Figure 4.1 shows a graphical interpretation of the genomic differences in female fitness across LDs. Inspection of this figure also revealed that the differential response of genomes to the density treatments was not random. The pattern of fitness changes in the fitness of individual genomes changes across densities appeared to be associated with the overall effects of genomes on female fitness. In particular, we observe that FB genomes experience relatively higher fitness at low LD but suffered a greater fitness reduction as LD was increased up to the high LD than did genomes that were N or MB. Numerically, the average proportional fitness loss for FB genomes between low and high LDs was 53.7%, as compared to 43.8% for MB and 39.7% for N genomes. To investigate whether the patterns of fitness change that we observe between the fitness classes are significant, we calculated the mean fitness of each genome at each LD and used this as the dependent variable in a two-way ANOVA with fitness class, density and their interaction as independent variables. This model confirmed overall differences in female fecundity between fitness classes across LDs (Table 4.1B, fitness class term), and corroborated the general decline of fitness across LDs (Table 4.1B, LD term). Most importantly, the ANOVA model confirmed that genomes of the different fitness classes responded differently to changes in LD, with the significant fitness class-by-LD term (Table 4.1B). As above, inspection of figure 4.1 suggested that this difference was most pronounced for the FB genomes, relative to N and MB genomes.

We integrated our male fitness data so that we could analyse the fitness effects of TGs between the sexes. We specifically investigated the pattern of SA fitness variation between our TGs using Spearman's rank correlations, as this is generally a robust test of association between variables when the sample size is small (in our case n = 9), and where the variables may not fulfill the assumption of bivariate normality. We correlated mean male and female fitness values across genomes, within LDs. Within the high LD, we obtained a significant negative correlation (rho = -0.767, n = 9, P = 0.021). Correlation within the low and intermediate LD treatments were also negative, but not statistically significant (low LD: rho = -0.2.66, n = 9, P = 0.493; intermediate LD: rho = -0.583, n = 9, P = 0.108). So overall our data tended to support the expected association between high female and low male fitness and high male and low female fitness across the genomes we studied here.

Finally, we tested for the correspondence between our measures of female fitness and those obtained previously by Innocenti and Morrow (2010) for the same genomes. We

detected significant positive relationships between these two fitness measures for all LDs (Spearman Rank Correlation; low LD; $r = 0.833$, $n = 9$, $P < 0.008$; intermediate LD; $r = 0.767$, $n = 9$, $P = 0.021$; high LD; $r = 0.733$, $n = 9$, $P = 0.031$). These results demonstrate that female fitness of these genomes was stable and could be measured in a repeatable manner. They further showed that the relative fitness rank of our nine TGs did not vary substantially in response to changes in LD.

4.3.2 Variation in wing morphology between the sexes

Having analysed female data in isolation, we combined the data sets for male and female wing morphology in a global analysis. This analysis covered both wing size and wing shape. Shape was represented by the 22 PC axes, each of which captured an independent aspect of shape variation among all the wings analysed (see Methods and Materials for detail). The data for each individual consisted of 22 PC scores (wing shape) and one centroid score (wing size), attached to a specific genome, LD, and sex. The total matrix consisted of 2000 individuals, 1430 females and 570 males.

We investigated how wing size varied between genomes, LDs and the two sexes, using a three-way ANOVA. We modelled TG wing size as the dependent variable with the variables sex, genome, LD, and all their interactions. This analysis showed that TG wing size varied very significantly with all model terms (Table 4.2A). Firstly, the significant effects of the individual terms, sex, genome, and LD, demonstrated that each of these factors independently contributed to variation in wing size across our whole sample. Specifically, the sex term confirmed the well-established presence of SSD, while the LD term showed that, across sexes and genotypes, wing size was sensitive to increased larval competition. Specifically, females showed mean centroid scores of 1164.34 ± 1.51 (mean ± SE) at low LD, 1081.32 ± 1.32 at intermediate LD, and 1009.12 ± 1.32 at high LD while males showed mean centroid scores of 1013.80 ± 2.25 (mean ± SE) at low LD, 935.72 ± 2.07 at intermediate LD, and 894.70 ± 2.53 at high LD. The differences between these means, relative to the small standard errors, indicated that wing size declined with every increase of LD for both sexes. Scrutiny of the sums of squares showed that the sex and LD terms explained 43.70% and 41.25%, respectively, of the total wing size variation within our dataset, demonstrating that sexual dimorphism and larval competition explained the bulk of wing size variation that we observed. The genome term explained the next highest proportion of wing size variation (3.88%), and its significant effect means that across the sexes and LDs, we

found consistent genomic effects on wing size. The remaining interaction terms collectively explained less than 2% of the total wing size variation. Specifically, the significant two-way interactions sex-by-LD and sex-by-genome suggested that the degree of SSD varies both between LD treatments and between genomes, even though the size variation caused by these effects was small. The significant genome-by-LD interaction revealed gene-by-environment interactions, whereby TGs differed across the sexes in their wing size response to changes in LD. Finally, the significant sex-by-genome-by-LD term suggested some level of sexual dimorphism in this gene-by-environment interaction.

For the interpretation of some of these interaction terms it is important to note that the ANOVA model was based on absolute size values. This meant that significant interaction terms could arise if the mean absolute size varied, despite the fact that proportional differences between groups remained constant. This is particularly relevant to the analysis of sexual dimorphism across the LDs since flies get bigger as LD decreases, so that the absolute difference between males and females will increase from high to low LD, even if relative dimorphism remains constant. The ANOVA fitted parameters on an absolute scale of wing size and so the combination of constant SSD and changes in absolute size between densities would lead to a significant sex-by-LD effect. To verify whether such scale effects confounded our results on wing SSD, we specifically tested for variation in proportional SSD between LDs. To do so, we calculated the ratio of the size differential (female / male) at each LD and each genome. Between LDs, we observed the greatest SSD within the intermediate LD (1.156), and the lowest within the high LD (1.128), while the low LD was intermediate (1.148). To establish whether these measures of SSD were significantly different between LDs, we transformed all TG wing size data onto the same scale by dividing all values (male and female) by the mean female wing size of their corresponding LD. These new values of relative TG wing size were then used as the dependent variable in a synonymous three-way ANOVA, as above, with the variables sex, genome, LD, and the interactions of these terms (Table 4.2B). The highly significant sex-by-LD term in this model confirmed that SSD did indeed vary between both LDs, even on a scale of relative sizes. Although the variation in SSD across the different LDs was a significant effect, it is important to acknowledge that this effect only explained 0.56% of the total variation of the whole model. This reflects that, even though the changes in SSD may be consistent, their scale was relatively small. The greatest change in SSD across LDs corresponded

to a 2.80% difference in relative female wing size. This difference was more than 20 times smaller than the wing size difference between low and high LDs, in both sexes.

To assess the relationship between male and female wing size across genomes we calculated the intersexual genetic correlation of wing size within each LD. We ran separate Spearman's rank correlations for each LD, to calculate the association between the mean wing sizes for males and females across the nine genomes. At the intermediate LD we observed as significant positive correlation (rho = 0.80, n = 9, P = 0.014). The correlations at the low and high LDs were also positive, but not quite significant (low: rho = 0.583, n = 9, P = 0.108; high: rho = 0.616, n = 9, P = 0.086). In addition to these individual correlations, we also analysed average relative sizes across the density treatment. To do this without confounding the size scores with density effects we calculated these averages from size values that were standardized relative to the mean of each LD for males and females. The standard scores for each replicate at each LD were calculated as the value measured for the replicate minus the mean size across the corresponding LD, divided by the standard deviation across the corresponding LD. Performing a Spearman's rank correlation between these scores revealed a highly significant positive correlation (rho = 0.933, n = 9, P < 0.001). These results demonstrated that the genomic size effect was highly correlated between the sexes across our sample. Accordingly, a genome which produced relatively large males also produced relatively large females, and vice versa.

It appeared that the sex-specific fitness effects of a given genome had a substantial impact on the phenotypic size response to changes in LD. This was highlighted in chapter 3 for males (section 3.3.2) and above for female (section 4.3.1). By combining the size and fitness data for the two sexes, we could directly compare how the size response to changes in LD varies between the sexes for genomes from each distinct fitness class (MB, FB, and N). This allowed us to investigate how the expression of the phenotypic trait wing size, in both sexes related to the sex-specific fitness effect of genomes. To test the wing size effects across LDs that are associated with the distinct fitness class of each genome, we fitted a three-way ANOVA. In grouping the genomes into their respective fitness classes, we again used genomes as our unit of replication. This reduced our sample to 54 data points (two sexes, times nine genomes, times three densities, with 27 points per sex and 18 per density). Our model used the dependent variable mean TG wing size with the variables, fitness class, sex, LD and all

interactions between these variables (Table 4.3A). The terms that we were interested in from this new model were those that include the fitness class variable, as all other terms and their interactions were dealt with above. Specifically, this initial model showed an almost significant effect associated with the term fitness class (P = 0.065). Considering the small sample sizes (three genomes per class), this can be taken to suggest that TG wing size tended to differ between our fitness classes across both sexes and all LDs, which demonstrated that these distinct fitness classes also generate a distinct pattern of size variation. The non-significant two-way interactions fitness class-by-sex and fitness class-by-LD, suggested that differences between the classes in sexual dimorphism and the response to changes in LD were not evident in our dataset, nor were differences in the extent of SSD between LDs (triple interaction). The fact that none of these effects were significant might once again be a consequence of the small sample sizes, combined with the small phenotypic differences between classes. Visual inspection of the data (Fig. 4.2) suggested that MB and FB genomes differed in their response to increasing LD, with FB genomes showing a more pronounced decrease in size with increasing density than MB genomes in both sexes. Contrary to the analysis presented in Table 3.4A, this effect was significant in an analyses on just the two extreme fitness classes, MB and FB (Table 4.3B). Finally, we tested whether the phenotypic response to changes in LD was correlated between the sexes. To do this we calculated the mean size differential per genome between the low and high LDs for males and females, and correlated these values using a Spearman's Rank between the sexes. This revealed a significant positive correlation (rho = 0.7, n = 9, P = 0.043), suggesting that the response phenotypic size response to LD is genetically correlated between males and females (Fig 4.3).

Overall variation in wing shape was investigated using MANOVA (see Methods and Materials for details). The dependent variable 'wing shape' was composed of the 22 PC axes. This wing shape was modelled with the covariate wing size, and the variables sex, genome, LD and all the interactions of these terms, plus the interaction terms wing size-by-sex and wing size-by-genome (Table 4.4). The inclusion of the covariate wing size allowed us to detect allometric changes in shape associated with changes in size, and removed their effects when analysing subsequent terms in our model. We also included the interaction terms wing size-by-sex, and wing size-by-genome to test how this allometry varied between the sexes and genomes. As with the equivalent MANOVA analysis on wing size, this model revealed highly significant variation in

overall wing shape with all terms. The significant wing size term showed that changes in wing size caused allometric variation in wing shape, across LDs and the two sexes. Furthermore the significant wing size-by-sex, and wing size-by-genome term suggested that this allometric effect of size varied between the sexes, and the genomes. The significant sex and LD terms confirmed the presence of SShD and wing shape variation between LD treatments. As above, the significant genome term showed consistent genomic effects on wing shape across the sexes and LDs. The interaction terms sex-by-LD, and sex-by-genome also suggested that the degree of SShD varied between the different LDs and genomes. Finally, the interaction term genome-by-LD and the triple interaction sex-by-genome-by-LD showed that the effect of LD on overall wing shape of TGs varied both across and between the sexes. It is important to note that the P-values for each of the terms were highly significant. We must therefore be careful not to over-interpret the importance of individual model terms. As wing shape can vary across multiple axes, the large number of different PC axes representing wing shape could have artificially increased the power of our model, by picking up various different patterns simultaneously.

To gain a better understanding of the relative importance of each term in the model described above, we used multiple univariate ANOVAs to estimate the proportion of wing shape variation that is explained by each term (see Methods, section 4.2.7, for more detail). This analysis showed that the bulk of the explained variance for wing shape was due to effects of the allometry (wing size term, 17.36%), and genetic differences (genome term, 22.59%). The sex term explained the next highest proportion of variance (4.00%), demonstrating that sex differences explained substantially less variation in shape than it did for size. The remaining terms, although significant in the overall analysis, explained only a small proportion of wing shape variation. Terms in the 1% - 2% category included LD, sex-by-genome and LD-by-genome. The small effect of LD can be explained by the strong effect of LD on size. So when size was already accounted for, the remaining effect of LD in altering wing shape was very low. Some terms, including size-by-sex, size-by-genome, sex-by-LD, and sex-by-LD-by-genome, explained less than 1% of shape variation. This showed that differences in allometric variation between the sexes and genomes were very small. Similarly, the shape response to changes in LD were similar between the sexes, and between the genomes across the sexes.

The analysis of wing shape suggested that allometric changes in shape with size were similar between the sexes. This was indicated by the large proportion of variance explained by the general size term, compared to the almost insignificant amount of variation that was attributable to the size-by-sex interaction. To verify whether the change of shape with varying size was indeed the same in both sexes, we compared the coefficients of shape change with size estimated in separate ANOVAs run on males and females. Again we used the 22 PC axes as the dependent variables, but this time, only included the single variable, TG wing size. We subsequently extracted the coefficients for shape change with size for each of the 22 PC axes from both analyses. We tested for their relationship between the sexes using a weighted Pearson's product-moment correlation. Each coefficient was weighted by the product of the complements of P-values obtained for the coefficients in the analyses on males and females, $(1-P_{male})*(1-P_{female})$. This metric gave the most weight to the coefficients that were significant in both sexes, less to those that were significant in only one sex, and least to those that were non-significant in both sexes. Performing the correlation analysis revealed a highly significant positive relationship between the allometry of male TG wings, and allometry of female TG wings ($r = 0.946$, $n = 22$, $P < 0.001$). This showed that despite their difference in size, allometric changes of wing shape with varying size were highly aligned between the sexes. Thus, the way in which the shape of a male wing changed with increasing size was very similar to how the shape of a female wing changed with increasing size.

To visualize how male and female wing shape changed as LD increased, we plotted the direction and relative extent of mean landmark movement across all genomes, from low to high LD (Fig. 4.4). This plot illustrated well the highly significant correlation of allometric variation between the sexes, above, by showing that the overall pattern of shape change across LDs was very similar for both males and females. The majority of shape change that was associated with an increase in wing size appears to have resulted from relative changes in the locations of the most distal landmarks, 8, 9, 10, and 11. To visualize the extent of sexual dimorphism for shape within our sample, we also plotted the relative scale and direction of landmark movement between males and females reared at the low LD (Fig. 4.5). This plot suggested that male wings were generally wider than female wings, caused predominantly by the differential locations of landmark 8 and landmark 11. Male wings were also slightly shorter than female wings which was caused by the movement of the most proximal landmarks (1, 2, and 3).

4.3.3 Fitness effects of sex-specific morphologies

We analysed the fitness effects of wing size and shape to investigate how variation in wing morphology affected reproductive success within and between the two sexes. Across the male and female data sets, when we included all morphological and fitness data, genomes were again the lowest common unit of replication. We used values of mean fitness and morphological traits (size, shape) for males and females, which were specific to a particular genome and LD. This dataset consisted of 54 data points, two sexes, times nine genomes, times three LDs, with 27 points per sex and 18 per LD.

We first modelled the effect of wing size and wing shape on sex-specific fitness using separate linear models on each of the male and female data sets. When considering the data set at our disposal, it is important to acknowledge that our experimental manipulation of LD resulted in large differences in the variables wing size and wing shape (through allometry and maybe direct effects of rearing density) between the different LD treatments. As a consequence of the discrete distribution of the variable values in the three LDs, any analysis of fitness was likely to be dominated by effects between these LDs and associated fitness variation with morphological variation between the different LDs. To avoid these unwanted LD effects interfering with our analysis of wing size and shape on fitness, we specifically accounted for density effects by including LD as the first term in our models. Consequently, subsequent model terms describing morphological variation captured fitness variation over and above the effect of these terms that was explained by differences in morphology associated with LD. The models therefore produced a minimum estimate of the fitness variation that was explained by wing morphology.

We modelled the dependent variable fitness with LD, and the linear predictors mean wing size and mean score for each of the 22 PC axes, which collectively represented mean overall shape (see Methods and Materials). As in chapter 3, the final models for analysing the fitness effects of wing morphology were generated by minimizing the number of terms based on AIC values (see Methods and Materials). Overall, we found that reproductive performance varied significantly with TG wing size and shape for both sexes (Table 4.5). Within each model we used the sum of squares values to estimate the proportion of fitness variation explained by wing size and wing shape. The final model for males included eight significant PC axes, but did not feature size.

Collectively these eight PC axes explained 89.50% of male fitness variation. The final model for females included wing size and seven significant PC axes. Wing size explained just 1.78% of female fitness variation, whereas the seven PC axes describing shape explained 45.31%. To investigate whether fitness variation with dimensions of wing shape was aligned between the sexes, we correlated coefficients describing the change of male and female fitness with each of the 22 PC axes derived from full linear models of fitness variation with LD, wing size and shape. Performing this analysis revealed a correlation that was weakly positive but not significantly different from zero ($r = 0.239$, $n = 22$, $P < 0.283$), suggesting that the fitness effects of wing shape variation across our three LDs were not aligned between the sexes. In order to establish whether selection on wing morphology differed between males and females, we further tested whether the correlation between the coefficients from male and female models differed significantly from 1 (the value expected if selection were perfectly aligned between the sexes). To do this we ran a linear model of the form $y_i - x_i = a + bx_i + \varepsilon_i$, where y_i, x_i and ε_i, were male and female coefficients for fitness change with PC axis i and normally distributed error, respectively. This analysis revealed highly significant difference from 1 ($b = -0.999$ $t_1 = -6927.466$, $P < 0.001$), demonstrating that our data showed no relationship at all for selection on wing shape between the sexes.

To understand how the fitness of each sex was affected by allometric variation of wing shape with wing size, we correlated the coefficients describing fitness change with varying shape estimated here, with the coefficients of allometric shape change estimated earlier. This specifically allows us to tests whether the fitness effects of wing shape are dependent on allometric changes in wing shape with wing size. Again coefficients were weighted by $(1-P_{male})(1-P_{female})$. Neither series of coefficients were correlated in females ($r = -0.067$, $n = 22$, $P < 0.765$) or males ($r = -0.046$, $n = 22$, $P < 0.838$), indicating that the directions of shape selection and allometric shape change were not aligned within each sex. Specifically, this suggested that fitness gains associated with altered shape were independent of allometric variation of wing shape and hence not confounded by changes in size. Collectively our correlation analyses were not indicative of sexually antagonistic selection on aspects of wing shape between the sexes. However, based on our models of fitness variation with different aspects of wing morphology, our results suggested that wing size was a more important fitness component for females than males, and that wing shape was more important for males than females.

It is worth noting that the proportion of fitness variation explained by the linear models fitted in these analyses was extremely high. This may in part be explained by the fact that we were using averaged values of fitness and morphology, thereby reducing sampling variance and amplifying the deterministic effects of morphological selection. However, it was also possible in these highly dimensional analyses that some of the morphological variation identified as important for fitness could by chance have reflected other phenotypic differences between genomes that are the true determinants of fitness. It is impossible to test this using our data, as genomes are our unit of replication and hence we could not separate morphological differences between TGs from differences in any other phenotypic attribute of these genotypes by including a genome term in the models.

4.3.4 Development time: size and fitness effects

For each genome, development time was measured as the mean number of hours between egg laying and eclosion of adult flies. In this experiment we only measured development time for the low and intermediate LDs (see Appendix 1). Across all genomes, we measured mean (±SE) female development times of 223.53 ± 0.28 hrs at low LD, and of 230.07 ± 0.52 hrs at intermediate LD. For males, mean development times were 231.87 ± 0.94 hrs for low LD and 237.64 ± 0.93 hrs for intermediate LD.

To test how development time varied between the sexes, LDs, and genomes we ran a global ANOVA on the full set of replicates for male and female development time across the LDs and sexes. All replicates available (see Methods and Materials, section 4.2.3) corresponded to a total sample size of 179, across two LDs, and two sexes. In males, this consisted of six and three replicates per genome for low and intermediate LD, respectively. In females it consisted of seven and four replicates per genome for low and intermediate LD, respectively. We fitted a three-way ANOVA, modelling the dependent variable development time with the variables sex, LD, genome, and the interactions of all these terms (Table 4.6A). This analysis revealed significant effects of the individual terms sex, LD and genome, but non-significant interactions between these terms. The significant sex effect supported the well-documented sexual dimorphism for development time and the LD term showed that increased larval competition increased development time for both sexes. Finally, the significant genome term showed that beyond the variation that we expected between the sexes and LDs, genetic effects still

added further variation to development time. However, the non-significant two-way interactions, genome-by-sex, and genome-by-density, showed that genomic effects on development time were the same, whether expressed in males or females, and when reared at low or intermediate LDs. The non-significant LD-by-sex term showed that across genomes, the effect of LD on development time was the same between the sexes. Finally, the non-significant triple interaction genome-by-LD-by-sex, showed that the effect of LD on development time was the same for genomes when they were expressed in either males or females. Collectively, these results suggested that development time was strongly determined by genetic effects, and these genetic effects were highly correlated between the sexes, and across LDs.

We subsequently tested for associations between development time, size and fitness. As the measures we collected for development time, size and fitness were not taken from the same individual flies, the genome was once again the unit of replication in this analysis. We fitted a three-way ANOVA with the dependent variable mean development time and independent variables mean wing size, mean fitness and sex, plus the interactions of all these terms. Models were fitted separately for low and intermediate LDs to minimize environmental variation across each analysis and directly measure genome effects in isolation. Both analyses revealed significant variation for mean development time between the sexes, but non-significant variation with all other terms (Table 4.6B). However, in the analysis for the intermediate LD, the terms sex-by-mean wing size, and sex-by-mean fitness, approached significance ($P < 0.1$). This suggested that development time may have been associated with size and fitness, but that the exact nature of this association differed between the sexes.

In addition to an effect of absolute development time on fitness, we also tested for an association between fitness and developmental rate. As a measure of developmental rate, we derived the residuals from a linear model of development time as a function of wing size and LD. This model was performed on each sex separately. Each residual described the relative rate at which each genome gained wing size during larval development. We subsequently modelled fitness as the dependent variable in an ANCOVA fitted on each sex separately with the residuals (development time, on wing size) and LD. For males (Table 4.7, Male) we observed a non-significant fitness difference between the low and intermediate LDs, which was established in Chapter 3. However, the effect of residual development time approached significance ($P = 0.083$).

The coefficient corresponding to this term was negative, indicating that higher male fitness was associated with faster development time, relative to size. This is the opposite pattern to what we would expect under the selection for perfection hypothesis. As a comparison, we also ran this analysis on data from females. Here, we observed a significant fitness difference between low and intermediate LDs, as established above (section 4.3.2), but a non-significant effect of residual development time (Table 4.7, Female). This suggested that developmental rate was not a determinant of female fitness.

Finally we used the data from the development time experiment to assess the intensity of larval competition experienced by male and female flies in the different LD treatments. As already mentioned in the Methods section 4.2.3, the mortality of offspring differed between the crosses that generated male and female experimental flies and this was due to the use of a stock with an attached X in the male cross. To compensate for this, the numbers of larvae used for the three LDs was not the same for males and females. In order to judge whether this resulted in equivalent rearing conditions for males and females in the different treatments, we analysed counts of emerging flies from the experiments conducted to measure development time under low and intermediate LD. We ran separate Generalised Linear Models (GLM) for Poisson distributed data (log link function) for the intermediate and low LD, and fitted the total number of TG adult flies as the dependent variable, with sex and genome as independent variables (Table 4.8). Likelihood ratio tests based on Chi-squared distributions were used to test statistical significance of the terms in each model. We observed a non-significant effect of sex within the intermediate LD, demonstrating that the total numbers of emerging adults are comparable between the sexes. However, within the low LD we found significant differences between the LD treatments applied to each sex. Specifically, we found that the mean number of adult flies that emerged within the low LD was 28 for males, and 42 for females.

4.4 Discussion

In chapter 4 our approach was to investigate the relationship between wing traits for male and female *D. melanogaster,* and measure their sex specific fitness effects. We built on our findings from chapter 3 on the fitness effects of male wing morphology by collecting an equivalent data set for females. We also added a third trait, development time. The duration of development correlates positively with size. It therefore potentially affects the fitness of both sexes and creates a link between larval and adult fitness. All our chosen traits show pronounced sexual dimorphism and are therefore likely candidates to be subject to divergent selection between the sexes in *D. melanogaster*. We specifically test whether the genetic interactions between these dimorphic traits contribute to genome-wide SA fitness variation. To establish the contribution of traits towards SA fitness variation we looked to fulfill two criteria, 1) a genetic correlation between the sexes for the particular trait of interest, and 2) divergent sexual selection on these morphological traits. In the following, we will explore whether these conditions are fulfilled for the traits studied here.

We first discuss the environmental and genetic effects on *D. melanogaster* wing morphology that we have documented across our sample of genomes. Within the discussion of the genetic architecture of wing shape and size, we also specifically look at the intersexual genetic correlations for our measurements of wing morphology, which are important in the context of sexual antagonism. We then assess the evidence for the contribution of these morphological components towards the fitness of each sex, and discuss the impact of wing morphology on SA fitness variation. We discuss morphological patterns that correspond to the distinct genomic fitness classes captured in our sample and finally we evaluate the relationship of development time with size, and in particular with male fitness.

4.4.1 Genetics of wing morphology

Our results on genetic size effects between the sexes support the accepted view that size is highly correlated between the sexes (e.g. Cowley et al. 1986; Fairbairn 1997). We demonstrate this effect with several different analyses. Specifically our global analysis, modelling the wing size effects between fitness classes, shows congruent effects of each fitness class on wing size between the sexes (Table 4.4). We also find a strong highly significant genetic correlation for size across all LDs (correlation coefficient

rho=0.933), using size scores that were standardized by the mean of their respective LD. Therefore, if a genome generates a relatively large female, it will also generate a relatively large male and vice versa. Overall the correlated size effects that we observe between males and females can be generally attributed to a sharing of the same genes that control growth and development (Roff 1997; Nijhout et al. 2006) which ultimately determine the adult size of the flies (see section 4.4.5 below for further detail). Combining the relative size effects of each genome across the LDs served as an effective means of increasing the power of our correlation analysis. The significant correlation of these averages demonstrates that the relative rank of genomes remains reasonably consistent between LDs. Therefore we infer that the genetic effects on size are relatively independent of the level of larval competition for resources during development. Our estimate of the male-female size correlation between the sexes is comparable to previous estimates of the inter-sexual genetic correlation for size which tend to be greater than 0.8 (Fairbairn 2007). However, it is important to note that these estimates are not strictly comparable to ours. Our genetic correlation was measured across environments and, more importantly, our sample of genotypes is based on specific genomic fitness effects and hence distinctly non-random. In contrast, quantitative genetic studies are based on random samples of genotypes drawn from a population.

Besides revealing sexually concordant genomic effects on wing size, our global analysis of size variation also provided some evidence that genomes differ in their size effect between the sexes, i.e., vary in their degree of sexual dimorphism (genome-by-sex effect, Table 4.4). These effects are relatively small compared to the overall wing size variation that we observe across our global model, and compared to the sexually concordant genomic effects on size. We are probably able to detect these small genetic differences because our large sample size (n = 2000) for the analysis of wing morphology provides our model with extremely high power. Genetic variation in sexual dimorphism is important, because it allows the independent evolution of the size of each sex. The effect of this sexually independent variation for body size is demonstrated in the seed beetle (*C. maculatus*). Although this species exhibits a high intersexual genetic correlation for size, it also showed a substantial degree of evolution in SSD after just 40 generations of selection in a novel environment (Messina 2004). This shows how a high intersexual genetic correlation for size does not totally impede the evolution of SSD.

Like wing size, we find that wing shape shows pronounced variation with effects of LD, genome, and sex. Between sexes, we find that the mean direction of SShD for wings is consistent with the general consensus that male wings are wider and shorter than female wings in *D. melanogaster* (Gilchrist et al. 2000; Gidaszewski et al. 2009; Abbott et al. 2010). Specifically, and consistently with previous studies (Gidaszewski et al. 2009), the distal part of the wing is wider in males relative to females, an effect we show is mostly due to an overall divergence of the wing veins that lead into LM 8 relative to LM 10 and LM 11 (Fig. 4.5). Previous studies have shown that another distinct feature of the SShD from females to males in *D. melanogaster* is a proximal shift of the two interior cross veins that run from LM 4 to LM 5, and LM 6 to LM 7 (Gidaszewski et al. 2009). Here we observe that these cross veins shift in the same proximal direction, but the overall scale of this movement is almost negligible compared with the movement of the distal LMs located on the perimeter of the wings, i.e. LMs 8, 10 and 11.

Previous studies suggest that the overall pattern of SShD is highly conserved both within *D. melanogaster* (Gilchrist et al. 2000), and across the *D. melanogaster* subgroup (Gidaszewski et al. 2009). In addition, research on *D. melanogaster* and other species has recognized that allometric effects of size on shape are an important component of SShD (O'Higgins et al. 1990; Bruner et al. 2005). Of particular relevance, Gilchrist et al. (2000) analysed morphological variation in *D. melanogaster* wing shape between three populations from different continents, each of which naturally occurred along similar environmental clines that create variation in body size. A very high conservation of SShD between populations from the different continents was observed, and a primary developmental constraint on wing shape was the allometric relationship between size and shape. As a consequence, they attributed shape variation between the sexes largely to variation in size and hence to SSD. Here we provide evidence that suggests the genetic architecture of SShD is in fact more complex and not just a product of size-related constraints. We find that the overall proportion of wing shape variation that is attributable to SShD across the sexes is low (4.00%), when the effects of shape variation with size (i.e. allometry 17.36%) are accounted for. At the same time, the overall proportion of size variation that is attributable to SSD across the sexes in the exact same data set is high (40.70%). We can estimate the proportion of the total shape variation that is attributable to sex differences in size by calculating the proportion of shape variation with size (17.36%) that is the result of sex differences in size (40.70%).

Performing this calculation provides an estimate that 7.06% of the total shape variation (i.e. 40.70% of 17.36%) is attributable to sex differences in size, i.e. allometric variation. This figure is larger than the 4.00% of shape variation that is due to non-allometric effects, demonstrating that wing size imposes some constraint on variation in wing shape. However, the difference between these values (4.00% and 7.07%) is not huge, suggesting that slightly more than a third of the total SShD is attributable to non-allometric effects. This result contradicts the idea that allometry imposes an almost complete constraint on wing shape, as proposed by Gilchrist et al (2000). Instead we provide more support for the findings of Gidaszewski et al. (2009), who found that non-allometric effects explain roughly half of the SShD across the *D. melanogaster* subgroup. Although our estimate of non-allometric SShD does not meet their estimation, the patterns of our data indicate that the effects of allometry are not the only factor that determines SShD.

However, our results specifically show that the allometric component of variation in wing shape is highly correlated between males and females. In other words, the change in shape associated with size variation is generally parallel between the sexes. Figure 4.4 depicts this pattern showing that the LMs, which move the most in response to increased wing size, do so in a parallel direction between the sexes. These findings suggest that the genetic architecture that determines the way in which wing shapes change with changes in wing size are very similar for both males and females. The underlying SShD is still preserved across a wide range of different size values because these shape changes are aligned between the sexes. Some authors have speculated that *D. melanogaster* wing shape is under tighter genetic control than wing size (Birdsall et al. 2000; Mezey and Houle 2005), especially under the fluctuating environmental conditions found in the wild (Gilbert et al. 1998). Consequently, in natural conditions we expect wing size to vary considerably more than wing shape, as it is largely determined by food availability and temperature during larval development. Here our results suggest that indeed wing shape will vary with environmental changes that affect wing size, but the allometric component of this shape variation is highly similar between the sexes.

An important finding is that the largest proportion of total wing shape variation, after accounting for affects of wing size, is explained by genetic differences (22%). In general this provides evidence in support of there being substantial standing genetic

variation for wing shape. This has been demonstrated in studies that have tested the response of wing shape to artificial selection, which frequently show a strong and rapid response to selection (Weber 1990; Hansen and Houle 2008). Here we specifically find that this genetic variation in wing shape is predominantly non-sex specific, which is indicated by the relatively small proportion of variation that is explained by genetic differences between the sexes (sex-by-genome interaction, < 2% variance). As a result, we expect that the genetic effects on shape are relatively similar between the sexes.

4.4.2 Fitness effects of wing morphology between the sexes

With regard to establishing the contribution of wing morphology to SA fitness variation, we have provided substantial evidence that variation for both wing shape and wing size are highly genetically correlated between the sexes. We now discuss our results on the selective pressures that operate on wing morphology in the two sexes. Firstly, we note that the analysis we carried out on wing size, wing shape, and fitness revealed that values of all these variables showed non-overlapping distributions between the three LDs. The discontinuities in the variation of these three phenotypic dimensions carried the risk of establishing spurious associations between morphology and fitness when modelling fitness as a function of wing shape and size. To avoid making such erroneous inferences, we included the LD term in our global analysis on the fitness effects of wing morphology, so as to account for such adverse affects. This means that our analyses cannot detect variation in fitness between larval densities that could be explained by variation between densities in size and/or shape. However it has the merit that by doing so we avoid wrongly attributing differences in fitness to differences in size and/or shape.

Between the sexes, we found that wing size is proportionately more important for the reproductive success of females than males. In particular the complete absence of wing size from the model on the fitness effects of male wing morphology highlights the relative insignificance that wing size has on male fitness when LD effects are accounted for. This particular result is in line with the general idea that female reproductive success (i.e. fecundity) in insects is more closely related to their body size, than it is for males (Honek 1993). However, the relative proportion of fitness variation explained by female wing size is also negligible compared with the effect of wing shape in both sexes. We find wing shape explains roughly two times the fitness variation of males compared with females, which in turn is more than 40 times the fitness variation

explained by wing size for females. This suggests that wing shape is more important for male fitness relative to wing size, which is more important for female fitness. We must note that despite seeking to account for effects of LD on size, the extremely high proportion of fitness variation explained by the wing shape components could be a product of correlated responses for some of the shape components (PC axes) with effects of wing size across the LDs. However, because allometric variation of wing shape with wing size is highly correlated between the sexes (see section 4.4.1), we expect that these shape changes are similar. As a result the inference that the relative importance of wing shape for males is double that of females is valid, because the variation in wing shape between the sexes is very small across the range of sizes that we imposed via manipulation of LD.

These findings serve as indirect evidence that selection on wing shape is relatively more important for males, and that selection on wing size is relatively more important for females. These results are generally supportive of our findings from chapter 3 on the fitness effects associated with male wing morphology. In chapter 3 we found that wing shape explained the majority of male fitness variation when size was not inhibitive, i.e. when a focal male was not smaller than the male that he was in competition with. This was specifically the case at low and intermediate LDs. Where we analysed the fitness effects of male wing morphology across all LDs we found that wing size and wing shape explained almost exactly the same proportion of male fitness variation. This, in part, contradicts our findings here, as wing size is not included in this minimum model. However, this difference between these results for males and females is likely due to the fitness effect of wing size being underestimated in the global model of male fitness with wing morphology in chapter 4. As above, this underestimation could be the result of wing size variation correlating with some shape components, and consequently reducing the proportion of independent fitness variation associated with the wing size component alone. Another possibility is that these differences were caused by the fact that it was necessary to use genome means at each LD as our unit of replication for measuring the fitness affects of wing morphology in chapter 4, which provides a relatively small amount of replication per genome. Previously (chapter 3), we were able to use fitness data for males at the level of the individual. It is possible that individual variation between the wing size of individuals from the same genome has a large effect on fitness on an individual level, but does not have a large influence on the average fitness between different genomes.

Authors have made functional arguments to suggest that aspects of wing shape could be important for male fitness in *Drosophila* species. Field research by Markow and Ricker (1992) on three separate species of *Drosophila* revealed the very high level of male wing flashing around mating sites leading the authors to suggest that aspects of wing shape could be an important component of sexual selection. Whilst both sexes use their wings for flight, only males use their wings to produce courtship song, providing the scope for functionally divergent selection on the same trait (wing morphology) between the sexes. We hypothesize, as have others (Gidaszewski et al. 2009), that differences in wing shape may impact on the specific characteristics of courtship song (Cowling and Burnet 1981) and that this may provide a foundation for strong selection towards particular wing shapes in males. Experimental evidence suggested that changes in sexual selection could generate modifications of courtship song (Snook et al. 2005) and demonstrated that the type of song generated is directly linked to sexual selection on males. Furthermore, the potential for strong selection on wing shape was highlighted by experiments using male-limited evolution, where *D. melanogaster* genomes had been limited to expression in males for more than 70 generations (Abbott et al. 2010). These experiments showed that increased male fitness was associated with masculinisation of wing size (smaller wings) and shape (shorter and wider wings). This serves as evidence that male wing shape comprises a selectable trait with consequences for male reproductive success.

Contrary to the fitness effects of wing shape, we find that wing size is relatively more important for the reproductive success of females. Overall we find that larger female size is associated with higher fitness, which is congruent with the widely accepted idea that female fecundity increases with body size (Knight and Robertson 1957). As wing size measurements serve as a good proxy for body size (Robertson 1953; David and Legay 1977), we can assume that our positive relationship between female wing size and female fitness represents this positive fecundity effect from increased female body size. In particular, our findings support the idea that females gain more in terms of fitness from increased size, relative to males (Charnov et al. 1981).

Whilst both sexes show a linear increase in overall mean size as LD is decreased, the change in fitness across larval density differs between the sexes. In females, fecundity increases in a linear fashion as LD decreases. In males, by contrast, increased size does

not appear to have any benefit beyond the mean size that is produced at the intermediate LD (see chapter 3 section 3.4.3) and male fitness increases only from the high to the intermediate LD. It therefore appears that once males are larger than their competitors they gain little fitness benefit form further size increases. Consequently, although the effect of relative males size does not directly show that males are not selected towards greater size, it does suggests that the balance of selection on size may differ between the sexes. While female size is under constant directional selection due to its effect on fecundity, the selective advantage of increased male size saturates and selection for increased size ceases to act once a male is larger than its competitors.

The overall pattern that we observe is reasonably consistent with the pattern observed by Prasad et al. (2007) on the fitness effects of wing size for males and females. Prasad et al. (2007) showed that when genome-wide selection was limited to males only, it produced a decrease in the average male size. This decrease in male size was associated with increased male fitness. The evolution of smaller male size in the absence of selection towards female specific trait values suggested that males may be displaced from their optimum size under normal evolutionary conditions, due to selection in females. The evolution of reduced male size provided strong evidence that wing size, and by proxy body size, is normally under divergent selection between the sexes, with males favored to become relatively smaller, and females relatively bigger.

We also tested the fitness effects of wing shape between the sexes. Our results from chapter 3 suggested that wing shape may be particularly important for male fitness, and functional arguments (see section 3.4.3), along with recent empirical evidence (Abbott et al. 2010) make a good case for the existence of strong, and potentially divergent selection on wing shape in *D. melanogaster*. Firstly, our results do not provide direct evidence that selection on aspects of male and female wing shape are opposed. This would be evident from a negative correlation between the coefficients describing the change of male and female fitness with measures describing shape. Our analyses do, however, demonstrate that selection on wing shape is not aligned between the sexes. Thus, the correlation between the coefficients of male and female fitness change with shape is not significantly different from zero, but is significantly smaller than unity. This result implies that selection on wing shape differs between the sexes. Consequently, the argument that divergent selection on wing shape contributes to SA fitness variation (chapter 3 section 3.4.4) is not disproved by our results here.

Overall, our experiments provide tentative support for a potential role of wing morphology in sexual antagonism in our study population. Our analysis of the genetic architecture of wing morphology provided strong evidence for positive intersexual genetic correlations between males and females for wing size and shape. Furthermore, we were able to demonstrate that despite sexual dimorphism in size and shape, allometric shape changes with increasing size are very similar in both sexes. Taken together, these results suggest that genetic correlations between the sexes could constrain the independent evolution of male and female wing morphology. A contribution of wing morphology to sexual antagonism is plausible because our fitness analysis demonstrated that selection on wing morphology is, if not opposed, at least divergent between the sexes. Although divergent selection pressures in males and females will impose less of a constraint on the simultaneous adaptation of the sexes than opposing selection, they may still limit their speed of adaptive evolution. Analysing cuticular hydrocarbons in the fly *D. serrata*, Gosden et al. (2012) showed that the rate at which males and females could adapt in response to divergent sex-specific selection pressures was constrained by a combination of inter-sexual correlations for the values of individual traits and genetic correlations between those traits within each sex. In light of these findings, the strong genetic correlation that we observe between wing size and wing shape within each sex, is likely to add to constraints on the morphological divergence between the sexes that occurs due to inter-sexual correlations for size or shape alone. Gosden et al.'s (2012) study also suggests that stronger sexual selection in males will tend to mean that correlated responses to sexually antagonistic selection will normally be more maladaptive for females than males. In relation to our findings this suggests that strong selection on male wing shape may displace females from their optimal wing morphology. However, as size is such an important component of female fecundity (Knight and Robertson 1957) this may not be the case for wing morphology in particular, as it is under strong directional selection for increased size in females. Consequently, it is possible that through the high intersexual correlation on allometry, the wing morphology of *D. melanogaster* is trapped in an evolutionary deadlock between selection on wing shape in males, and selection on wing size in females.

4.4.3 Environmental effects on size and SSD

Our measurements of male and female wing size across the three different LDs show that increased competition for food during larval development results in a corresponding

decrease in overall size. We find that due to these effects of LDs the size distribution in males and female shows three distinct peaks. Specifically, this demonstrates that changes in the level of competition for finite amount of food during larval development has a strong effect on the size of both sexes. Here we uphold the general consensus that the volume of food available is one of the most important determinants of adult size in ectothermic animals (Nylin and Gotthard 1998). In holometabolous species, such as *D. melanogaster*, adult size is completely determined by growth during larval development (Ashburner and Thompson 1978). In particular, the effects of food restriction during larval stages via manipulation of LD have been repeatedly shown to reduce the size of both sexes in *D. melanogaster* (Sang 1949; Lints and Lints 1969; Santos et al. 1994), just as we observe here.

As is normal for *D. melanogaster* we find that SSD is female biased. Specifically we find that the greatest SSD occurs within the intermediate LD, which best matches the normal rearing conditions which our nine TGs have adapted for. Here, we find that females are on average 1.156 times larger than males. The scale of this SSD is extremely consistent with that previously measured for other traits that have been used to estimate SSD in *D. melanogaster,* such as wing length (1.16) and thorax length (1.15) (David et al. 2003). It is notable that we find the magnitude of SSD varies significantly between the different LDs that we imposed. Although we demonstrate that this is not just a by-product of the actual size differences between LDs, it is important to emphasize that the relative scale of the SSD differences between LDs is small compared with the size differences generated by the different LDs themselves. As a result, the proportion of total size variation explained by SSD is also extremely low. We also note that the males and females, which we compare to calculate SSD, are not reared from the same vials, nor from equivalent parental crosses. In particular, the LDs applied to males were increased in order to account for early larval death during development (caused by a triple X chromosome karyotype - see Appendix 1). In comparing SSD across LDs it is important that the rearing conditions experienced by males and females are equivalent. Our analysis of the total number of emerging adults from the intermediate LD, suggests that the rearing conditions experienced by each sex are almost exactly equivalent for this treatment.

Although we have no data on the total number of emerging adults at the high LD, because the male:female rearing ratio is very similar to that of the intermediate LD, we

can assume that the rearing conditions at high LD are also comparable between the sexes. Finally, there are numerical differences between the low LD treatments that were applied to each sex. Nevertheless, the purpose of these particular treatments is to provide conditions where there is minimal larval competition. Typically, larval survival can be used as a measure of competition amongst larvae. Importantly, our low LDs for both sexes fall within the minimal mortality range predicted by Sang (1949). Specifically, from the equivalent LDs imposed by Sang (1949), the difference in mortality between our male and female low LD treatments would be ~ 1.4%. We can therefore assume that these low LDs provide equivalent larval conditions for each sex. Although it is unlikely that variation between the larval treatments applied to each sex explains all of the variation in SSD that we observe between LDs, we cannot be totally sure that the observed pattern is not, in part, the result of slightly different rearing conditions between the sexes. Typically, differences in SSD caused by environmental variation are attributed to differences in the phenotypic plasticity of body size between the sexes. For example, the seed beetle *C. maculatus* shows pronounced differences in phenotypic robustness to variation in rearing temperature. Males reared at 20°C were on average 63% larger than those reared at 35°C, whereas females were 38% larger across the same temperature range (Stillwell and Fox 2007). Interestingly, an equivalent explanation does not suffice to explain the differences that we observe for SSD in response to nutritional manipulation. Specifically we observe an opposite pattern by which phenotypic robustness for size appears not sex-specific, but instead largely determined by the fitness effects of each TG.

4.4.4 Phenotypic robustness for size

An interesting finding, which builds on our results from chapter 3, is that this phenotypic robustness for size appears linked to the fitness effects of a given TG. In chapter 3 we found that across environments, MB genomes conferred relatively greater robustness for size in males, compared with FB genomes, which conferred relatively low robustness for size in males (section 3.4.6). Within the same data set we also observed a negative association between male fitness and the size differential (used as a measure of robustness) between low and high LDs (Fig. 3.2), which suggested that phenotypic robustness for size could be an important component of male fitness. In the previous study, where we only had size data for TG expression in males, we demonstrate that relatively high phenotypic robustness of MB genomes was a trait directly associated with high male fitness, or the result of a heavy SA load in males

expressing FB genomes, thus making them relatively less robust (Fig. 3.4). By expressing the same MB genomes in females and subjecting them to the same range of rearing conditions we demonstrate here that phenotypic robustness for size within our sample of genomes is not sex specific, but is in fact ubiquitously expressed between the sexes (Fig. 4.2). Specifically, our results show that high MB genomes confer relatively high phenotypic robustness for size to both sexes. Conversely, we find that FB genomes provide relatively low robustness for size when expressed in either sex.

These findings suggest that a high phenotypic robustness for size is indeed an important component of male fitness. This provides evidence that adds to the general idea that males are selected towards greater phenotypic robustness such that they are buffered from changes that could displace them from their optimal trait values (L Partridge personal communication, Nunney 1996). Here, we relate this 'selection for perfection' theory to male size. Specifically, we have demonstrated that when males are smaller than their immediate competition then it can have a high reproductive cost. Hence robustness for size is likely to be an important component of male fitness.

Our results also provide some support for theories explaining the extent of phenotypic robustness for particular traits. The greater robustness conferred by MB genomes supports the 'adaptive canalization hypothesis', which predicts that phenotypic robustness should be greatest for traits that are subject to the strongest selection pressure (Fairbairn 2005). In our case this trait under strong selection could be male size. We demonstrated the importance of male size at the level of the individual in chapter 3, specifically showing that male mating success increased when a male was relatively larger than the male that they were in immediate competition with. This is especially interesting because increased robustness for size in males would ensure that a male is able to maximize his relative size compared with other males that are reared under the same conditions. This would theoretically ensure that male with high phenotypic robustness for size can maintain this relative size advantage with local competitors. However, Stillwell and Fox (2009) predicted that traits subject to the strongest stabilizing selection should be the most phenotypically robust (least plastic), and instead traits subject to the strongest directional selection should in fact show the most phenotypic plasticity (Cotton et al. 2004; Bonduriansky 2007). This alternative, know as the 'condition dependence hypothesis' suggests that individuals will invest a substantial proportion of resources into a trait, such as size, if it is important for fitness.

However, this heavily weighted resource allocation can make the trait highly dependent on resource availability, and therefore highly plastic in response to shortages of resources. We observe a similar pattern for FB genomes, which show relatively high plasticity compared with MB genomes. With our limited sample size it is not possible to determine whether the differences in phenotypic robustness we observe between MB and FB genomes are driven by either or both of the mechanism postulated by these hypotheses. Gaining a more detailed understanding of the relationship between phenotypic robustness and sex specific fitness would require a larger sample of genomes with known fitness. In particular, a future experiment could look to detect a difference between the mean level of phenotypic size robustness across a larger sample, and compare this with the robustness of similar MB and FB genomes. If adaptive canalization mediated by MB genomes causes the observed divergence in robustness then we would expect MB genomes to show significantly greater robustness relative to all other genomes. Conversely, if condition dependence mediated by FB genomes causes the observed divergence in robustness, then we would expect FB genomes to display significantly lower robustness relative to all other genomes.

The finding that robustness for size is not differentially expressed between the sexes is consistent with the positive genetic correlation for size between the sexes. Specifically, the correlated size effects that we observe between males and females can be generally attributed to a sharing of the same genes that control growth and development (Roff 1997; Nijhout 2003), which ultimately determine the adult size of flies. Here we provide further evidence, suggesting that the genes involved in determining the phenotypic response of size to environmental stress are also shared between sexes. However, our findings specifically contradict a quantitative genetic study by Fraser & Schadt (2010) in mice. These authors showed that the expression of hundreds of alleles that affect phenotypic robustness were polymorphic, and could be mapped to discrete genomic loci. They then demonstrated that the expression of these alleles involved in phenotypic robustness with environmental variation is predominantly sex-specific, and hence should result in differential patterns of phenotypic robustness between males and females. It is also worth noting that results from male-limited genomes in *D. melanogaster* also imply that loci responsible for environmental robustness should be sex-specific in their expression (Abbott et al. 2010). However, this latter experiment by Abbott et al. (2010) is not directly comparable to ours. Their measurement of phenotypic robustness used fluctuating asymmetry (FA) of wings in a single

environment. This measurement of micro-environmental developmental stability is not necessarily comparable to our measures of phenotypic change that result from environmental manipulation.

Due to the differing phenotypic robustness for size conferred by MB and FB genomes, we find that where flies are under the most nutritional stress (high LD), females expressing FB genomes are on average smaller than those expressing MB genomes. However, despite FB genomes conferring relatively small size at high LDs, they still maintain their relative fitness rank (Fig. 4.6). This suggests that FB genomes encode traits other than large body size that allow them to gain a reproductive advantage over other females. Previous studies have also demonstrated a breakdown in the link between fecundity and size, when size is altered by environmental manipulation. For example, Nunney and Cheung (1997) used manipulation of rearing temperature, a factor that is negatively correlated with size, to environmentally alter the size of female *D. melanogaster* upon eclosion. They then showed that females reared at lower temperatures had increased adult size and did not experience any corresponding increase in lifetime fecundity, as would be expected based on the usual assumption that size correlates positively with fecundity. This study corroborates our observation, and collectively the two similar observations highlight the condition dependence of the fitness effects that are associated with size.

In our results the inversion of female size relative to female fitness generates a negative genetic correlation between female fitness and female size within the high LD ($r = -0.704$, $n = 9$, $P = 0.034$), against the classic result that female fitness is positively correlated with size in *D. melanogaster* (Knight and Robertson 1957) and many other insect species (e.g. Honek 1993). This is especially intriguing given that unpublished data (Reuter et al.) on a slightly larger subset of TGs, including the nine that we have used, has shown a positive genetic correlation for female size with female fitness at low LD. The results also provide evidence for the reversal of a genetic correlation between life traits at different LDs, and is indicative of the sensitivity of genetic trait correlations to environmental variation. For example, a previously well-established negative correlation between early fertility and late fertility in female *D. melanogaster* first identified by Rose (1984), can disappear (Chippindale et al. 1993; Leroi et al. 1994a) and re-appear (Leroi et al. 1994b) depending on subtle environmental variations. Here we provide evidence that further stresses the importance of considering the

environmental context of interactions between life-history traits (Chippindale et al. 2003; Prasad and Shakarad 2004).

4.4.5 Development time

In *D. melanogaster* and most other insect species, the size of both sexes has been shown to positively correlate with development time (Roff 1980; Fairbairn 1990; Partridge and Fowler 1992; Zwaan et al. 1995). Our results are in line with this broad consensus, showing a positive association between longer development time and larger size across sexes and environments. We also find that there is a strong genetic component to development time, which has the same effect in both sexes. As above, this highlights the fact that the genes controlling growth and development are largely shared between the sexes (Roff 1997; Nijhout 2003). In other words, it takes time to get large for both sexes, and hence the genes determining size are predominantly involved in controlling development time.

We also find that males develop for longer than females, which is a well known phenomenon in *D. melanogaster*. Why the smaller sex should take longer to develop is not well understood. Certainly a large proportion of the difference is attributable to females having a faster growth rate than males. This has been recorded at 25% to 33% more dry mass per hour of development (Chippindale et al. 2003). However, in the context of differing developmental rates between the sexes it is important to explain why females grow so much faster than males. Roper et al. (1993) proposed that females could be subject to greater selection for early eclosion. Their argument, relating to developmental rates, specifically suggested that males were under less intense selection for rapid development than females because they take less time to become reproductively mature after eclosion. Roper et al. (1993) predicted that females are under comparatively strong selection for fast development so that they can eclose and mature in time to maximize their reproductive success. But this argument of relative developmental rates fails to apply where generations overlap. There would then be no conceivable advantage accruing from relatively early eclosion, and would therefore not be applicable to the natural life history of *D. melanogaster*.

An alternative, and more robust idea, is the selection for perfection hypothesis. This suggests that males should be subject to selection towards more accurate development, and are hence more perfectly formed, with likely benefits in intrasexual competition to

acquire matings (Chippindale et al. 2003; Nunney 1996). Our data support the basic premise of this hypothesis, in that male-beneficial genomes appear to have greater phenotypic robustness for size in the face of environmental stress. However, analysis of our development time data suggests that higher male fitness is in fact associated with shorter development time (Table 4.7). This directly contradicts the selection for perfection hypothesis, and instead suggests that males within our population are selected for faster development. In particular, research by Shakarad et al. (2001) on *D. melanogaster* failed to show that developmental stability (estimated via fluctuating asymmetry) decreased in response to artificial selection for rapid development in males. Our results generate a similar picture in which phenotypic robustness for size does not appear directly related to development time.

One possible explanation for this paradoxical result could come from the selection dynamics imposed by the rearing regime of the population (LHm) from which our sample of TGs was derived (see section 3.2.1). Of note, the LHm rearing regime has non-overlapping generations. As a result, males that eclose first should have the greatest access to the most virgin females, which will readily mate. However, such an explanation is unsatisfactory because research on the patterns of sperm precedence in the LHm population suggest that early mating males rarely sire any offspring (Morrow et al. 2005). Instead, mating success and mating order largely determine the overall reproductive success of LHm males. In other words, the last successful mating usually sires the most offspring for a given female (Pischedda and Rice 2012). An alternative possibility is that robustness for size is not related to developmental rate, but instead to the genetic mechanism that determines the threshold of critical weight for pupation (De Moed et al. 1999). In other words, robustness for size when under nutritional stress could be the result of developing for longer to reach a larger larval size before pupation. Specifically this would mean that individuals with higher robustness would develop for longer when nutrition is limited during the larval life stage. Certainly further research into the development of *D. melanogaster* and the links between development time, size and fitness is warranted. In particular future studies should focus on the association between developmental rate and its effect on both size and fitness, with a view to better understanding their inter-relationships.

4.4.6 Conclusions

The primary aim of this study was to establish whether intersexual genetic correlations for wing morphology contribute to SA fitness variation in *D. melanogaster*. We show clear evidence of genetic correlations for both wing size and wing shape between the sexes, which serves as evidence that there is some constraint on the independent evolution of wing morphologies between the sexes. We also demonstrate that selection on wing morphology is, at least in part, divergent between the sexes. This suggests that the genetic architecture and selective pressures necessary for wing morphology to contribute to SA fitness variation are present in our sample. More data that directly relates fitness effects with specific wing morphologies across a larger sample is required to make more accurate predictions about the selective pressures operating on the morphology of both male and female wings.

We also find a strong indication that phenotypic robustness of size is related to the fitness effects of a given genome, and that this robustness is non sex-specific. In particular our data indicates that high robustness may be associated with high male genetic fitness and that low robustness may be associated with high female genetic fitness.

Finally, we found that male fitness appears to be associated with a faster larval development time. This is the opposite result to what we expected based on a priori predictions in relation to the selection for perfection hypothesis, according to which higher male fitness is theoretically associated with slower development. A clear explanation for this finding is difficult to establish, and further study of the sex-specific effects of development rate on size and fitness of males is required.

4.5 References

Abbott JK, Bedhomme S, Chippindale AK (2010) Sexual conflict in wing size and shape in *Drosophila melanogaster*. J Evolution Biol 23(9): 1989-1997.

Ashburner and Thompson (1978) The laboratory culture of *Drosophila*. In: The Genetics and Biology of *Drosophila* Vol. 2a. Edited by M. Ashburner and T. R. F. Wright. Academic Press, London & New York. pp. 1-109.

Birdsall K, Zimmerman E, Teeter K, Gibson G (2000) Genetic variation for the positioning of wing veins in *Drosophila melanogaster*. Evol Dev 2(1): 16-24.

Bonduriansky R (2007) The evolution of condition-dependent sexual dimorphism. Am Nat 169(1): 9-19.

Brommer JE, Kirkpatrick M, Qvarnstrom A, Gustafsson L (2007) The intersexual genetic correlation for lifetime fitness in the wild and its implications for sexual selection. Plos One 2(8).

Bruner E, Costantini D, Fanfani A, Dell'Omo G (2005) Morphological variation and sexual dimorphism of the cephalic scales in *Lacerta bilineata*. Acta Zool-Stockholm 86(4): 245-254.

Chapman T, Arnqvist G, Bangham J, Rowe L (2003) Sexual conflict. Trends Ecol Evol 18(1): 41-47.

Charnov EL, Losdenhartogh RL, Jones WT, Vandenassem J (1981) Sex ration evolution in a variable environment. Nature 289(5793): 27-33.

Chippindale AK, Gibson JR, Rice WR (2001) Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. P Natl Acad Sci USA 98(4): 1671-1675.

Chippindale AK, Ngo AL, Rose MR (2003) The devil in the details of life-history evolution: instability and reversal of genetic correlations during selection on *Drosophila* development. J Genet 82(3): 133-145.

Chippindale AK, Leroi AM, Kim SB, Rose MR (1993) Phenotypic plasticity and selection in *Drosophila* life-history evolution .1. Nutrition and the Cost of Reproduction. J Evolution Biol 6(2): 171-193.

Cotton S, Fowler K, Pomiankowski A (2004) Heightened condition dependence is not a general feature of male eyespan in stalk-eyed flies (Diptera : Diopsidae). J Evolution Biol 17(6): 1310-1316.

Cowley DE, Atchley WR, Rutledge JJ (1986) Quantitative genetics of *Drosophila melanogaster* .1. Sexual dimorphism in genetic-parameters for wing traits. Genetics 114(2): 549-566.

Cowling DE, Burnet B (1981) Courtship songs and genetic-control of their acoustic characteristics in sibling species of the *Drosophila melanogaster* subgroup. Anim Behav 29(Aug): 924-935.

David J, Legay JM (1977) Relationship between genetic variations of egg and female sizes - comparison of 3 geographic races of *Drosophila melanogaster*. Arch Zool Exp Gen 118(3): 305-314.

David JR, Gibert P, Mignon-Grasteau S, Legout H, Petavy G et al. (2003) Genetic variability of sexual size dimorphism in a natural population of *Drosophila melanogaster*: an isofemale-line approach. J Genet 82(3): 79-88.

De Moed GH, Kruitwagen CLJJ, De Jong G, Scharloo W (1999) Critical weight for the induction of pupariation in *Drosophila melanogaster*: genetic and environmental variation. J Evolution Biol 12(5): 852-858.

Fairbairn DJ (1990) Factors influencing sexual size dimorphism in temperate waterstriders. Am Nat 136(1): 61-86.

Fairbairn DJ (1997) Allometry for sexual size dimorphism: Pattern and process in the coevolution of body size in males and females. Annu Rev Ecol Syst 28: 659-687.

Fairbairn DJ (2005) Allometry for sexual size dimorphism: Testing two hypotheses for Rensch's rule in the water strider *Aquarius remigis*. Am Nat 166(4): S69-S84.

Fairbairn DJB, W. U. Szekely, T. (2007) Sex, size & gender roles: Evolutionary studies of sexual size dimorphism: Oxford University Press.

Felsenstein J (1976) Theoretical population genetics of variable selection and migration. Annu Rev Genet 10: 253-280.

Foerster K, Coulson T, Sheldon BC, Pemberton JM, Clutton-Brock TH et al. (2007) Sexually antagonistic genetic variation for fitness in red deer. Nature 447(7148): 1107-U1109.

Fraser HB, Schadt EE (2010) The quantitative genetics of phenotypic robustness. Plos One 5(1).

Gavrilets S, Rice WR (2006) Genetic models of homosexuality: generating testable predictions. P Roy Soc B-Biol Sci 273(1605): 3031-3038.

Gidaszewski NA, Baylac M, Klingenberg CP (2009) Evolution of sexual dimorphism of wing shape in the *Drosophila melanogaster* subgroup. Bmc Evol Biol 9.

Gilbert J, Woods SM, Tekauz A (1998) Relationship between environmental variables and the prevalence and isolation frequency of leaf-spotting pathogens in spring wheat. Can J Plant Pathol 20(2): 158-164.

Gilchrist AS, Azevedo RBR, Partridge L, O'Higgins P (2000) Adaptation and constraint in the evolution of *Drosophila melanogaster* wing shape. Evol Dev 2(2): 114-124.

Gillespie JH, Turelli M (1989) Genotype-environment interactions and the maintenance of polygenic variation. Genetics 121(1): 129-138.

Gosden TP, Shastri KL, Innocenti P, Chenoweth SF (2012) The B-Matrix harbors significant and sex-specific constraints on the evolution of multicharacter sexual dimorphism. Evolution 66(7): 2106-2116.

Greenspan RJ, Ferveur JF (2000) Courtship in *Drosophila*. Annu Rev Genet 34: 205-232.

Hansen TF, Houle D (2008) Measuring and comparing evolvability and constraint in multivariate characters. J Evolution Biol 21(5): 1201-1219.

Honek A (1993) Intraspecific variation in body size and fecundity in insects - a general relationship. Oikos 66(3): 483-492.

Karan D, Morin JP, Gravot E, Moreteau B, David JR (1999) Body size reaction norms in *Drosophila melanogaster*: temporal stability and genetic architecture in a natural population. Genet Sel Evol 31(5-6): 491-508.

Knight GR, Robertson A (1957) Fitness as a measurable character in *Drosophila*. Genetics 42(4): 524-530.

Leroi AM, Chippindale AK, Rose MR (1994a) Long-term laboratory evolution of a genetic life-history trade-off in *Drosophila melanogaster* .1. The role of genotype-by-environment interaction. Evolution 48(4): 1244-1257.

Leroi AM, Chen WR, Rose MR (1994b) Long-term laboratory evolution of a genetic life-history trade-off in *Drosophila melanogaster* .2. Stability of genetic correlations. Evolution 48(4): 1258-1268.

Lints FA, Lints CV (1969) Respiration in *Drosophila* .3. Influence of preimaginal environment on respiration and ageing in *Drosophila melanogaster* hybrids. Exp Gerontol 4(2): 81-&.

Long TAF, Rice WR (2007) Adult locomotory activity mediates intralocus sexual conflict in a laboratory-adapted population of *Drosophila melanogaster*. P Roy Soc B-Biol Sci 274(1629): 3105-3112.

Mainguy J, Cote SD, Festa-Bianchet M, Coltman DW (2009) Father-offspring phenotypic correlations suggest intralocus sexual conflict for a fitness-linked trait in a wild sexually dimorphic mammal. P Roy Soc B-Biol Sci 276(1675): 4067-4075.

Markow TA, Ricker JP (1992) Male size, developmental stability, and mating success in natural-populations of 3 *Drosophila* species. Heredity 69: 122-127.

Matos M, Rose MR, Pite MTR, Rego C, Avelar T (2000) Adaptation to the laboratory environment in *Drosophila subobscura*. J Evolution Biol 13(1): 9-19.

Messina FJ (2004) Predictable modification of body size and competitive ability following a host shift by a seed beetle. Evolution 58(12): 2788-2797.

Mezey JG, Houle D (2005) The dimensionality of genetic variation for wing shape in *Drosophila melanogaster*. Evolution 59(5): 1027-1038.

Mokkonen M, Kokko H, Koskela E, Lehtonen J, Mappes T et al. (2011) Negative frequency-dependent selection of sexually antagonistic alleles in *Myodes glareolus*. Science 334(6058): 972-974.

Morrow EH, Stewart AD, Rice WR (2005) Patterns of sperm precedence are not affected by female mating history in *Drosophila melanogaster*. Evolution 59(12): 2608-2615.

Nijhout H (2003) The control of body size in insects. Dev Biol 261(1): 1-9.

Nijhout HF, Reed MC, Anderson DF, Mattingly JC, James SJ et al. (2006) Long-range allosteric interactions between the folate and methionine cycles stabilize DNA methylation reaction rate (vol 2, pg 81, 2006). Epigenetics 1(3): 115-115.

Nunney L (1983) Sex-differences in larval competition in *Drosophila melanogaster* - the testing of a competition model and its relevance to frequency-dependent selection. Am Nat 121(1): 67-93.

Nunney L (1996) The response to selection for fast larval development in *Drosophila melanogaster* and its effect on adult weight: An example of a fitness trade-off. Evolution 50(3): 1193-1204.

Nunney L, Cheung W (1997) The effect of temperature on body size and fecundity in female *Drosophila melanogaster*: Evidence for adaptive plasticity. Evolution 51(5): 1529-1535.

Nylin S, Gotthard K (1998) Plasticity in life-history traits. Annu Rev Entomol 43: 63-83.

O'higgins P, Collard M (2002) Sexual dimorphism and facial growth in papionin monkeys. J Zool 257: 255-272.

Ohiggins P, Johnson DR, Moore WJ, Flinn RM (1990) The variability of patterns of sexual dimorphism in the hominoid skull. Experientia 46(7): 670-672.

Partridge L, Fowler K (1992) Direct and correlated responses to selection on age at reproduction in *Drosophila melanogaster*. Evolution 46(1): 76-91.

Partridge L, Fowler K, Trevitt S, Sharp W (1986) An examination of the effects of males on the survival and egg-production rates of female *Drosophila melanogaster*. J Insect Physiol 32(11): 925-929.

Pischedda A, Rice WR (2012) Partitioning sexual selection into its mating success and fertilization success components. P Natl Acad Sci USA 109(6): 2049-2053.

Prasad NG, Shakarad MN (2004) Genetic correlations: transient truths of adaptive evolution. J Genet 83(1): 3-6.

Prasad NG, Bedhomme S, Day T, Chippindale AK (2007) An evolutionary cost of separate genders revealed by male-limited evolution. Am Nat 169(1): 29-37.

Rice WR (1984) Sex chromosomes and the evolution of sexual dimorphism. Evolution 38(4): 735-742.

Robertson FW (1953) Mass selection for large and small size in *Drosophila*. Heredity 7(3): 443-443.

Roff D (1980) Optimizing development time in a seasonal environment - the ups and downs of clinal variation. Oecologia 45(2): 202-208.

Roff D (1997) Evolutionary Quantitative Genetic: Springer-Verlag GmbH.

Roper C, Pignatelli P, Partridge L (1993) Evolutionary effects of selection on age at reproduction in larval and adult *Drosophila melanogaster*. Evolution 47(2): 445-455.

Rose MR (1984) Genetic covariation in *Drosophila* life history - Untangling the data. Am Nat 123(4): 565-569.

Sang JH (1949) The ecological determinants of population growth in a *Drosophila* Culture .3. Larval and pupal survival. Physiol Zool 22(3): 183-202.

Sang JH (1950) Population growth in *Drosophila* Cultures. Biol Rev 25(2): 188-219.

Santos M, Fowler K, Partridge L (1994) Gene-environment interaction for body size and larval density in *Drosophila melanogaster* - an investigation of effects on development time, thorax length and adult sex ratio. Heredity 72: 515-521.

Sgro CM, Partridge L (2000) Evolutionary responses of the life history of wild-caught *Drosophila melanogaster* to two standard methods of laboratory culture. Am Nat 156(4): 341-353.

Shakarad M, Prasad NG, Rajamani M, Joshi A (2001) Evolution of faster development does not lead to greater fluctuating asymmetry of sternopleural bristle number in *Drosophila*. J Genet 80(1): 1-7.

Snook RR, Robertson A, Crudgington HS, Ritchie MG (2005) Experimental manipulation of sexual selection and the evolution of courtship song in *Drosophila pseudoobscura*. Behav Genet 35(3): 245-255.

Stillwell RC, Fox CW (2007) Environmental effects on sexual size dimorphism of a seed-feeding beetle. Oecologia 153(2): 273-280.

Stillwell RC, Fox CW (2009) Geographic variation in body size, sexual size dimorphism and fitness components of a seed beetle: local adaptation versus phenotypic plasticity. Oikos 118(5): 703-712.

Svensson EI, McAdam AG, Sinervo B (2009) Intralocus sexual conflict over immune defense, gender load, and sex-specific signaling in a natural lizard population. Evolution 63(12): 3124-3135.

Tantawy AO, Vetukhiv MO (1960) Effects of size on fecundity, longevity and viability in populations of *Drosophila pseudoobscura*. Am Nat 94(879): 395-403.

Weber KE (1990) Selection on wing allometry in *Drosophila melanogaster*. Genetics 126(4): 975-989.

Zwaan B, Bijlsma R, Hoekstra RF (1995) Artificial selection for developmental time in *Drosophila melanogaster* in relation to the evolution of aging - direct and correlated responses. Evolution 49(4): 635-648.

4.6 Tables and Figures

Table 4.1 Female fitness. Results of a two-way ANOVA on the dependent variable female fitness with the variables, LD, genome and their interaction (A), followed by a two-way ANOVA of female fitness with the variables LD, fitness class and their interaction (B).

A.

	Df	*Sum Sq*	*F-value*	*P-value*
LD	2	192425	59.033	<0.001
Genome	8	162850	12.49	<0.001
LD*Genome	16	66638	2.556	0.005
Residuals	54	88009		

B.

	Df	*Sum Sq*	*F-value*	*P-value*
LD	2	42500	20.0919	<0.001
Fitness class	2	64142	30.3228	<0.001
LD*Fitness class	4	14958	3.5357	0.027
Residuals	18	19038		

Table 4.2 Wing size in both sexes. Results of a pair of three-way ANOVA models fitting the dependent variable wing size, with the independent variables, sex, genome, LD, and their interactions. Initial analysis used the absolute wing size values (A), followed by an analysis that used measures of wing size, relative to the mean of each sex at each LD (B).

A.

	Df	*Sum Sq*	*F-value*	*P-value*
Sex	1	6501762	9344.539	<0.001
Genome	8	577949	103.831	<0.001
LD	2	6136543	4409.818	<0.001
Sex*Genome	8	23453	4.213	<0.001
Sex*LD	2	116131	83.454	<0.001
Genome*LD	16	119106	10.699	<0.001
Sex*Genome*LD	16	48671	4.371	<0.001
Residuals	1946	1353992		

B.

Transformed size scores

	Df	*Sum Sq*	*F-value*	*P-value*
Sex	1	6.4482	10251.654	<0.001
Genome	8	0.2908	57.792	<0.001
LD	2	0.0167	13.277	<0.001
Sex*Genome	8	0.022	4.375	<0.001
Sex*LD	2	0.0461	36.683	<0.001
Genome*LD	16	0.1003	9.961	<0.001
Sex*Genome*LD	16	0.0443	4.401	<0.001
Residuals	1946	1.224		

Table 4.3 Wing size, effect of sex-specific fitness classes. Results of a pair of three-way ANOVA models fitting the dependent variable wing size with the independent variables fitness class, sex, and LD. Initial analysis (A) included all three fitness classes (N, MB, and FB). A subsequent analysis (B) excluded genomes of the N fitness class, in order to enable direct comparisons between the fitness classes of extreme sex specific fitness, MB and FB.

A.

	Df	*Sum Sq*	*F-value*	*P-value*
Fitness class	2	1645	2.944	0.065
Sex	1	246515	882.228	<0.001
LD	2	157725	282.234	<0.001
Fitness class*Sex	2	282	0.504	0.608
Fitness class*LD	4	2312	2.068	0.105
Sex*LD	2	3485	6.236	0.004
Fitness class*Sex*LD	4	880	0.787	0.541
Residuals	36	10059		

B.

	Df	*Sum Sq*	*F-value*	*P-value*
Fitness class	1	1093	5.774	0.024
Sex	1	170084	898.437	<0.001
LD	2	94976	250.847	<0.001
Fitness class*Sex	1	134	0.707	0.408
Fitness class*LD	2	1498	3.955	0.032
Sex*LD	2	1286	3.395	0.05
Fitness class*Sex*LD	2	212	0.558	0.579
Residuals	24	4543		

Table 4.4 Overall variation of wing shape. Results of MANOVA using the 22 PC axes as the dependent variable with the covariate term wing size, the terms sex, LD, genome, plus all their interactions, and the interaction terms wing size-by-sex, and wing size-by-genome. Values in the percentage (%) explained variance column are estimates of the percentage of total wing shape variance explained by the corresponding term from 22 univariate ANOVA models (see Methods, section 4.2.7 for more detail).

	Df	*Pillai*	*Approx-F*	*P-value*	*% explained variance*
Wing Size	1	0.8411	457.7	<0.001	17.36
Sex	1	0.974	3254.3	<0.001	3.99
LD	2	1.1896	128.1	<0.001	1.35
Genome	8	3.5477	69.5	<0.001	22.58
Sex*LD	2	1.144	116.7	<0.001	0.51
Sex*Genome	8	1.1817	15.1	<0.001	1.73
LD*Genome	16	1.3513	8.1	<0.001	1.11
Wing size*Gex	1	0.4464	70.2	<0.001	0.08
Wing size*Genome	8	0.6674	8	<0.001	0.23
Sex*LD*Genome	16	1.2161	7.3	<0.001	0.57
Residuals	1936				50.44

Table 4.5 Fitness effects of sex-specific wing morphologies. Table showing the final ANOVA models on the dependent variable fitness with the independent variables LD, mean TG wing size and mean PC scores across 22 axes (representing wing shape). This model was applied to the male (A) and female (B) data separately. For each model the proportion of fitness variation explained by TG wing size, and TG wing shape ("Explained variance") is calculated from the corresponding sum of squares values.

A. Fitness effects of male wing morphology

	Df	*Sum Sq*	*F-value*	*P-value*	
LD	2	0.233664	21.103	0.0036512	**
PC1	1	0.012748	2.3026	0.1896085	
PC2	1	0.050861	9.1868	0.0290489	*
PC3	1	0.012769	2.3064	0.189307	
PC5	1	0.02751	4.969	0.0762485	.
PC6	1	0.204435	36.9265	0.001744	**
PC7	1	0.007365	1.3304	0.3008719	
PC8	1	0.298649	53.9441	0.0007342	***
PC9	1	0.045436	8.207	0.035209	*
PC10	1	0.00723	1.3059	0.3048945	
PC11	1	0.023951	4.3262	0.0920676	.
PC12	1	0.108224	19.5482	0.0068834	**
PC13	1	0.00352	0.6358	0.4614145	
PC14	1	0.004681	0.8455	0.4000236	
PC15	1	0.005779	1.0439	0.3537883	
PC16	1	0	0	0.9947623	
PC17	1	0.000526	0.095	0.7703033	
PC18	1	0.063838	11.5309	0.0193408	*
PC19	1	0.090926	16.4237	0.0098001	**
PC20	1	0.044441	8.0272	0.0365354	*
Residuals	5	0.027681			
Explained variance: shape		**0.89502713**			

B. Fitness effects of female wing morphology

	Df	*Sum Sq*	*F-value*	*P-value*	
LD	2	64142	70.6594	8.24E-06	***
Mean wing size	1	2509	5.5284	0.046589	*
PC1	1	10201	22.4759	0.001462	**
PC2	1	1371	3.0206	0.120414	
PC3	1	951	2.0945	0.185856	
PC5	1	5813	12.8063	0.007203	**
PC9	1	802	1.7681	0.220283	
PC10	1	44	0.0959	0.764718	
PC12	1	30374	66.9197	3.72E-05	***
PC13	1	750	1.6526	0.234562	
PC14	1	843	1.8572	0.210066	
PC15	1	3026	6.667	0.032512	*
PC16	1	385	0.8493	0.3837	
PC17	1	4037	8.8941	0.017539	*
PC19	1	3606	7.9442	0.022544	*
PC20	1	1487	3.2767	0.107863	
PC21	1	6666	14.6876	0.005001	**
Residuals	8	3631			
Explained variance: size		**0.017840129**			
Explained variance: shape		**0.453099447**			

Table 4.6 Analysis of development time variation within and between the sexes. Initially we modelled development time as the dependent variable in a three-way ANOVA with the independent variables, genome, LD, sex, and their interactions (A). Subsequently, we used three-way ANOVA on the dependent variable mean development time, with the independent variables mean wing size, sex, mean fitness, and their interactions (B). This latter analysis was divided into two separate models within the low LD and intermediate LD.

A.

	Df	*Sum Sq*	*F-value*	*P-value*
Genome	8	384.98	3.0993	0.003
LD	1	1464.48	94.3193	< 0.001
Sex	1	2868.79	184.7631	< 0.001
Genome*LD	8	84.81	0.6827	0.706
Genome*Sex	8	130.32	1.0491	0.402
LD*Sex	1	5.6	0.3606	0.549
Genome*LD*Sex	8	81.2	0.6537	0.731
Residuals	143	2220.34		

B.

Low LD

	Df	*Sum Sq*	*F-value*	*P-value*
Mean wing size	1	280.495	49.4107	< 0.001
Sex	1	36.27	6.3892	0.029
Mean fitness	1	0.276	0.0486	0.83
Mean wing size*Sex	1	0.201	0.0354	0.854
Mean wing size*Mean fitness	1	0.567	0.0998	0.758
Sex*Mean fitness	1	5.279	0.9299	0.357
Mean wing size*Sex*Mean fitness	1	2.668	0.47	0.508
Residuals	10	56.768		

Intermediate LD

	Df	*Sum Sq*	*F-value*	*P-value*
Mean wing size	1	261.719	68.428	< 0.001
Sex	1	1.17	0.3059	0.592
Mean fitness	1	3.67	0.9595	0.35
Mean wing size*Sex	1	14.51	3.7937	0.08
Mean wing size*Mean fitness	1	0.818	0.2139	0.653
Sex*Mean fitness	1	14.601	3.8174	0.079
Mean wing size*Sex*Mean fitness	1	6.019	1.5737	0.238
Residuals	10	38.247		

Table 4.7 Fitness effects of developmental rate. Results of a pair of ANOVAs fitting the dependent variable mean fitness, with the variables LD and a measure of developmental rate (residuals of a linear model of development time as a function of size and LD). Data on each sex was analysed independently..

Male

	Df	*Sum Sq*	*F-value*	*P-value*
Developmental rate	1	0.13317	3.4356	0.083
LD	1	0.02546	0.6568	0.43
Residuals	15	0.58141		

Female

	Df	*Sum Sq*	*F-value*	*P-value*
Developmental rate	1	4070	0.9523	0.344
LD	1	21563	5.0453	0.04
Residuals	15	64108		

Table 4.8 Comparison of larval treatments applied to each sex. Results of a pair of Poisson GLMs (log link function) using the total number of emerging adults as the dependent variable, with the independent variables sex and genome and their interaction. Chi-squared tests were used to determine statistical significance. The two models were fitted within the low and intermediate LDs separately.

Low LD

	DF	*Deviance*	*Resid. Dev*	*P-Value*
Sex	1	38.751	159.45	<0.001
Genome	1	18.265	141.18	0.019
Sex*Genome	8	15.456	125.73	0.05

Intermediate LD

	DF	*Deviance*	*Resid. Dev*	*P-Value*
Sex	1	0.4549	81.582	0.5
Genome	1	23.0902	58.492	0.003
Sex*Genome	8	10.0213	48.471	0.263

Figure 4.1 Plot of mean female fitness per genome, across the three LDs. Line colour represents the fitness class of each genome (FB = red, MB = blue, N = black), and error bars represent the standard deviation for each mean fecundity score.

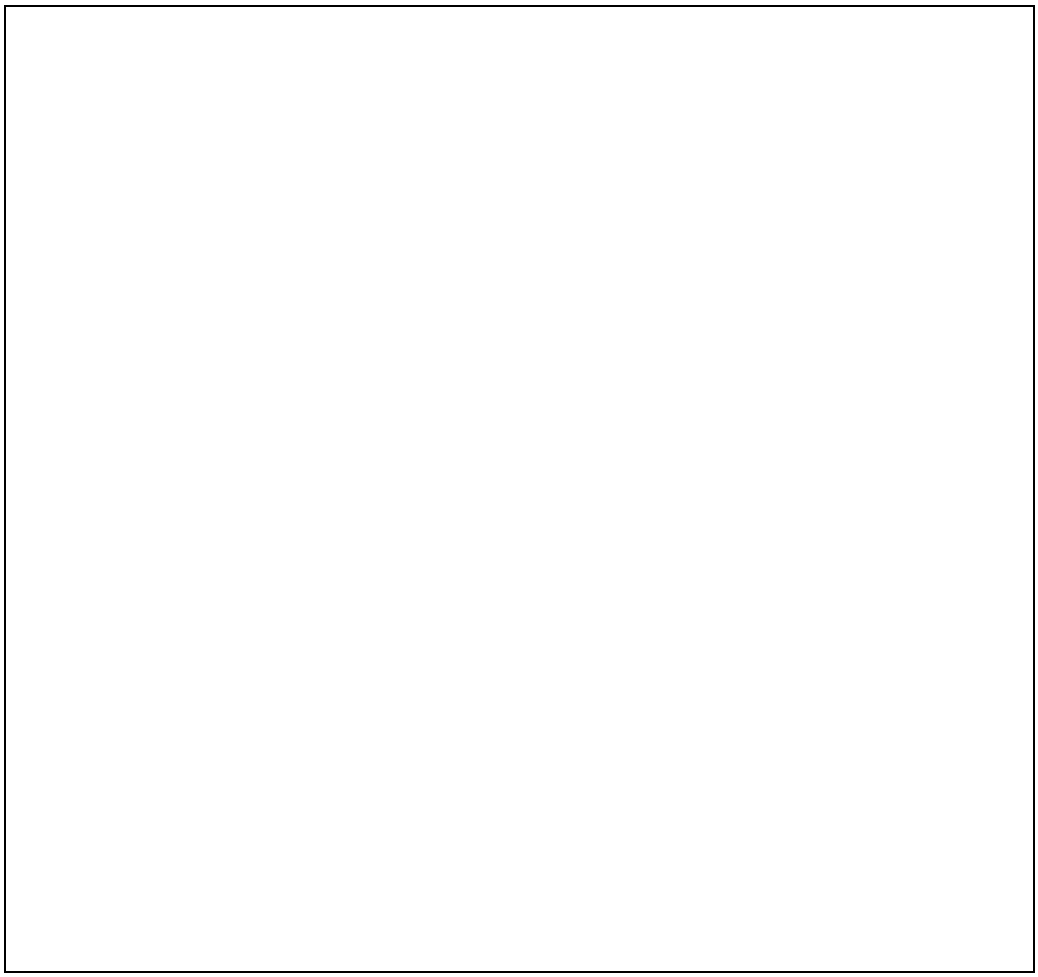


Figure 4.2 Plot of mean size values for genomes within the fitness classes MB (blue) and FB (red), across the three LDs. Dashed lines represent TG expression in males, and solid lines TG expression in females.

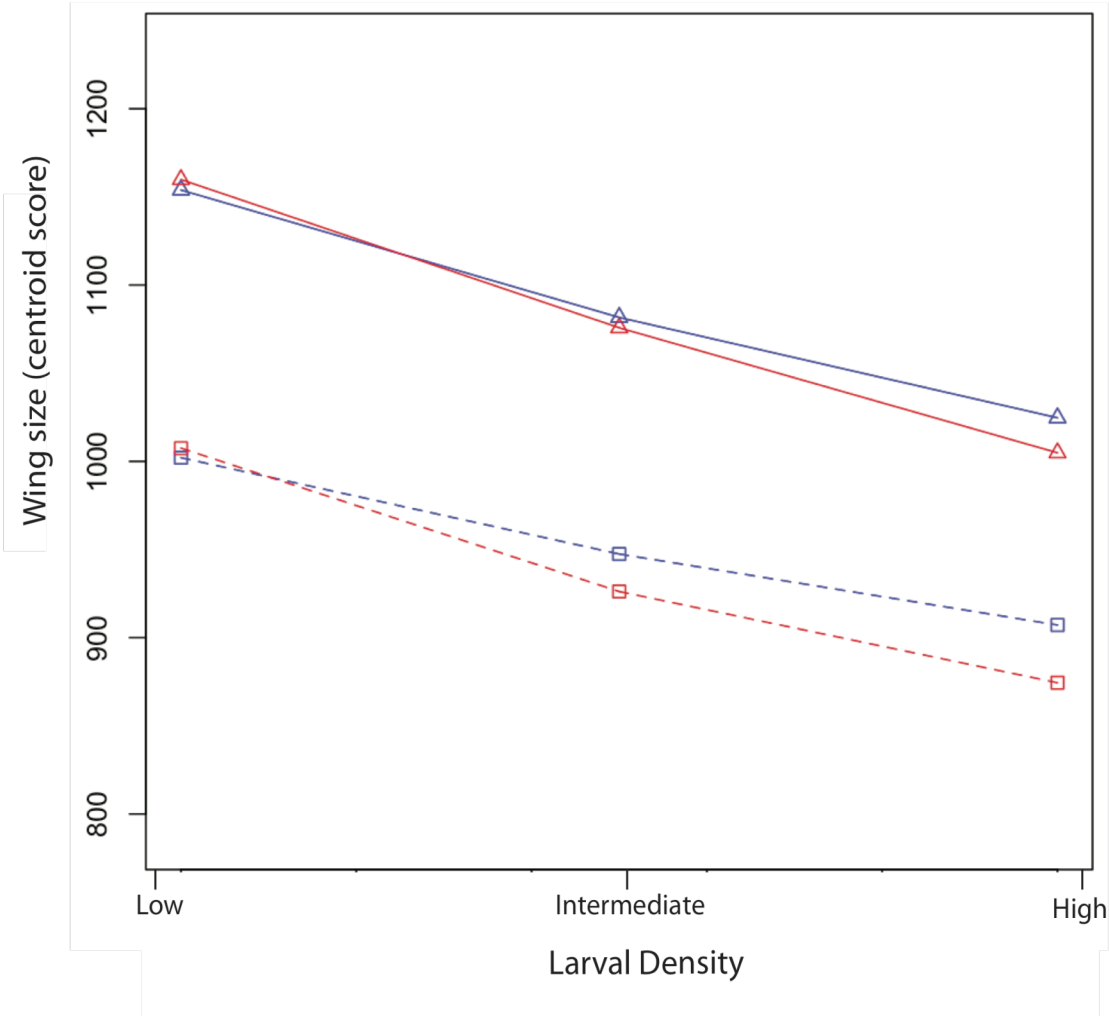


Figure 4.3 Plot showing the relationship between the size differential (L – H) of each genome when expressed in male (X axis) and female (Y axis) backgrounds. Colours indicate the fitness class of each genome (MB = blue, FB = red, and N = black).

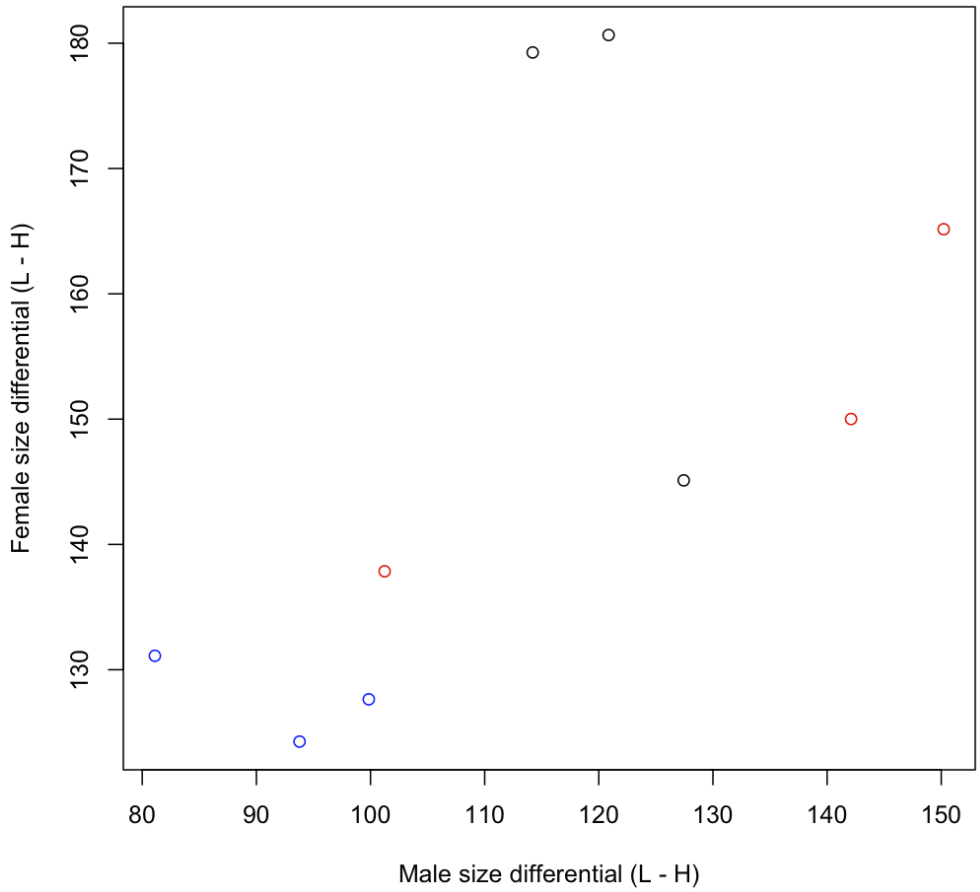


Figure 4.4 Plot showing the mean direction of LM movement from high to low LD for both males (green), and females (black). Arrow tails indicate the relative scale of movement for each LM, and are amplified to 4x the actual distance for better visualization. The X and Y axes represent the range of coordinate values after performing Generalized Procrustes Superimposition (see Methods, section 4.2.7 for detail).

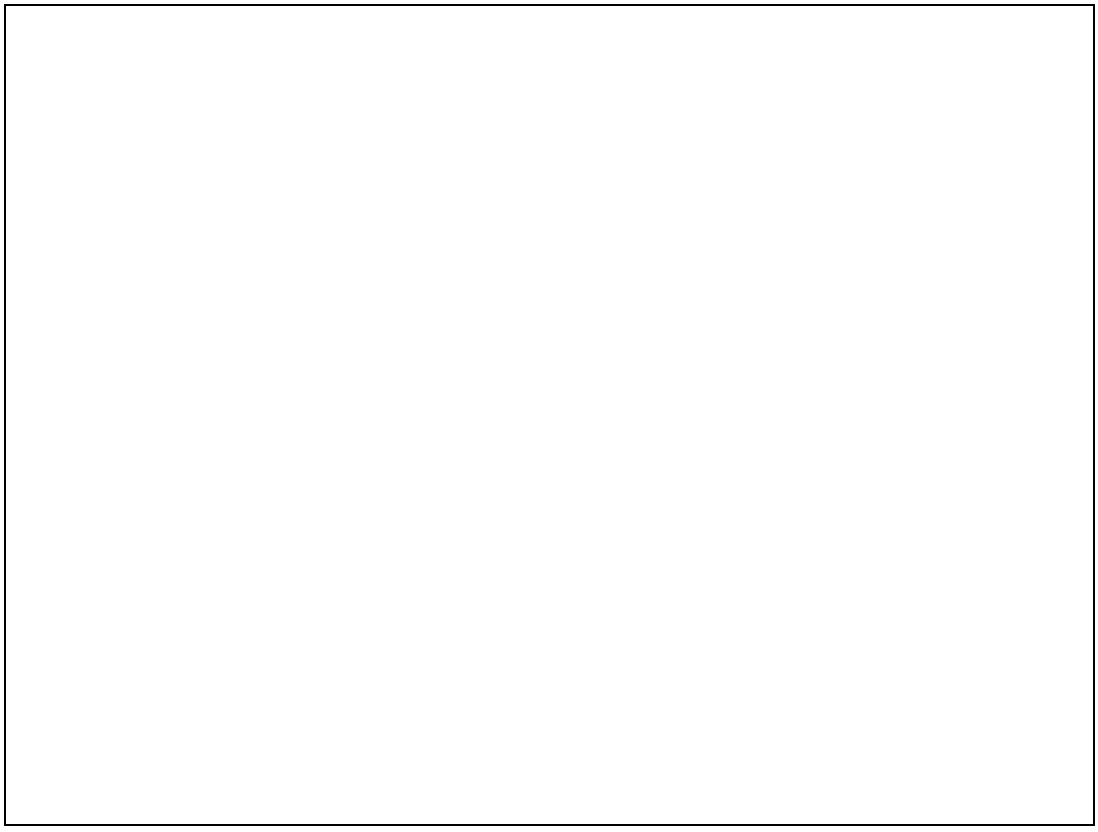


Figure 4.5 Extant sexual dimorphism. Plot showing the difference between the mean LM locations of males and females at low LD. Arrows point from the mean female coordinate to the mean male coordinate. The X and Y axes represent the range of coordinate values after performing Generalized Procrustes Superimposition (see Methods, section 4.2.7 for detail).

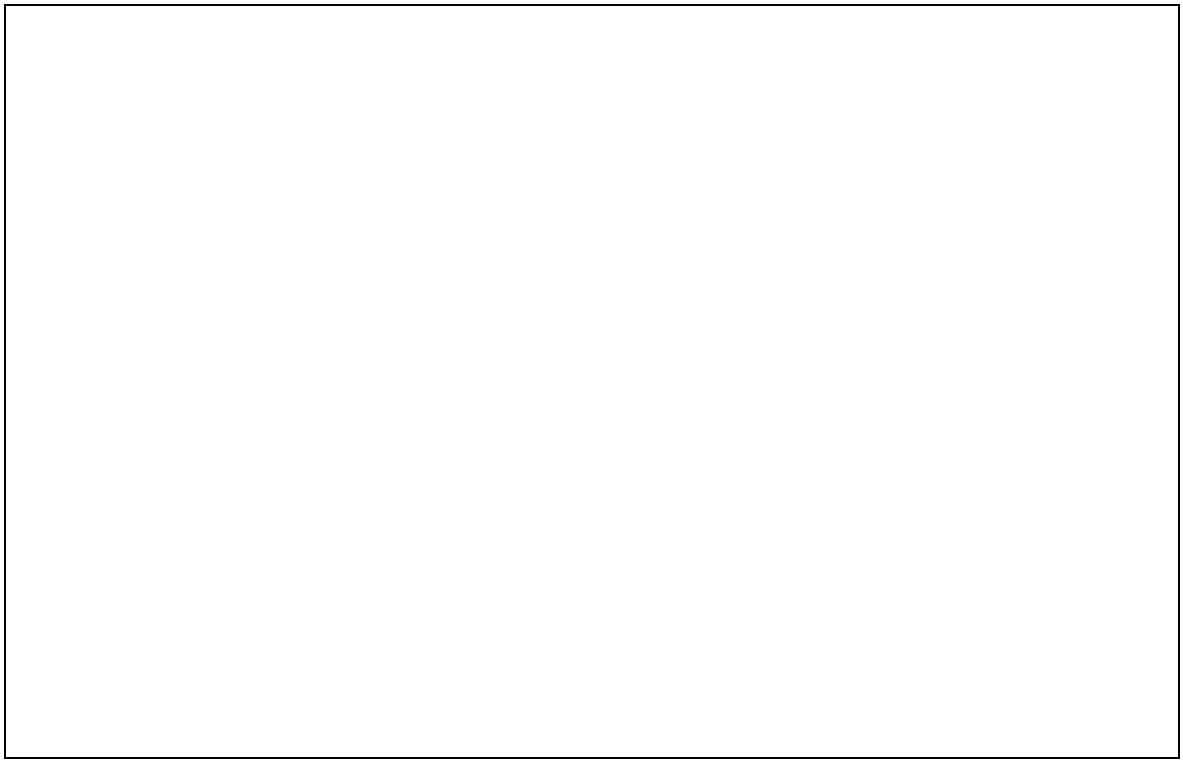
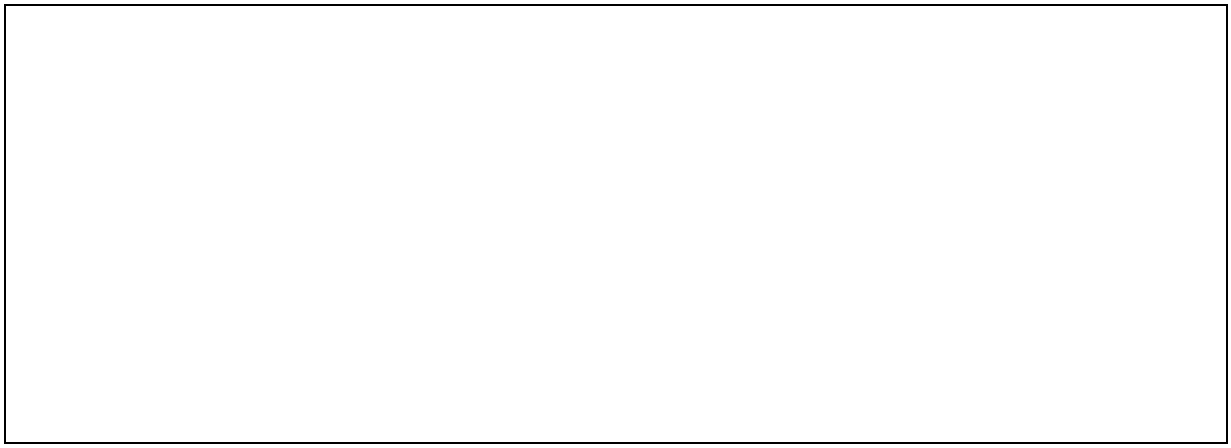


Figure 4.6 Rank of female fecundity and wing size at low and high LDs. Diagram representing the rank of mean female fecundity and mean female wing size values at low (L) and high (H) LD. Line colors represent the fitness-class of each genome (FB = red, MB = blue, N = black).



5

General discussion

5.1 Overview

In this general discussion I start by providing a summary of how the findings of my three empirical chapters relate to each other. I will describe briefly how my results combine to enhance our understanding of variation in sex-specific fitness, at genome-wide and trait-specific levels. Then I go on to describe some valuable future studies that constitute interesting extensions of my findings.

5.2 Summary of principal findings

In this thesis I explored the topics of genome-wide SA fitness variation, and components of sex-specific fitness in *D. melanogaster*. I investigated the effects of genetic drift on SA fitness variation in chapter 2, and the effects of wing morphology on male fitness in chapter 3. Then I complemented my data on males with a further suite of data for measures of fitness and wing morphology in females (chapter 4). This allowed me to evaluate the contribution of wing morphology to genome-wide SA fitness variation. Finally I investigated the associations of development time with fitness and wing morphology in male and female *D. melanogaster* (chapter 4).

I obtained experimental data on the effects of genetic drift on SA fitness variation by assaying 4 small, independent, populations of the Dahomey population of *D. melanogaster*. My experimental design mimicked that of Chippindale et al. (2001) and estimated sex-specific fitness in both larval and adult life stages. My primary finding was that the small independently evolving populations diverged significantly in their sex-specific adult fitness. Importantly, the divergence in fitness was not consistent with the random fixation of deleterious mutations, as would normally be expected under genetic drift in small populations. Instead the divergence in sex-specific fitness occurred along a SA fitness continuum, where an increase in the fitness of one sex was associated with a corresponding decrease in the fitness of the other sex, within each of the separate populations.

I adopted a trait-focused approach to the study of male fitness by exploiting a sample of 9 genomes that belonged to three classes of fitness patterns: low male/high female fitness, high male/low female fitness, and an intermediate level of fitness in both sexes. I measured the wing morphology (size and shape), and a component of male fitness (mating success) of males expressing genomes from each fitness class across three different larval density environments. I found that wing morphology substantially

affected the fitness of males. Using male wing size as a proxy for overall body size I showed that larger males gained a fitness advantage from being bigger than their immediate competitors. However, I did not detect any further increase in male fitness beyond a relative size advantage. In other words male size only inhibited reproductive success when a male was relatively smaller than their competitor. Wing shape did vary allometrically with wing size, but also wing shape variation that was independent of size affected fitness. Wing shape explained more male fitness variation than did wing size when a male was of equal or greater size than their competitor. I also provided some tentative evidence that the mean direction of selection of male wing shape, in part, opposed the mean direction of wing shape change with increased size, suggesting that males were unable to achieve their optimal wing shape when wings were large. Finally, I demonstrated that male fitness across my sample of genomes was correlated with the phenotypic robustness of size, specifically suggesting that higher male fitness may be associated with increased phenotypic robustness. I discussed these findings with respect to selection on male wing morphology and the potential for SA fitness effects associated with the wing morphologies of both sexes.

I obtained a homologous data set for female fitness and wing morphology by expressing the same genomes from each fitness class that were previously used to analyse variation between males in a female background. As for males, I measured female fitness, wing size, and wing shape in three larval density environments. In addition, I performed experiments to measure development time in both sexes. The new suite of measures enabled the investigation of the inter-relationships of fitness, wing morphology, and development time between the sexes. I tested for evidence of SA selection on wing size and shape and asked whether males may have been subject to selection for decreased development rate owing to selection for a more accurate morphology. I found evidence that appropriate genetic architecture exists for the sexes to be genetically restricted in their evolution of sex-specific wing morphology. Although my data suggested that wing morphology was subject to sexually divergent selection, I did not find strong direct evidence that this was the case. Investigating environmental robustness for size across the dataset with both sexes revealed the interesting results that male-beneficial genomes conferred high robustness not only to males expressing them but also to females. Thus, females expressing these genomes showed increased developmental buffering against intense larval competition, despite the fact that they had low fitness. Finally I found, counter intuitively, that increased male fitness was associated with a

higher developmental rate. I discussed these findings with respect to selection on the wing morphology of males and females.

5.3 Sexual antagonism in laboratory-adapted D. melanogaster

An interesting finding in Chapter 2 was that I detected sexually antagonistic fitness variation in the Dahomey laboratory population of *D. melanogaster*, for which SA had never previously been reported. To my knowledge SA fitness variation has been detected in two other laboratory populations of *D. melanogaster*, namely the IV population (Connallon and Jakubowski 2009) and the LHm population, which I used here in the experiments reported in chapters 3 and 4. The LHm population in particular has been used in numerous other experiments that have investigated SA (e.g. Chippindale et al. 2001; Prasad et al. 2007; Abbott et al. 2010; Innocenti and Morrow 2010). Some authors (e.g. Chapman et al. 2003) have argued that the relative ease of detecting SA fitness variation in laboratory-adapted populations suggests that it could be an artifact of unusually consistent rearing conditions. The notion is that under constant laboratory conditions then any adult fitness variation that is attributable to unconditionally deleterious mutations is likely to get purged from a given population. Consequently, the relative contribution of alleles with SA fitness effects could become inflated. This argument is potentially applicable to the LHm population because it has been reared for many generations under a strict two-week cycle with non-overlapping generations (see Methods, section 3.2.1). However, in chapter 2 I found evidence of the presence of SA fitness variation in the ancestral Dahomey population from which the four selection lines that I tested had been derived. The Dahomey population is reared in large population cages of several thousand individuals and with overlapping generations. Therefore my study and Connallon and Jakubowski's (2009) experiments using the IV population demonstrate that the detection of SA fitness variation is not an oddity of the LHm population, nor is it an inevitable consequence of a history of highly controlled rearing conditions.

A striking difference between the studies detecting SA in the LHm population (e.g., Chippindale et al. 2001), and my findings in chapter 2, is that I found evidence of some adaptive conflict over optimal larval and optimal adult phenotypes. My experiment showed that, averaged across the sexes, higher mean population fitness at the larval stage was associated with lower mean population fitness at the adult stage. One possible explanation for such an effect in a sample derived from the Dahomey population is that

there is substantial variation in life history traits in that population. Given that larval density in the Dahomey population is uncontrolled, there is considerable scope for alternative life history strategies. Consequently the potential for the persistence of genetic variation in larval fitness is much greater for the Dahomey population than for the LHm population. In general terms, this highlights the fact that larval fitness may have important effects on sex-specific adult fitness..

The trade-off between larval and adult fitness that I detected in chapter 2 was also mirrored in my analyses on male development and fitness in chapters 3 and 4. There, I found tentative evidence that male adult fitness correlated positively with larval developmental rate. This finding suggests a trade-off between larval and adult fitness when combined with evidence from other studies demonstrating that a higher development rate is associated with reduced pre-adult survival (Chippindale et al. 1994; Chippindale et al. 1997). Collectively, these results suggest that a higher rate of development could increase adult fitness in males, whilst reducing larval survival, and hence by our measure also reduce larval fitness. However, more experimental data will be necessary to verify these hypotheses and to establish the generality of the patterns I observed.

5.4 Genetic drift and sexual antagonism

In chapter 2, we detected genetic drift in sexually antagonistic variation by directly assaying the sex-specific fitness of genomes sampled from the four populations. Measuring fitness is often difficult in wild populations and it would therefore be interesting to explore how genetic drift would affect the expression of phenotypic traits associated with antagonistic fitness effects by causing frequency changes in the alleles underlying them. The question is, how would we expect small, isolated populations to differ in the expression of antagonistic phenotypic traits, such as wing morphology? In chapter 2 I found evidence that genetic drift causes a population divergence in sex-specific fitness variation. This means that under genetic drift populations which evolved high male fitness did so at the expense of the level of female fitness, and vice versa. One prediction I can make from this result is that the morphology of these independently evolving populations should also diverge in a sex-specific manner. For example, a population that has evolved to have high male/low female fitness should evolve a more masculinised phenotype across both sexes whereas population with a high male/low female fitness should show more feminised traits values in both sexes.

With respect to wing morphology, I would predict that the variation in wing morphology between populations should reflect in the direction corresponding to the sex with high fitness in the given population. Furthermore, I would expect more phenotypic variation between, than within populations. It may in theory be possible to detect increased between-population variation in candidate antagonistic traits, by comparing them to traits that are known to be under stabilising selection. However, it is unclear whether this comparison would make it possible to use such a test to support the association of these traits with antagonism, as neutral evolution would lead to similar patterns of phenotypic divergence between populations.

An interesting theoretical study that relates directly to findings in chapter 2 was conducted by another member of our laboratory. This study by Mullon et al. (2012), investigated how the interplay between selection and genetic drift should affect the genomic distribution of SA alleles. Their model specifically predicted that genetic drift should lead to the accumulation of SA alleles on the X chromosome in male heterogametic (XY) species, and on the autosomes in female heterogametic (ZW) systems. Furthermore, this effect should be especially pronounced when sexual selection is strong among males.

5.5 Wing morphology

The design of my study of wing morphology prevented me from using standard quantitative genetic analysis. I worked with a sample restricted to 9 genomes and in addition these genotypes were not a random sample of the available fitness variants in the population but were specifically chosen for their unusual phenotype. In the future it will be important to characterize wing traits associated with the full range of fitness values, to measure genetic variances of and genetic covariances between those traits with the aim of being able to predict evolutionary change, and its constraints, on the different wing morphological traits. One such study has recently been conducted in my research group (Reuter laboratory). It has used a random sample of 30 genomes to encompass a broad snapshot of fitness variation. Genomes were expressed in both sexes and flies reared in a controlled low density larval environment to measure wing morphology of each sex. Future analysis of the results of this study will be valuable in the context of the association between variation in wing morphology and fitness.

Finally, I propose a future study to further our knowledge of selective pressures on wing phenotypes. In agreement with functional arguments (Gidaszewski et al. 2009; Abbott et al. 2010), one of my key findings was that the high proportion of male fitness variation explained by wing shape suggests that there may be direct benefits for males if they have wings of particular morphologies. Wings are used to generate the male courtship song so that the morphology of this 'instrument' could be important for male success in intersexual competition. To date there is no direct evidence for a mechanism by which wing shape could alter male fitness. However, research does suggest that wing shape comprises an important component of intersexual competition among males. Specifically, Snook et al. (2005) have demonstrated that courtship song may rapidly evolve under conditions of strong sexual selection in *D. pseudoobscura*. I hypothesize that variation in wing shape may generate corresponding variation in the pitch of the generated song and this could be subject to sexual selection among males. To my knowledge, the variation in the pitch of *Drosophila* courtship song with variation in wing shape has not been assayed. It would be interesting to investigate how wing shape affects the characteristics of courtship song, especially the pitch of the generated sound. If courtship song did indeed vary with aspects of wing shape, then one could subsequently isolate the effect of courtship song using auditory playback of previously recorded courtship song. This would permit the identification of those wing shapes that generate courtship song favoured by females. One way to do this would be to run multiple mating assays, each with one randomly selected female and a randomly selected male with clipped wings, so that the experimental male cannot perform courtship song. By playing an array different variants of courtship song across multiple mating assays, one could use the latency to copulation as a direct measure of the most favored variations of courtship song.

5.6 References

Abbott JK, Bedhomme S, Chippindale AK (2010) Sexual conflict in wing size and shape in *Drosophila melanogaster*. J Evolution Biol 23(9): 1989-1997.

Chapman T, Arnqvist G, Bangham J, Rowe L (2003) Sexual conflict. Trends Ecol Evol 18(1): 41-47.

Chippindale AK, Gibson JR, Rice WR (2001) Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. P Natl Acad Sci USA 98(4): 1671-1675.

Chippindale AK, Hoang DT, Service PM, Rose MR (1994) The evolution of development in *Drosophila melanogaster* selected for postponed senescence. Evolution 48(6): 1880-1899.

Chippindale AK, Alipaz JA, Chen HW, Rose MR (1997) Experimental evolution of accelerated development in *Drosophila* .1. Developmental speed and larval survival. Evolution 51(5): 1536-1551.

Connallon T, Jakubowski E (2009) Association between sex ratio distortion and sexually antagonistic fitness consequences of female choice. Evolution 63(8): 2179-2183.

Gidaszewski NA, Baylac M, Klingenberg CP (2009) Evolution of sexual dimorphism of wing shape in the *Drosophila melanogaster* subgroup. Bmc Evol Biol 9.

Innocenti P, Morrow EH (2010) The sexually antagonistic genes of *Drosophila melanogaster*. Plos Biol 8(3).

Mullon C, Pomiankowski A, Reuter M (2012) Molecular evolution of *Drosophila* Sex-lethal and related sex determining genes. Bmc Evol Biol 12.

Prasad NG, Bedhomme S, Day T, Chippindale AK (2007) An evolutionary cost of separate genders revealed by male-limited evolution. Am Nat 169(1): 29-37.

Snook RR, Robertson A, Crudgington HS, Ritchie MG (2005) Experimental manipulation of sexual selection and the evolution of courtship song in *Drosophila pseudoobscura*. Behav Genet 35(3): 245-255.

Appendix 1

Pilot studies on the effects of LD variation on aspects of *D. melanogaster* development time and wing size

The material described in this appendix comprises evaluations of the effects of the experimental manipulation of LD on the development time and wing size of *D. melanogaster*. Our general approach was to measure the variation in these traits that results from systematic experimental manipulations of larval density (LD). In this way we were able to gauge the range of phenotypic variation created by specific conditions and so rigorously plan future experiments. In addition, we estimated the amount of larval mortality associated with *D. melanogaster* matings that generate a proportion of progeny with a triple X chromosome haplotype. These pilot studies were an essential pre-requisite for determining the details of the experimental design of the larger scale studies, which are fully documented in Chapters 3 and 4. Throughout the pilot studies we carried out trials using the LHm base population, or using cloned haploid target genomes (TGs) derived from the LHm population.

A1.1 Development time and larval density (LD)

For *D. melanogaster* the term development time describes the period from egg laying through to eclosion of an adult fly, during which period larval growth and pupation occur. Typically this process takes roughly 8.5 days under standard laboratory culture, with optimal food at 25^{o}C (Ashburner and Thompson 1978). Development time shows considerable variation between the sexes (Bonnier 1926), with the eclosion of females usually beginning before that of males from the same brood. When LD is increased (i.e. less food is available to each individual larva), there are corresponding increases in both development time, and the length of the eclosion period (the time from the first to the last flies to emerge) (Peters and Barbosa 1977).

Our goal was to carry out a large-scale experiment that included the measurement of development time across a range of different LDs. It was therefore important to first conduct a pilot study to directly measure the relationship between development time and LD with the stock populations, so as to establish the best empirical parameters for the later experiment.

Several aspects of the measurement of developmental time required consideration. First, we must determine exactly how development time varies with changes in LD. Previous research suggests there is a linear response in development time to increases in LD (Santos et al. 1994). Despite this background knowledge, it could not be assumed that the stock population (LHm) used in our study would show an identical response to manipulations of LD. Of particular significance, the rearing regime of the LHm

population involves strict control of LD (175 eggs per vial, with 10ml of culture media). As the LHm population has been adapting to these rearing conditions for more than 450 generations, we might expect the response to changes in LD to be relatively different compared with other stock populations that experience more variable LD, and hence requires a priori testing. Second, in measuring development time it is important to use a sampling interval that provides optimal resolution, and that is logistically possible to implement based for the scale of a given experiment. Previous assays of development time have used relatively long intervals (up to 12 hours), between the sampling of flies during eclosion. Given that eclosion will last a maximum of 48hrs when larval competition is very low, an interval of this length provides a limited resolution of only 4 intervals from which to attain data. It was therefore important to establish the length of interval necessary to make accurate and efficient measurements of development time, across a range of different LDs. Finally, we must consider the total duration of the period of eclosion, and how it varies (or not) with increasing LD. Specifically, the total duration of eclosion at each LD determines the logistical feasibility of testing development time at a particular LD. In addition to these factors involved in the measurement of development time, we also wanted to estimate when the peak periods of eclosion occur for flies reared across a range of LDs. This latter point relates to the acquisition of flies for morphological analysis from different LDs. In particular, it is important to standardize the period over which flies are extracted for morphological analysis to minimize morphological variation caused by micro-environmental differences that can change across the eclosion period. A second pilot assay (see below *A1.2*) covered some aspects of this in detail. Here, by closely inspecting the eclosion rates during development at different LDs, we could make estimations, as to when it was optimal to extract adult flies for morphological analysis.

To test these factors, we set up a series of fly cultures at different LDs sired by groups of flies randomly selected from the LHm population. Each LD was achieved by transferring newly hatched first instar larvae from eggs that had been laid on grape juice media over 12 hours (see Methods. Section 3.2, of Chapter 3), into vials containing exactly 10ml of cornmeal-molasses-yeast media. Three LDs were set up; 50 (low LD), 180 (intermediate LD), and 350 (high LD), with four replicate vials for each LD. As soon as the wave of eclosion commenced, all vials were cleared at 12 hour intervals, once in the morning (10am), and once in the evening (10pm). Each interval, is herein referred to as a time-frame, and is denoted T1 for the first time-frame, T2 for the second

time-frame, and so on. The adults collected at each time-frame were frozen and subsequently counted. From this data we calculated the proportion of emerging flies during each time-frame.

To check for differences in development time between the four replicates within each LD we fitted three, one-way ANOVAs, for the data at each LD. Each ANOVA modeled the number of flies eclosing at each time-frame as the dependent variable, with the term, replicate, as the independent variable (Table A1.1). Within all LDs this analysis showed non-significant variation between replicates across all time-frames. These results confirmed that flies reared in independent vials at equivalent LDs displayed similar rates of development. To investigate how the profile of adult emergence varied between LDs we plotted bar charts of the mean proportion of flies that emerged during each time-frame, one for each LD tested (Figure A1.1). These plots showed clearly that within the low and intermediate LDs more than 50% of the flies emerged during T2, and hence this represented the peak eclosion period for both of these LDs. By the end of T3 (36 hours after eclosion begins) 100% of the low LD and 90% of intermediate LD flies had eclosed. The emergence of flies from high LD was more protracted, with the largest proportion, 34% eclosing during T3 (peak eclosion at high LD), and more than 50% eclosing between T4 and T6.

Firstly, our results demonstrated that the manipulation of LD to particular levels generated repeatable effects on development time. Replicates within each LD showed a very high level of synchrony, meaning that they serve as a reliable source of replication in scaling-up experimentation on development time at fixed LDs. However, the observed asynchrony between the emergence of flies from high LD, relative to the low and intermediate LDs makes tests at the high LD logistically difficult to include in a larger experiment. This was compounded by the fact that these high LDs stretch the total eclosion period up to ~3.5 days, making the eclosion period so long that it was not directly relevant to the rearing regime conditions of the LHm population (see Methods, section 3.2.1, Chapter 3), in which flies that eclose more than 48 hours after the beginning of eclosion cannot make it into the next generation. Given these findings, it was decided that large scale experiments focused on developmental time data would not be conducted at the high LD. Instead the intermediate and low LDs were used. Specifically, the intermediate LD provides larval conditions that are comparable to those of the LHm rearing regime, and therefore can be used to generate development

time data that is relevant to the conditions that the flies have adapted to for more than 450 generations. The low LD constituted, what we predict were optimal conditions for development, with very low larval mortality, and was used to provide a baseline measure of optimal development time under conditions of very little or no larval competition.

For low and intermediate LDs, the interval of 12 hours between collections of newly eclosed flies, provides just three time-frames (T1, T2, and T3) from which more than 90% of the flies would eclose. This resolution is likely to be insufficient for measuring accurate differences between the genomic effects on development time, as we would like to. A logistically viable solution was to divide each 12 hour interval into three, to make 4 hour intervals. This should provide a minimum of 7-9 time intervals during which flies are collected to measure development time and so enable a more accurate comparison of genomic effects on development time.

A1.2 Size and larval density

In holometabolous insects, such as *Drosophila,* growth is restricted to the larval stages. Larval nutrition is a key factor in determining adult size (Nijhout 2003; Edgar 2006; Mirth and Riddiford 2007). Larvae reared at higher LD have reduced access to food due to increased competition, and on average will develop into smaller adults (Miller and Thomas 1958). Size is an important component of fitness for both sexes in *D. melanogaster*, and is correlated with numerous other traits, such as fecundity (Partridge and Farquhar 1983), sexual attractiveness (Long et al. 2009), and development time (Robertson 1963; Partridge et al. 1999). One of the principal questions addressed in chapters 3 and 4 is how fitness is affected by environmental size manipulation, in the context of genotypes that confer extreme values of sex-specific fitness. Prior to conducting large scale experiments addressing that question, it was important to establish the size range of adult males and females that would be generated by a suite of LDs. In particular, we required that there were repeatable and significant size differences between each different LD, but also that each density range should overlap with the next closest LD. In this way we could measure the effect of LD on size, i.e. how phenotypic size varied across LDs for a given genome, and the overall effect of size on the fitness of particular genomes.

To measure how LD affected size, a series of different LDs were evaluated. We chose a range of LDs that we predicted would generate an appropriate range of variable, but overlapping size groups in both males and females. To obtain results that would be of general applicability to our future studies with nine particular TGs, we pooled these TGs spanning the range of sexually antagonistic fitness effects (see Methods, section 3.2.2). These TGs were expressed in male and female backgrounds, by crossing mixed TG-CG males with virgin DxLHm and LHm females, respectively (see Methods, section 3.2.3). Each LD was set up using the transfer of first instar larvae (see Methods, section 3.2.4), into vials containing 10ml of culture media. For evaluation of TG expression in a female background, three repeats of three different LDs were set up, low; 50, intermediate; 200, and high; 400. For evaluation of TG expression in a male background, we set up three repeats at three similar LDs; low; 50, intermediate; 160, and high; 350. Once the adult flies had fully eclosed, the LHm-TG flies were identified and frozen for subsequent wing mounting. For assessing male and female expression we mounted more than 10 wings per replicate at low LD, and more than 20 wings per replicate at each of the intermediate and high LDs. Wings were detached using tweezers and mounted onto glass slides with isopropanol, and sealed with a glued cover slip. Wing area was then calculated using image J (http://rsb.info.nih.gov/ij/), on a computer attached to a light microscope.

To investigate the size affect of each LD on TGs expressed in males and females we fitted a one-way ANOVA within each sex on the dependent variable, wing size, with the independent variable, LD (Table A1.2). For females there was a highly significant difference in wing size between LDs, and a *post hoc* Tukey test revealed that the wing sizes were significantly different between all LDs (Table A1.3). This result is depicted clearly in figure A1.2, which shows the consistent separation of the effects of LD on female size. For males a one-way ANOVA also revealed highly significant differences in wing size between LDs. However the *post hoc* Tukey test showed that there was no significant difference between the low and intermediate LDs (Table A1.3). These differences are shown in figure A1.2. The mean wing size for male flies reared at the low and intermediate LDs were 0.999 cm^2 and 0.986 cm^2 respectively, both of which were significantly different to the mean wing size value at high LD, 0.856 cm^2. Here we showed that increases in LD did cause corresponding decreases in wing size, but the level at which LD began to reduce the achievable size for males was unclear. In particular the intermediate LD (160 larvae per 10ml of media) produced male wing

sizes that were almost identical to those produced at low LD. However, it is likely that the lack of wing size variation between the low and intermediate LDs was in part the result of death during, or after, the 1st instar larval stage. Specifically, a subsequent assay revealed that ~ 27.8% of the larvae produced by crosses for the expression of TGs in males died during the first instar stage of development (see below, Appendix 1, section A1.3). This increased larval mortality is caused by a triple X chromosome karyotype, which is present in one third of the progeny produced by DxLHm virgin females. One of the consequences of this increased larval mortality was that the effective LD of the intermediate treatment (160 larvae per 10ml of media) is much lower than intended.

As with the effect of development time above, it is important that there is not substantial variation between the size effects of replicates that are set at the same LD. This way we can ensure LD treatments serve as a repeatable way of consistently manipulating adult size. To test for differences between the size effect of replicates within each LD, we fitted three, one-way ANOVA's for each sex, one for each LD. We modelled the dependent variable, wing size, with the independent variable, replicate. For nearly every LD for both sexes we found no difference in mean wing size between replicates, with the exception of the female low LD (Table A1.4). This significant variation within the low LD for female size could be the result of genome effects. In the case of this particular larval treatment these effects may be evident here for two reasons. First, the relative proportion of each genome at this low LD could be highly variable between replicates, leading to over or under representation of some TGs. As a result of this sampling variation we could actually observe significant differences between replicates due disproportionate genomic effects. Second, each genome may be able to reach optimal size at this low LD (Sang 1949) because they have access to an excess of nutrition, this in turn may lead to greater differences between genomes as they can achieve their maximum phenotypic size. Furthermore, this would exaggerate the effect of disproportionate representation of TG's between replicates. Overall, the consistency of wing size variation between the replicates for the majority of LDs showed that manipulation of LD served as a reliable way to manipulate adult size for TGs that were expressed in either sex.

Our results provided information that is essential for choosing a range of LDs that are appropriate for larger experiments that manipulate adult body size of particular TGs

(see Chapters 3, and 4). This pilot assay showed reassuringly, that we could generate repeatable changes in the size of TGs with changes in LD. As a result, size data acquired from flies reared within different replicates did not need to be collected or analysed separately. However, our results also showed that where there were differences between the size effects of LDs, there was very little overlap. As a result future assays should use smaller differences between LDs. Perhaps most importantly we showed that the number of larvae picked for TG expression in males must be scaled up by a factor of ~ 30% to account for disproportionately high levels of larval mortality. Taking into account these factors, we decided on the following LDs to create a range of three equivalent larval conditions each sex; for females, low = 50, intermediate = 175, high = 300, and for males, low = 60, intermediate = 240, and high = 400.

A1.3 DxLHm genotype: estimating larval mortality

Where we manipulated LD for the experiments implemented in Chapters 3 and 4, it was of particular importance that we could impose equivalent and consistent rearing conditions between and within each sex. However this could not be achieved by a uniform standardization of LDs at the start of the first larval instar. A specific problem was caused by the expression TG's in a male LHm background. To generate these TG-LHm males required crossing TG-CG males with DxLHm virgin females. The compound X chromosome (Dx) ensured paternal transmission of the TG X chromosome from father to son (see Methods, section 3.2, Chapter 3). However one third of the offspring from this cross inherit the compound X, with an additional wild type X chromosome. Individuals with this triple X karyotype did not develop into adults, but the majority of them are viable until the first instar stage. Given that our strategy for assaying male and female fitness across environments involved standardising LD we needed to ensure that LDs genuinely correspond to the intended values.

First, we confirmed the exact stage of this larval mortality by carrying out a small observational study. This mini assay was performed by transferring 20 sets of 10 larvae, sired by DxLHm mothers, every 24 hours for 5 consecutive days. Each set of 10 larvae were maintained on grape juice media, with a small volume of live yeast paste. Direct observations revealed that all larvae that died, did so towards the end of the first larval instar stage. As we picked larvae to control LDs during the first instar stage, it is essential that we account for any mortality that occurred after this stage, so that we can

adjust our methods accordingly to create equivalent rearing conditions. In other words, initial LDs for the expression of TGs in males needed to be increased to take account of the early death of one of the progeny genotypes.

We estimated by how much we needed to increase the initial LD of the male TG expression crosses to generate final LDs that were equivalent to those generated by female TG expression crosses. To calculate this parameter we used data from a previous experiment where we controlled LD. In this previous study we crossed each of our 9 TG's with DxLHm virgin females. LDs were controlled by transferring groups of 50 first instar larvae into vials with 10 ml of standard culture media. For each genome this was replicated 10 times. Once the adult flies eclosed, they were counted and the mean proportion of survivors to adulthood was calculated for each TG (Table A1.5). These 9 values were used to calculate the mean deviation across all vials, from the expected value of 50 adult flies per vial. On average there were 27.8% fewer adult flies than would be expected if there was no mortality. We subsequently used this value as a benchmark for the proportion of larval death for experimental crosses that express any of our 9 TG's in a male LHm background.

A1.4 References

Ashburner and Thompson (1978) The laboratory culture of *Drosophila*. In: The genetics and biology of *Drosophila* Vol. 2a. Edited by M. Ashburner and T. R. F. Wright. Academic Press, London & New York. pp. 1-109.

Bonnier G (1926) Temperature and time of development of the two sexes in *Drosophila*. Br J Exp Biol 4(2): 186-195.

Edgar BA (2006) How flies get their size: genetics meets physiology. Nat Rev Genet 7(12): 907-916.

Long TAF, Pischedda A, Stewart AD, Rice WR (2009) A cost of sexual attractiveness to high-fitness females. Plos Biol 7(12): e1000254.

Miller RS, Thomas JL (1958) The effects of larval crowding and body size on the longevity of adult *Drosophila melanogaster*. Ecology 39(1): 118-125.

Mirth CK, Riddiford LM (2007) Size assessment and growth control: how adult size is determined in insects. Bioessays 29(4): 344-355.

Nijhout H (2003) The control of body size in insects. Dev Biol 261(1): 1-9.

Partridge L, Farquhar M (1983) Lifetime mating success of male fruitflies (*Drosophila melanogaster*) is related to their size. Anim Behav 31(Aug): 871-877.

Partridge L, Langelan R, Fowler K, Zwaan B, French V (1999) Correlated responses to selection on body size in *Drosophila melanogaster*. Genetics Research 74(1): 43-54.

Peters TM, Barbosa P (1977) Influence of population density on size, fecundity, and developmental rate of insects in culture. Annu Rev Entomol 22: 431-450.

Robertson FW (1963) Ecological genetics of growth in *Drosophila* .6. Genetic correlation between duration of larval period and body size in relation to larval diet. Genet Res 4(1): 74-&.

Sang JH (1949) The ecological determinants of population growth in a *Drosophila* culture .3. Larval and pupal survival. Physiol Zool 22(3): 183-202.

Santos M, Fowler K, Partridge L (1994) Gene-environment interaction for body size and larval density in *Drosophila melanogaster*: an investigation of effects on development time, thorax length and adult sex ratio. Heredity 72: 515-521.

Table A1.1. Differences in development time between replicates at each LD. Table shows the results of onc-way ANOVA models using the number of eclosing flies per time-frame as the dependent variable, with replicate as the independent variable. Three ANOVAs were used to analyse, one for each of the low (D-50), intermediate (D-180), and high (D-350) LDs.

D-50	*Df*	*Sum Sq*	*F-value*	*P-value*
Replicate	3	4.14	0.0117	0.998
Residuals	24	2839.71		
D-180				
Replicate	3	154	0.0538	0.983
Residuals	24	22911		
D-350				
Replicate	3	225	0.0793	0.971
Residuals	24	22698		

Table A1.2. Effects of LD manipulation on wing size for males and females. Results of a pair of one-way ANOVA tests, modeling this dependent variable, wing size, with the independent variable, LD. Male and female wing scores were analysed separately.

Male

	Df	*Sum Sq*	*F-value*	*P-value*
LD	2	0.71565	80.798	< 0.001
Residuals	162	0.71744		

Female

	Df	*Sum Sq*	*F-value*	*P-value*
LD	2	3.1022	384.22	< 0.001
Residuals	287	1.1586		

Table A1.3. Effects of LD manipulation on wing size. Results of a *post hoc* Tukey tests on the ANOVA models (described in Table A1.2) to establish location of significant differences in wing size.

Male

	Difference	*Upper limit*	*Lower limit*	*P-value*
D300-D160	-0.13036293	-0.15797886	-0.10274699	0
D50-D160	0.01281364	-0.02010218	0.04572946	0.628
D50-D300	0.14317656	0.110083	0.17627013	0

Female

	Difference	*Upper limit*	*Lower limit*	*P-value*
D400-D200	-0.1278258	-0.1469588	-0.1086927	0
D50-D200	0.1707933	0.1436417	0.1979448	0
D50-D400	0.298619	0.2722396	0.3249985	0

Table A1.4. Differences between the wing size effects of replicates within LDs for each sex. Results of series of one-way ANOVAs modeling the dependent variable, wing size, with the independent variable, replicate. Each ANOVA analyses variation within one sex, and within one LD.

Male

D-50	*Df*	*Sum Sq*	*F-value*	*P-value*
Wing size	2	0.007137	1.2675	0.295
Residuals	32	0.090096		

D-160	*Df*	*Sum Sq*	*F-value*	*P-value*
Wing size	2	0.010489	1.8508	0.1655
Residuals	63	0.178517		

D-350	*Df*	*Sum Sq*	*F-value*	*P-value*
Wing size	2	0.0281	2.1259	0.128
Residuals	61	0.4031		

Female

D-50	*Df*	*Sum Sq*	*F-value*	*P-value*
Wing size	2	0.035509	6.8449	0.003
Residuals	39	0.101158		

D-200	*Df*	*Sum Sq*	*F-value*	*P-value*
Wing size	2	0.007854	1.4642	0.235
Residuals	107	0.286958		

D-400	*Df*	*Sum Sq*	*F-value*	*P-value*
Wing size	2	0.00337	0.3144	0.731
Residuals	135	0.72377		

Table A1.5. Larval mortality in crosses used to generate flies that express TGs in males. Table shows the mean number of adult flies emerging from crosses of nine specific TG-LHm males to DxLHm virgin females. At the first larval instar, progeny of each cross are established in controlled density cultures of exactly 50 larvae per 10ml of culture media. From this data we calculated the mean proportion of larval mortality, based on the expected total number of emerging adult flies (50).

Genome ID	*Mean No. of adults*
P50	39
H14	38
P48	41
P18	40
H13	37
H12	34
P7	23
P22	36
H7	37
mean mortality rate	27.78%

Figure A1.1. Profile of adult emergence across the eclosion period of each LD tested. Bar charts showing the mean proportion of flies emerging, during each consecutive time frame, for low (50), intermediate (180), and high (350) LDs. Each time frame corresponds to a 12 hour interval.

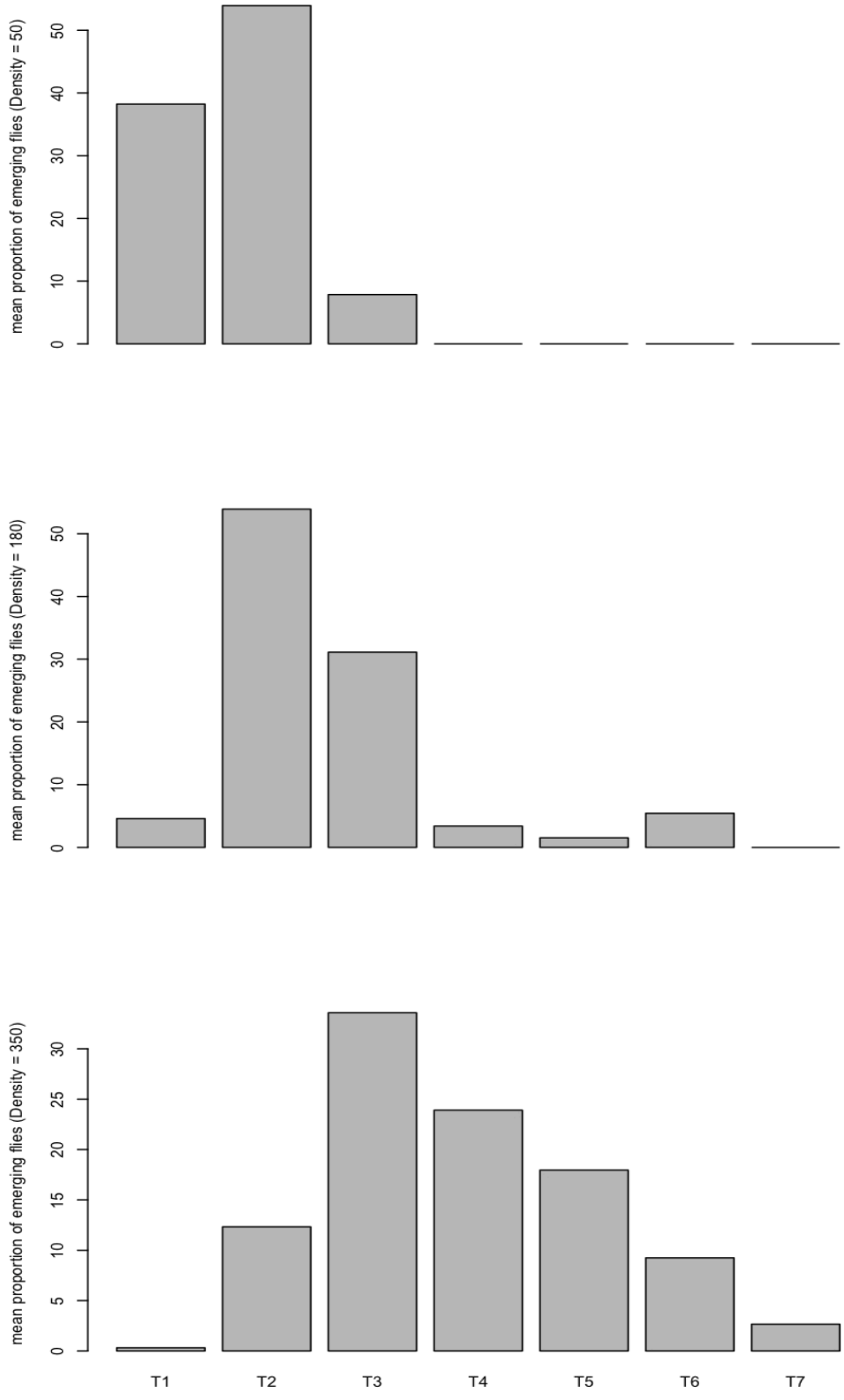
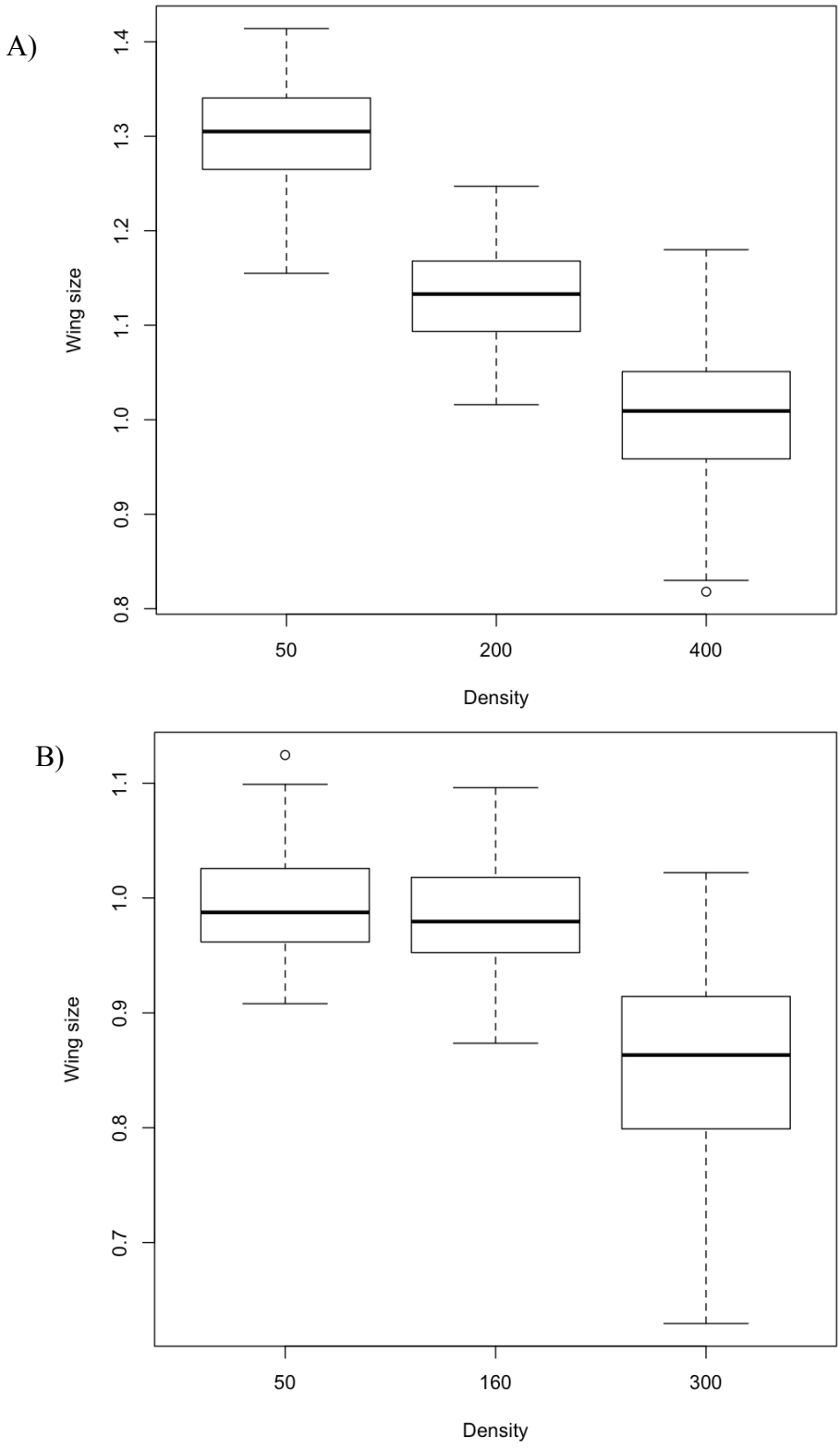


Figure A1.2. Wing size across LDs for females and males. Box plots showing the wing size and distribution of size, for wings from females (A), and males (B), across three LDs. Mean wing size (black horizontal line), 2nd & 3rd quartiles (boxed area), and total range of wing size (T – bars), are displayed for each LD.



Appendix 2

Resolution of sexual antagonism in a laboratory population of *Drosophila melanogaster*

This appendix contains the first draft of a manuscript written by Julie Collet, a former postdoctoral researcher in the Reuter laboratory. The text describes experiments to which I made a substantial contribution during my PhD. This work is intended for submission to *Evolution* with authors J. Collet, S. Fuentes, J. Hesketh, K. Fowler & M. Reuter.

A2.1 Introduction

Due to their different reproductive roles, male and female adults are often selected for different optimal phenotypes. However, the response to this divergent selection is complicated by the fact that both sexes typically share a large part of their genomes and new mutation frequently affect the phenotype of males and females in a similar way. The resulting genetic correlation between male and female phenotypes, in combination with divergent selection on the sexes sets the scene for intra-locus sexual conflict or sexual antagonism, where mutations that increase the fitness in one sex do so at the expense of the fitness in the other sex (Rice 1984; Bonduriansky and Chenoweth 2009; van Doorn 2009). Sexually antagonistic genetic variation has been shown to segregate in natural and laboratory populations of a wide range of organisms, including insects (Rice and Chippindale 2001; Gay et al. 2011; Berg and Maklakov 2012), vertebrates (Brommer et al. 2007; Foerster et al. 2007; Mokkonen et al. 2011) and plants (Kohorn 1994; Delph et al. 2011a). This growing body of evidence demonstrates that the common genetic basis of male and female phenotypes poses a constraint on adaptive evolution of sex-specific traits, even those that already show pronounced sexual dimorphism (Forsman 1995; Robinson et al. 2006; but see Bedhomme et al. 2011; Mills et al. 2012). By limiting male and female evolution towards their respective fitness optima, sexual antagonism is considered a powerful agent for the maintenance of genetic variation for fitness (Patten et al. 2010).

Despite the fact that sexual antagonism is both widespread and recognised as an important for in organismal evolution, we know relatively little about the long-term evolutionary fate of antagonism itself. It is generally assumed that the adaptive conflict between the sexes can be resolved by expressing antagonistic genes differentially in the two sexes. This would then allow the two sexes to diverge towards their respective phenotypic optima (Lande 1980; Rice 1984; Ellegren and Parsch 2007; Bonduriansky and Chenoweth 2009). On a proximate level, this scenario raises the question which processes allow resolution to happen. A number of possible mechanisms have been proposed. Models, both verbal and mathematical, have been used to predict that the resolution of antagonism and the evolution of differential gene expression could be aided by the duplication of antagonistic loci, thereby creating ‘permanent heterozygotes’ that carry both male- and female-beneficial alleles (Proulx and Phillips

2006; Connallon and Clark 2011; Gallach and Betrán 2011). From a conceptual point of view, it is a matter of debate whether this process would allow for antagonism to be completely resolved (Bonduriansky and Chenoweth 2009; Hosken 2011), but it would be expected to at least weaken adaptive conflicts between the sexes. Strong empirical support for the role of gene duplication in the resolution of antagonism is so far lacking. Genomic data suggest certainly suggests that duplicate genes are frequently expressed in a sex-specific manner, with expression being commonly testis-specific (Betran et al. 2002; Vinckenbosch et al. 2006; Wyman et al. 2012), however the link to antagonism is not clear. For particular classes of genes, such as duplicates of nuclearly encoded mitochondrial genes, pattern of testis-biased duplicate expression have been found and interpreted as in line with divergent sex-specific selection pressures (Gallach et al. 2010). However, the connection to sex-specific fitness has so far not been experimentally verified, and alternative mechanistic explanations for sex-biased expression of duplicates (Vinckenbosch et al. 2006; Fontanillas et al. 2007) have not been ruled out. In addition to gene duplication, epigenetic mechanisms such as imprinting could mediate antagonistic fitness effects and contribute to the resolution of conflict (Bonduriansky and Chenoweth 2009). Evidence for the role of imprinting comes from experimental results in the fly *Prochyliza xanthostoma*, where sexually dimorphic traits were transmitted to the offspring only through same sex parents (Bonduriansky and Rowe 2005).

Independently of which mechanisms allow for sexual antagonism to be resolved, the more fundamental question arises as to the timescale over which resolution takes place (van Doorn 2009). Sexual dimorphism is certainly ubiquitous, both at the level of the phenotype (Badyaev 2002) and the transcriptome (e.g., Parisi et al. 2004). Some evidence is also available in support for adaptation in the genetic basis of dimorphic traits. Bonduriasky and Rowe (Bonduriansky and Rowe 2005) showed that in the fly *P. xanthostoma*, the inter-sexual genetic correlation between homologous traits in males and females is inversely proportional to the degree of sexual dimorphism of the traits. This was interpreted as a break-down of genetic correlations under the influence of divergent selection on the two sexes at the level of the transcriptome. Furthermore, comparative transcriptomic studies in fruitflies have shown that the identity of sex-biased genes and the degree to which their expression differs between males and females changes along the phylogeny (Zhang et al. 2007). Assuming that the evolution of sexual dimorphism involves a temporary phase of sexual antagonism, these results

would indicate that resolution of antagonism occurs readily, at least on an evolutionary timescale. Experimental evidence from plants further suggest that resolution can occur relatively rapidly. Using artificial disruptive selection on flower size in male and female of *Silene latifolia*, Delph et al. (2011b) were able to significantly reduce the previously strong intersexual genetic correlation between these traits after a mere five generations of selection.

The data described above paints a picture of a dynamic and evolvable genetic architecture of male and female traits. This strikes a puzzling contrast with the antagonism that has been documented in populations of a growing number of organisms. It is currently unclear how these two observations could be reconciled. One possible solution would be that extant dimorphism reflects traits and genetic loci for which the genetic un-coupling between male and female phenotypes is relatively easy (van Doorn 2009). Sexual antagonistic variation would then be made of polymorphism at loci for which sex-specific is more difficult or impossible to evolve, for example because of deleterious pleiotropic effects (Mank et al. 2008). Alternatively, the antagonism we observe might reflect a dynamic equilibrium between rapid resolution of antagonism, combined with the input of new antagonistic variants, either at the same or different loci (Morrow et al. 2008; van Doorn 2009). To understand the resolution of sexual antagonism and its dynamics we must not merely document changes in dimorphism or genetic architecture of individual traits, but also take into account the selective pressures on these traits. Only then will it be possible to establish a causal link between divergent selection and a change in genetic architecture of traits across the sexes.

In this article we present evidence for a change in genetic architecture that is driven by divergent selection on males and females. We compare the genetic correlation between male and female fitness in two replicates of a laboratory population of *Drosophila melanogaster*, LHm. Both were established from the original LHm population that is maintained in the laboratory of W. Rice at the University of California at Santa Barbara (hereafter 'LHm-UCSB'). LHm-UCSB had been used in a number of pioneering studies documenting sexually antagonistic genetic effects (Rice 1984) and standing antagonistic variation (Chippindale et al. 2001). The two population studies here are descendant of the LHm-UCSB populations, currently maintained in laboratories at UCL (LHm-UCL) and the University of Uppsala (LHm-UU). Importantly for the comparison between the

different replicates, all LHm populations are maintained under identical, tightly controlled and repeatable conditions and fitness is measured in assays that mimic this rearing regime as far as possible.

The study presented here compares existing fitness data for LHm-UU (Innocenti and Morrow 2010) with newly generated data for LHm-UCL. Both datasets contain measures of male and female fitness across large samples of randomly drawn genotypes from each of the populations. The samples allow us to estimate and compare the genetic architecture of fitness in these two population that evolve independently under near-identical environmental selection pressures. In order to cross-validate fitness measures obtained from the two populations, a number of genotypes from the LHm-UU study were included in the fitness assays conducted on LHm-UCL. Furthermore, we complemented the quantitative genetic comparison with a population genetic analysis, based on microsatellite genotypes of flies sampled in both populations. Estimates of genetic diversity and differentiation allow us to make inferences about the genetic histories of the populations and rule out catastrophic losses of genetic diversity that would affect genetic architecture of fitness.

A2.2 Materials and Methods

A2.2.1 Study populations

The populations used here are descendants of the LH population, established in 1991 by L. Harshman from 400 wild-caught females and subsequently maintained in the laboratory, at large population size. In 1996, the LH population was put onto a controlled and standardised 14-day rearing regime with constant larval and adult densities and has since been maintained under identical conditions in the laboratory of W. Rice (University of California Santa Barbara) under the name LHm (here LHm-UCSB). In this study, we analysed data for two independent descendants of LHm-UCSB, LHm-UCL at University College London and LHm-UU at the University of Uppsala. The first is derived from a duplicate of LHm-UCSB that was taken to Queens University (Kingston, Canada) by A. Chippindale in February 2002. Subsequently, a duplicate of this population was shipped to the Reuter group, University College London (UCL) in May 2009 to establish the LHm-UCL population used in the present study. Independently, a replicate of LHm-UCSB was taken to the Morrow group, University of Uppsala (UU) in December 2005 to establish the other population analysed here, LHm-UU (Fig. A2.1).

Starting with the establishment of LHm-UCSB in 1996, all LHm populations have been maintained under an identical, strictly regimented rearing regime. Each population consists of 56 vials, each of which contains 150-200 eggs at the start of a generation. Eleven days later, the newly eclosed adults from these vials are mixed and placed into 56 ‘adult competition’ vials in groups of 16 sexually mature males and 16 mature females. After 48 hours in these vials, adult flies are transferred to the ‘larval competition’ vials in which they can lay eggs for 18 hours. At the end of the oviposition period, flies are removed and the egg density is standardised to 150-200 eggs to grow the next generation. A more detailed description of the rearing conditions can be found in Rice et al. (2005).

A2.2.2 Genetic architecture of male and female fitness

A2.2.2.1 Genome extraction

We used hemiclonal analysis to measure the effects of haploid genomes on male and female fitness (see Abbott and Morrow 2011for a review of the approach). A hemiclone is a group of individuals that have a copy of the chromosomes X, II and III in common, thus sharing 99.5% of an identical haplotype (all genes except for the 0.5% of the genome located on the 'dot' fourth chromosome). We created 113 hemiclonal lines from the LHm-UCL population and we used the previously published data from 100 lines in the LHm-UU population. In order to extract and manipulate X-II-III chromosome sets, we use a 'clone-generator' stock (more details in Chippindale et al. 2001; Rice et al. 2005; Abbott and Morrow 2011). Females of the clone-generator stock carry a Y chromosome, an attached X (CD(1)DX, *y*, *f*) and a translocation of chromosomes II and III (T(2;3) *rdgC st in ri* p^{P} *bw*). The attached X consists of two X chromosomes that co-segregate together in females, enabling a father-to-son transmission of the X chromosome (and a mother-to-son transmission of the Y chromosome). The translocated chromosome II and III enforces the co-segregation of those two chromosomes at each generation. These properties allow us to manipulate X-II-III chromosome sets. When maintained in males (which, in *Drosophila*, do not have recombination), these chromosome sets will remain intact. X-II-II sets were randomly sampled from LHm-UCL by crossing individual LHm-UCL males with a virgin clone generator female and back-crossing a single randomly chosen male offspring of this cross to another virgin clone-generator female. All male offspring of this second cross carry an identical X-II-III chromosome set and can be multiplied and maintained by further back-crosses to clone-generator female ('cytogenetic cloning').

A2.2.2.2 Fitness measurements for LHm-UCL

LHm-UCL hemiclone lines were established in August 2008 and their fitness was measured between July 2010 and September 2011. For all lines, fitness was assayed three times in each sex. Experiments were conducted in a blocked design. In each block, we measured fitness of one sex of all hemiclonal lines under investigation and we alternated assays of male and female fitness. To measure fitness in hemiclones we (i) performed crosses to express the target haplotypes in an outbred genetic background and the appropriate sex and (ii) measured the fitness of the individuals carrying the target hemiclones in conditions similar to those of the rearing regime. In addition to the 113 hemiclonal lines created from the LHm-UCL population, we also assessed the fitness of nine of the most sexually antagonistic lines created in the LHm-UU

population (Fig. 1 in Innocenti and Morrow 2010). This allowed us to directly compare the fitness of these lines when measured at UCL or at the University of Uppsala.

Female fitness assay: We created females carrying the target haplotypes and a LHm-UCL background by crossing 10 hemiclone carrier males with 15 LHm females. These parental flies were tossed onto new vials every day for three consecutive days and egg density was standardised to 150-200 in each vial, as in the LHm rearing regime. Virgin hemiclone females emerging from these crosses were collected on a single day, corresponding to 9-11 days after egg laying, depending on the vial of origin. The following day, we set up the 'adult competition' vials containing 10 virgin target females, 20 virgin LHm-*bw* competitor females and 30 virgin LHm-*bw* males. Fly density and yeast amount were doubled compared to rearing regime conditions, in order to reduce sampling variance and obtain more repeatable fitness scores. Forty-eight hours later, target and competitor females were anaesthetized on CO_2 and isolated individually in new vials to lay eggs. After 19.5 hours (1.5 hour more than the 18 hours egg-laying period to compensate for the post-anaesthesia recovery time), females were removed and the vials were stored for offspring to complete their development. Once the offspring had emerged, vials were frozen and the offspring counted. Raw female fitness scores were calculated as the average number of progeny produced by the hemiclone females of the same 'adult competition' vial of origin. Across the three blocks of female fitness assays that were performed in the UCL population, we measured the fitness of a total of 30 individual females per hemiclone line.

Male fitness assay: To introduce the target haplotypes into a male LHm-UCL background, 10 hemiclone carrier males were crossed with around 30 females of a DX-LHm stock. These DX-LHm females carry an attached X in a LHm background (Chippindale et al. 2001) and allow for father-to-son transmission of the X chromosome. As for the female assays, vials from these crosses were tossed on to fresh vials for three consecutive days. Due to the compound X, around half of the eggs laid by DX-LHm females are not viable. Therefore the egg density of those vials was standardised to twice the normal density, i.e. 300-400 eggs per vial. Virgin males carrying the target haplotype were collected on a single day, corresponding to 9-11 days after egg laying, depending on the vial of origin. The following day, we set up the 'adult competition' vials consisting of 10 virgin target males, 20 virgin LHm-*bw* competitor males and 30 virgin LHm-*bw* females. Once again, the fly density and yeast amount

was doubled compared to the LHm rearing regime to reduce sampling variance. Males and females were let to interact during 66 hours, corresponding to the 48 hours of adult competition and 18 hours of oviposition in the LHm rearing regime. Subsequently, females were isolated in vials containing yeast and let to lay eggs for at least 30 hours. Females were then removed and offspring allowed to complete their development. Vials containing the emerged offspring were frozen, after which offspring checked for eye-colour (wildtype or *bw*) and counted. Raw male fitness scores were the average proportion of offspring produced by target hemiclone males (with wildtype eye-colour) over the total number of offspring that emerged in all 30 egg-laying vials. Over the three blocks of male fitness assays that were performed in the LHm-UCL population, we measured the fitness of a total of 30 individual males per hemiclone line.

A2.2.2.3 Fitness data for LHm-UU

The dataset for the LHm-UU population comprised fitness measures obtained from 100 hemiclones extracted from LHm-UU in October 2007 and had previously been used to investigate the relationship between sex-specific fitness and gene expression (Innocenti and Morrow 2010). Fitness data were obtained in a similar manner as described above for the UCL population. Small differences included that fitness trials were performed on half the number of flies (five target individuals in competition with 10 *bw* flies) per competition vial. In addition, flies in the male assay were allowed to interact for 48+9 hours (instead of 48+18, as in the UCL assays) before laying females were isolated (more details in Innocenti and Morrow 2010). Six male assays and four female assays were performed in the LHm-UU population, thus testing fitness of a total of 30 individual males and 20 individual females per hemiclone (compared to 30 of each in the LHm-UCL dataset).

A2.2.2.4 Statistical analysis of fitness

We used the fitness data to estimate the contribution of additive genetic effects of hemiclones to the variation in male and female fitness, as well as the covariance between these genetic effects on fitness in males and females. Before analysis, raw male and female fitness scores were standardised by a transformation to a z-score to facilitate comparisons between sexes and populations. Prior to further analysis we also removed one outlying hemiclone from the UCL dataset. This genome had a very low male and female fitness (see Fig. A2.2), compatible with the effects of a strongly deleterious mutation in both males and females. Note that removing this data point was

conservative with respect to the quantitative genetic estimations made here because, if included, the outlier hemiclone artificially increased the estimates of heritabilities and the intersexual genetic correlation.

With male and female fitness scores transformed separately into z-scores, average fitness in both sexes is zero and we do not need to account for difference in male and female average fitness. We would, however, like to remove differences in average fitness between assay blocks. We do so by including fitness assay as a fixed effect into our analysis. Components of genetic variance and co-variance were then estimated by analysed the following linear mixed model

$$y_{i,j,k} = \beta_k + b_{i,j} + \varepsilon_{i,j,k},$$

where $y_{i,j,k}$ is the standardised fitness measured for sex i in genome j in the assay k, β_k is the fixed effect describing the average deviation of the average fitness score in assay k (with the index k running through male and female assays) from zero, the average standardised fitness score, $b_{i,j}$ is the random effect of the genome j within sex i with $b_{i,j}$ $\sim N(0, \sigma_i^2)$, σ_i^2 being sex specific, and $\varepsilon_{i,j,k} \sim N(0, \sigma_r^2)$ is the error term associated with this particular fitness measure. This linear model was run and analysed with the help of procedures implemented in the function lme (library nlme, Pinheiro et al. 2011) in R version 2.15.1 (R Development Core Team 2006). We estimated model parameters using REML and allowed for different residual variances in males and females with the option varIdent (Pinheiro and Bates 2000).

Heritabilities of male and female fitness were calculated from the variance components as $h_i^2 = 2\ \sigma_{i,g}^2/(\ \sigma_{i,g}^2 + \sigma_{i,r}^2)$ (Becker 1992), where $\sigma_{i,g}^2$ and $\sigma_{i,r}^2$ are the genetic and residual variances for sex i. The intersexual genetic correlation (r_{MF}) was calculated as $cov(f,m)/ (\sigma_{f,g} * \sigma_{m,g})$, where cov(f,m) is the genetic covariance between male and female fitness.

To further investigate the shape of intersexual correlation in both populations, we broke down the total genetic covariance into covariance along a sexually concordant and along a sexually antagonistic axis and for each population determined the percentage of covariance along these two axes. We estimated a confidence interval of these proportions of variance with a jackknife method. In each set of simulations, we created

10 datasets randomly removing 10% of the data set, thus creating 10 values of positive and negative male-female covariance. We then ran 100 of these simulations randomly assigning the groups of genomes to be removed.

For the LHm-UCL population, the data collected in the fitness assays was more complete than for the LHm-UU population and the female fitness assays included both the number of eggs laid by target females and by their *bw* competitors. A linear model showed that across assays, the standardised raw female fitness scores of target females in a vial co-varied significantly (and positively) with the average number of offspring produced by the competitor LHm-*bw* females in the same vial (linear regression of fitness z-score as a function of competitor fecundity; competitor fecundity: $F_{1,334}$ = 110.28, $p<0.001$). This result suggests that some variation in female fitness is due to between-vial effects that influenced target and competitor females to equal measures. For selected analyses of the LHm-UCL dataset we removed these environmental effects by using an alternative measure of female fitness. The vial-corrected adjusted fitness scores were calculated as the residuals of a linear regression of fitness z-score as a function of competitor fecundity. These residuals were then used as adjusted female fitness values in the mixed model described earlier to estimate the variance-covariance matrix of genetic effects on male and female fitness.

A2.2.3 Genetic diversity and divergence between both populations

We performed a microsatellite analysis to assess the level of genetic diversity within the two populations and the degree to which they had diverged genetically.

A2.2.3.1 Microsatellite genotyping

Ninety-six female adult flies were sampled from each of the populations. Flies from the LHm-UU population were collected in April 2011, those from the UCL population in June 2011. Total genomic DNA from homogenised individual flies was extracted using DNeasy® Blood and Tissue Kit (Qiagen) according to manufacturer's protocol. We genotyped the flies at 23 microsatellite markers (Supplementary Table A2.S1), selected from previously published studies and pilot analyses based on their genomic location and their polymorphism in the study populations. PCR reactions (20µl) contained 1x PCR buffer, 2.5mM $MgCl_2$, 0.05mM of total dNTPs, 0.15µM each of forward and

reverse primers, 5% Trehalose, 0.1µl Platinum® *Taq* DNA Polymerase (Invitrogen) and 2µl (ca. 1-2ng) of suspended gDNA. PCR amplification was carried out in a MJ Research PTC-200 thermal cycler with the following program: initial denaturation at 94 C for 2 min followed by 30 cycles with 94 C for 30 s, 55 C annealing temperature for 30 s, 72 C for 20 s and finishing with an elongation step at 72 C for 1 min. Marker resolution was performed using an ABI 96-capillary 3730xl DNA Analyzer and GeneMarker software (v.2.2.0). Allele sizes at all loci were estimated relative to an internal lane GS500 LIZ size standard.

A2.2.3.2 Statistical analysis of microsatellite data

In order to compare the levels of genetic diversity in the two populations, we used FSTAT (Goudet 1995) to calculate expected heterozygosity (H_e), allele number per locus (N_A) and allelic richness (AR, a standardised measure of the number of alleles per locus independent of sample size) separately for each locus in each population. These values were compared with a paired t-test.

We also calculated F_{ST} as a measure of genetic differentiation between the two populations. The estimate was tested against a null expectation of no differentiation using the permutation test implemented in FSTAT. We used a test that does not rely on the assumption of random mating within populations and ran 5000 permutations. Because the F_{ST} statistic is based on population genetic models that assume an infinite allele model (Weir and Cockerham 1984), it is not ideally suited for microsatellites that tend to follow a stepwise mutation model. Therefore, F_{ST} can underestimate the degree of genetic differentiation between populations (Slatkin 1995). For this reason we also calculated R_{ST}, a measure of genetic differentiation that assumes a stepwise mutation model. This was done using the R_{ST} Calc software (Goodman 1997).

A2.3 Results

A2.3.1 Genetic architecture of male and female fitness

For the nine most antagonistic hemiclonal lines from the LHm-UU population that were tested at the University of Uppsala and UCL, both sets of fitness assays showed significant positive correlation ($r = 0.813$, $p = 0.008$, Fig. A2.2). For males, where fitness has greater residual variation (Merila and Sheldon 2000; Pischedda and Chippindale 2006), the correlation between fitness scores obtained in University of Uppsala and UCL was also positive and close to significant ($r = 0.641$, $p = 0.063$, Fig. A2.2). This indicates that fitness can be measured in a repeatable manner in different laboratory settings and that the metrics of fitness used in the two datasets were comparable.

Re-analysing the fitness data for the LHm-UU population, we found a fitness heritability of 0.709 (CI: 0.533; 0.907) in females and of 0.193 (0.096; 0.367) in males. These figures are slightly higher than the heritability estimates obtained in the previous analysis of the UU dataset (Table 1, Innocenti and Morrow 2010). This improvement is consistent with the fact that the method used here to estimate genetic effects on fitness removed some environmental variance by using a statistical model that accounted for the environmental effect of the assay. While both heritability estimates increased, the estimate of heritability for female fitness remained significantly higher than for male fitness (Table A2.1).

Heritabilities of female and male fitness estimated in the LHm-UCL population were 0.394 (0.224; 0.648) and 0.407 (0.235; 0.658) respectively. Contrary to what was found in LHm-UU, these estimates were not significantly different from each other (Table A2.1) but the estimate of female heritability in LHm-UCL was significantly lower than in the LHm-UU population (Table A2.1). For the LHm-UCL population, we also calculated heritability using alternative female fitness scores that removed environmental variation between female competition vials, captured by the fecundity of LHm-*bw* competitor females. Using these vial-corrected adjusted fitness scores in the LHm-UCL population, we obtained an estimate of heritability for female fitness that

was not significantly different from the female heritability calculated for LHm-UU (Table A2.1).

In addition to estimating heritabilities, we calculated the intersexual genetic correlation for fitness in both populations. A negative correlation of -0.406 (-0.696; -0.001) between male and female fitness was confirmed in the LHm-UU population (Table A2.1). In contrast, the intersexual genetic correlation estimated for the LHm-UCL population was not different from zero (r_{MF}=0.129; CI: -0.294, 0.510, Fig. A2.2) and significantly different from that in LHm-UU (Table 1). Removing environmental variance in female fitness measures by using adjusted fitness scores resulted in a slightly narrower confidence interval but did not produce a lower or even negative estimate of the correlation coefficient (r_{MF}=0.219; -0.168; 0.548).

The change in the genetic architecture of fitness between the populations reflected in the altered intersexual correlation can further be illustrated by projecting male and female fitness values onto a coordinate system consisting of sexually concordant and antagonistic axes. For LHm-UU, 33.6 (28.7; 38.4) % of the fitness variation between genomes falls onto the sexually concordant axis, while 66.4 (61.9; 71.3) % fall onto the antagonistic axis. In the UCL population, these proportions had shifted to 56.4 (49.0; 65.0) % of concordant variation and 43.6 (35.0; 51.0) % of antagonistic variation.

A2.3.2 Genetic diversity and divergence

We did not detect any differences between LHm-UU and LHm-UCL in the level of standing genetic diversity within the populations. Across loci, the populations did not differ in any of the three measures of diversity, expected heterozygosity ($H_{e,UU}$=0.371±0.228 (mean across loci ± SD), $H_{e,UCL}$=0.402±0.216, difference $H_{e,UU}$-$H_{e,UCL}$=-0.031±0.175; paired t-test: t=-0.85, df=2, P=0.41), allele number ($N_{A,UU}$ = 2.652±1.071, $N_{A,UCL}$ = 3.000±1.537, difference $N_{A,UU}$-$N_{A,UCL}$=-0.348±1.335; paired t-test: t=-1.25, df=2, P=0.22) or allelic richness (AR_{UU}=2.588±1.002, AR_{UCL}=2.933±1.447, difference AR_{UU}-AR_{UCL}=-0.349±1.264; paired t-test: t=-1.32, df=2, P=0.20).

Whilst the level of genetic diversity was similar, tests showed that the two populations had diverged in allele frequencies. Across all loci, differentiation was evident in a value

of F_{ST}=0.236 that was significantly different from zero (permutation test, P<0.0002). The value of R_{ST} was similar (R_{ST}=0.202) and also significant (P<0.001). At the level of individual loci, all but three markers (114, 7 and 89; Fig. A2.3; Supplementary Table A2.S1) showed significant allele frequency differences between the two populations. The loci, even those with significant frequency divergence between populations, varied considerably in the value of F_{ST} with markers 4, 117 (on chromosome X), 24, 121, 28 (on chromosome II) and 60 (in chromosome III) showing the highest F_{ST} values (Fig. A2.3). However this variation in F_{ST} values was not associated with chromosomal location. F_{ST} values per locus did not differ between chromosomes X, II and III (Kruskal-Wallis test, Chi^2=3.61, df=2, P=0.16), nor were they higher on the X chromosome than the autosomes (Kruskal-Wallis test, Chi^2=0.11, df=1, P=0.74).

A2.4 Discussion

In this study we have demonstrated that sexual antagonism can be a very dynamic force as two populations evolving with the same selection conditions showed different pattern of intersexual genetic correlation of adult fitness after about 200 generations of independent evolution (Table A2.1). After a relatively short period of separation, resolution of sexual antagonism was observed in a replicate of the LHm population evolving in laboratory conditions. It is interesting to note that this resolution was recorded despite the fact that some natural selective forces (such as predation) which may be expected to be in the same direction for males and females were removed. By removing such natural selective forces, one might have expected sexual antagonism to be strengthened making resolution more difficult to be achieved than under natural conditions (Chapman et al. 2003; Bonduriansky and Chenoweth 2009). A microsatellite analysis revealed that the resolution of overall genomic sexual antagonism was due to parallel evolution, but no difference in genetic diversity between populations could be observed.

Our results also showed different heritabilities of male and female fitness between both LHm replicates (Table A2.1). Previous work showed that male fitness in the LHm population is strongly influenced by genes on the Y chromosome (Chippindale and Rice 2001). In our experimental design, the Y chromosome is transmitted to males by the DxLHm females. We have not taken into account the consequence of any polymorphism of the Y chromosome in the DxLHm population. However, we know that our DxLHm stock has been maintained in varied population sizes with limited gene flow. The increase of male fitness heritability in LHm-UCL could be an artefact of the conditions of maintenance of the DxLHm stock. The cause of the change in female fitness heritability is more difficult to identify. However, we have been able to measure the limited effect of this decrease of heritability in LHm-UCL by removing the environmental effects arising from variation in the fecundity of female competitors. This implies that that the lack of intersexual genetic correlation in the LHm-UCL population was not due to a decrease in female fitness heritability.

Our study also identified a candidate chromosomal area that could carry most of the genes or regulatory sequences responsible for sexual antagonism (Fig. A2.3). We note

that, despite some divergence between both populations in the X chromosome, most of the observed divergence was associated with a region on the arm of an autosome. Theoretical models make contrasting claims that genes affecting sexual antagonism should mainly be carried by sexual chromosomes (Rice 1984; Jordan and Charlesworth 2012) or autosomes (Fry 2010). Our data provides an indication that autosomes could indeed harbour a non-negligible proportion of genetic variation underpinning sexual antagonism. Interestingly, the area of most divergence between the LHm-UU and LHm-UCL populations corresponds to the area where most of the expressed genes of antagonism were found in the previous study of LHm-UU (Innocenti and Morrow 2010).

Theoretical models have long made contrasting claims about whether genes affecting sexual antagonism should mainly be carried by sexual chromosomes (Rice 1984; Jordan and Charlesworth 2012) or autosomes (Fry 2010). Our data provides an empirical indication that autosomes could indeed harbour a non-negligible proportion of genetic variation underpinning sexual antagonism (Fry 2010). First, no significant differences in *FST* values were found between the X chromosome and the autosomes suggesting that the resolution of sexual antagonism might not be due to sex chromosome linkage. Previous empirical evidence in *Drosophila* had already suggested that autosomal-linked loci might be responsible for sexually dimorphic phenotypes such as abdominal bristle pigmentation (Williams and Carroll 2009). Furthermore, 68 genomic regions enriched for sexually antagonistic loci located both in autosomes and X chromosomes had also been previously identified (Innocenti and Morrow 2010). And, interestingly, one of these genomic regions is located within an area of high genetic divergence between markers 121 and 28 (2L) in the LHm-UU and LHm-UCL populations (Fig. A2.3).

Our results support the notion that sexual antagonism can be viewed as a transient state of a population rather than being under long term constraints that cannot be overcome (van Doorn 2009). This study involved measures obtained after more than 200 generations of isolation between the replicates of the LHm-UCSB population of origin. It remains feasible that the lack of correlation between male and female fitness may have been present for an unknown number of generations in the LHm-UCL branch of the LHm population, as suggested by previous results (Pischedda and Chippindale 2006).

Our observations allow us to make informed speculations about the cause of the loss of sexual antagonism in LHm-UCL population. First, the LHm-UCL population may only represent a sample of the variation of fitness present in the LHm-UU population. In the LHm-UCL population there may have been fixation of alleles that benefitted either males, or females, or with fitness effects between those two extremes. However, this scenario is not supported by the patterns of our data. The fitness distribution of the extreme genomes of the LHm-UU population also falls within the extremes of the LHm-UCL population (Fig. A2.2). Furthermore, the microsatellite data showed a similar amount of genetic diversity in both populations.

Second, we cannot completely rule out the possibility that small differences between the rearing environments of LHm-UU and LHm-UCL may have been sufficient to increase the sexually concordant selection in LHm-UCL. This shift of selective forces may be related to uncontrollable minor variations that arose despite all possible precautions having been taken to ensure an identical culturing environment in both replicates of the LHm population. It is expected that sexually concordant forces would mainly be due to natural selection (Chapman et al. 2003). So that even if natural selection conditions changed between the two locations, the rearing conditions should emphasise sexual antagonism.

Thirdly, sexual antagonism may have been partially or completely resolved during the independent evolution of LHm-UCL. As the intersexual genetic correlation was neither positive nor negative, a partial resolution of sexual antagonism would mean that sexually antagonistic selection relatively decreased compared to sexually concordant selection, as illustrated by the projections on the antagonistic- concordant axes. A complete resolution of sexual antagonism would have occurred if sexual dimorphism reached such an extent that male and female fitness became completely independent. Unfortunately, the study of r_{MF} alone is not sufficient to disentangle both hypotheses, which if caused by the same phenomena (increase in sexual dimorphism), would have different consequences in the short and long term evolution of the LHm-UCL population. Sexual dimorphism can rapidly evolve. Artificial disruptive selection showed that a trait with a high r_{MF} could become undetectable in fewer than five generation (Delph et al. 2011b). More generally, it is entire G matrices than can shift within a few generation of drift (Whitlock et al. 2002). Several processes can increase sexual dimorphism and reduce sexual antagonism. Gene duplication followed by sex-

specific expression is a lengthy process that is unlikely to have occurred within the relatively few generations used here (Näsvall et al. 2012). Alternative sex-splicing can be controlled by sex-specific regulators that could be easily modified within a medium-term evolution as seen in this study and could easily be tested by microarray analysis. The mechanisms underlying genomic imprinting that could result in the resolution of sexual antagonism are less well understood. However, in our protocol, both males and females receive a copy of an X chromosome from the father. Thus, if genomic imprinting was present in the X chromosome it would not be detectable in our study. Moreover, the flies we tested received half of their genome from mothers from different populations (LHm-UCL for females, DxLHm-UCL for males). Thus, resolution through genomic imprinting should have evolved in both populations (LHm-UCL and DxLHm-UCL). Altogether, the unusual pattern of inheritance tested in our study makes it unlikely for genomic imprinting to be a strong hypothesis for the reduction of sexual antagonism.

Finally, the timing of resolution of sexual antagonism observed in LHm-UCL provides information on the potential mechanisms controlling sexual antagonism. Rapid resolution of sexual antagonism suggests a major role of few mutations which would have a big effect on male and female fitness . It is estimated that 8% of the *D. melanogaster* genome is sexually antagonistically expressed (Innocenti and Morrow 2010). It is unlikely that sex-specific beneficial mutations occurred on this 8% portion of the genome. A more likely scenario is that sexually antagonistic traits were already controlled by multiple sex-specific regulated genes , including sex-hormone receptors, sex-specific transcription regulators or Y-linked transcription regulators (Stewart et al. 2010). Thus, resolution may have occurred through a relatively small number of mutations targeting sex-specific regulatory mechanisms controlling sexually-antagonistically expressed genes.

In conclusion, our results reveal the potential for sexual antagonism to rapidly evolve and the necessity to consider sexual antagonism as a dynamic process which understanding can greatly benefit from long-term rather than one-of studies.

A2.5 References

Abbott, J. K., and E. H. Morrow. 2011. Obtaining snapshots of genetic variation using hemiclonal analysis. Trends Ecol Evol 26:359-368.

Badyaev, A. V. 2002. Growing apart: an ontogenetic perspective on the evolution of sexual size dimorphism. Trends Ecol Evol 17:369-378.

Becker, W. A. 1992. Manual of quantitative genetics. Academic Entreprises, Pullman, WA.

Bedhomme, S., A. K. Chippindale, N. G. Prasad, M. Delcourt, J. K. Abbott, M. A. Mallet, and H. D. Rundle. 2011. Male-limited evolution suggests no extant intralocus sexual conflict over the sexually dimorphic cuticular hydrocarbons of *Drosophila melanogaster*. Journal of Genetics 90:443-452.

Berg, E. C., and A. A. Maklakov. 2012. Sexes suffer from suboptimal lifespan because of genetic conflict in a seed beetle. Proceedings of the Royal Society B: Biological Sciences 279:4296-4302.

Betran, E., K. Thornton, and M. Long. 2002. Retroposed new genes out of the X in *Drosophila*. Genome Res. 12:1854-1859.

Bonduriansky, R., and S. F. Chenoweth. 2009. Intralocus sexual conflict. Trends Ecol Evol 24:280-288.

Bonduriansky, R., and L. Rowe. 2005. Intralocus sexual conflict and the genetic architecture of sexually dimorphic traits in *Prochyliza xanthostoma* (Diptera: Piophilidae). Evolution 59:1965-1975.

Brommer, J. E., M. Kirkpatrick, A. Qvarnström, and L. Gustafsson. 2007. The intersexual genetic correlation for lifetime fitness in the wild and its implications for sexual selection. PLoS ONE 2:e744.

Chapman, T., G. Arnqvist, J. Bangham, and L. Rowe. 2003. Sexual conflict. Trends Ecol Evol 18:41-47.

Chippindale, A. K., J. R. Gibson, and W. R. Rice. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. Proc. Natl. Acad. Sci. U.S.A. 98:1671-1675.

Chippindale, A. K., and W. R. Rice. 2001. Y chromosome polymorphism is a strong determinant of male fitness in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U.S.A. 98:5677-5682.

Colson, I., S. J. Macdonald, and D. B. DGoldstein. 1999. Microsatellite markers for interspecific mapping of *Drosophila simulans* and *D. sechellia*. Molecular Ecology 8:1951-1955.

Connallon, T., and A. G. Clark. 2011. The resolution of sexual antagonism by gene duplication. Genetics 187:919-937.

Delph, L. F., J. Andicoechea, J. C. Steven, C. R. Herlihy, S. V. Scarpino, and D. L. Bell. 2011a. Environment-dependent intralocus sexual conflict in a dioecious plant. New Phytol 192:542-552.

Delph, L. F., J. C. Steven, I. A. Anderson, C. R. Herlihy, and E. D. Brodie III. 2011b. Elimination of a genetic correlation between the sexes via artificial correlational selection. Evolution 65:2872-2880.

Ellegren, H., and J. Parsch. 2007. The evolution of sex-biased genes and sex-biased gene expression. Nat Rev Genet 8:689-698.

Foerster, K., T. Coulson, B. C. Sheldon, J. M. Pemberton, T. H. Clutton-Brock, and L. E. B. Kruuk. 2007. Sexually antagonistic genetic variation for fitness in red deer. Nature 447:1107-1110.

Fontanillas, P., D. L. Hartl, and M. Reuter. 2007. Genome organization and gene expression shape the transposable element distribution in the *Drosophila melanogaster* euchromatin. Plos Genet. 3:2256-2267.

Forsman, A. 1995. Opposing fitness consequences of colour pattern in male and female snakes. J Evolution Biol 8:53-70.

Fry, J. D. 2010. The genomic location of sexually antagonistic variation: some cautionary comments. Evolution 64:1510-1516.

Gallach, M., and E. Betrán. 2011. Intralocus sexual conflict resolved through gene duplication. Trends Ecol Evol 26:222-228.

Gallach, M., C. Chandrasekaran, and E. Betrán. 2010. Analyses of nuclearly encoded mitochondrial genes suggest gene duplication as a mechanism for resolving intralocus sexually antagonistic conflict in *Drosophila*. Genome Biology and Evolution 2:835-850.

Gay, L., E. Brown, T. Tregenza, D. Pincheira-Donoso, P. E. Eady, R. Vasudev, J. Hunt, and D. J. Hosken. 2011. The genetic architecture of sexual conflict: male harm and female resistance in *Callosobruchus maculatus*. J Evolution Biol 24:449-456.

Goldstein, D. B., and A. G. Clark. 1995. Microsatellite variation in north american populations of *Drosophila melanogaster*. Nucleic Acids Research 23:3882-3886.

Goodman, S. J. 1997. R-ST Calc: a collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. Molecular Ecology 6:881-885.

Goudet, J. 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. J. Hered. 86:485-486.

Hosken, D. J. 2011. Gene duplication might not resolve intralocus sexual conflict. Trends Ecol Evol 26:556-557.

Innocenti, P., and E. H. Morrow. 2010. The sexually antagonistic genes of *Drosophila melanogaster*. PLoS Biol 8:e1000335.

Jordan, C. Y., and D. Charlesworth. 2012. The potential for sexually antagonistic polymorphism in different genome regions. Evolution 66:505-516.

Kohorn, L. U. 1994. Shoot morphology and reproduction in jojoba - advantages of sexual dimorphism. Ecology 75:2384-2394.

Lande, R. 1980. Sexual dimorphism, sexual selection and adaptation in polygenic characters. Evolution 34:292-305.

Mank, J. E., L. Hultin-Rosenberg, M. Zwahlen, and H. Ellegren. 2008. Pleiotropic constraint hampers the resolution of sexual antagonism in vertebrate gene expression. Am. Nat. 171:35-43.

Merila, J., and B. C. Sheldon. 2000. Lifetime reproductive success and heritability in nature. Am. Nat. 155:301-310.

Mills, S. C., E. Koskela, and T. Mappes. 2012. Intralocus sexual conflict for fitness: sexually antagonistic alleles for testosterone. Proceedings of the Royal Society B: Biological Sciences 279:1889-1895.

Mokkonen, M., H. Kokko, E. Koskela, J. Lehtonen, T. Mappes, H. Martiskainen, and S. C. Mills. 2011. Negative frequency-dependent selection of sexually antagonistic alleles in *Myodes glareolus*. Science 334:972-974.

Morrow, E. H., A. D. Stewart, and W. R. Rice. 2008. Assessing the extent of genome-wide intralocus sexual conflict via experimentally enforced gender-limited selection. J Evolution Biol 21:1046-1054.

Näsvall, J., L. Sun, J. R. Roth, and D. I. Andersson. 2012. Real-time evolution of new genes by innovation, amplification, and divergence. Science 338:384-387.

Parisi, M., R. Nuttall, P. Edwards, J. Minor, D. Naiman, J. Lu, M. Doctolero, M. Vainer, C. Chan, J. Malley, S. Eastman, and B. Oliver. 2004. A survey of ovary-, testis-, and soma-biased gene expression in *Drosophila melanogaster* adults. Genome Biology 5:R40.

Patten, M. M., D. Haig, and F. Úbeda. 2010. Fitness variation due to sexual antagonism and linkage disequilibrium. Evolution 64:3638-3642.

Pinheiro, J., D. Bates, S. DebRoy, D. Sarpar, and R. D. C. Team. 2011. nlme: Linear and nonlinear mixed effects models.

Pinheiro, J. C., and D. M. Bates. 2000. Mixed-effects models in S and S-PLUS. Springer, New York.

Pischedda, A., and A. K. Chippindale. 2006. Intralocus sexual conflict diminishes the benefits of sexual selection. Plos Biology 4:2099-2103.

Proulx, S. R., and P. C. Phillips. 2006. Allelic divergence precedes and promotes gene duplication. Evolution 60:881-892.

R Development Core Team. 2006. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rice, W. R. 1984. Sex chromosomes and the evolution of sexual dimorphism. Evolution 38:735-742.

Rice, W. R., and A. K. Chippindale. 2001. Intersexual ontogenetic conflict. J Evolution Biol 14:685-693.

Rice, W. R., J. E. Linder, U. Friberg, T. A. Lew, E. H. Morrow, and A. D. Stewart. 2005. Inter-locus antagonistic coevolution as an engine of speciation: assessment with hemiclonal analysis. Proc. Natl. Acad. Sci. U.S.A. 102:6527-6534.

Robinson, M. R., J. G. Pilkington, T. H. Clutton-Brock, J. M. Pemberton, and L. E. B. Kruuk. 2006. Live fast, die young: trade-offs between fitness components and sexually antagonistic selection on weaponry in Soay sheep. Evolution 60:2168-2181.

Schug, M. D., K. A. Wetterstrand, M. S. Gaudette, R. H. Lim, C. M. Hutter, and C. F. Aquadro. 1998. The distribution and frequency of microsatellite loci in *Drosophila melanogaster*. Molecular Ecology 7:57-70.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457-462.

Stewart, A. D., A. Pischedda, and W. R. Rice. 2010. Resolving intralocus sexual conflict: genetic mechanisms and time frame. J Hered 101:S94-99.

van Doorn, G. S. 2009. Intralocus sexual conflict. Annals of the New York Academy of Sciences 1168:52-71.

Vinckenbosch, N., I. Dupanloup, and H. Kaessmann. 2006. Evolutionary fate of retroposed gene copies in the human genome. Proc. Natl. Acad. Sci. U.S.A. 103:3220-3225.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-Statistics for the analysis of population-structure. Evolution 38:1358-1370.

Whitlock, M. C., P. C. Phillips, and K. Fowler. 2002. Persistence of changes in the genetic covariance matrix after a bottleneck. Evolution 56:1968-1975.

Wyman, M. J., A. D. Cutter, and L. Rowe. 2012. Gene duplication in the evolution of sexual dimorphism. Evolution 66:1556-1566.

Zhang, Y., D. Sturgill, M. Parisi, S. Kumar, and B. Oliver. 2007. Constraint and turnover in sex-biased gene expression in the genus *Drosophila*. Nature 450:233-U232.

Table A2.1 Intersexual genetic correlation (r_{MF}) and female and male heritability (h^2) of fitness in both populations.

	female h^2 (CI)	**male h^2 (CI)**	**r_{MF} (CI)**
Fitness LHm-UU	0.709 (0.533; 0.907)	0.193 (0.096; 0.367)	-0.406 (-0.696; -0.001)
Fitness LHm-UCL	0.394 (0.224; 0.648)	0.407 (0.235; 0.658)	0.129 (-0.294; 0.510)
Adjusted fitness LHm-UCL	0.586 (0.402; 0.810)	0.408 (0.235; 0.659)	0.219 (-0.168; 0.548)

Figure A2.1 History of the different samples of the LHm population. A corresponds to the date of the first report of sexual antagonism in the LHm-UCSB population (Chippindale et al., 2001). B and C correspond to the date when genomes were extracted in the LHm-UU (Innocenti and Morrow, 2010) and the LHm-UCL population, respectively. The LHm-UU and LHm-UCL samples were separated in 2002, thus 150 to 220 generations before the extraction of genomes for the present study.

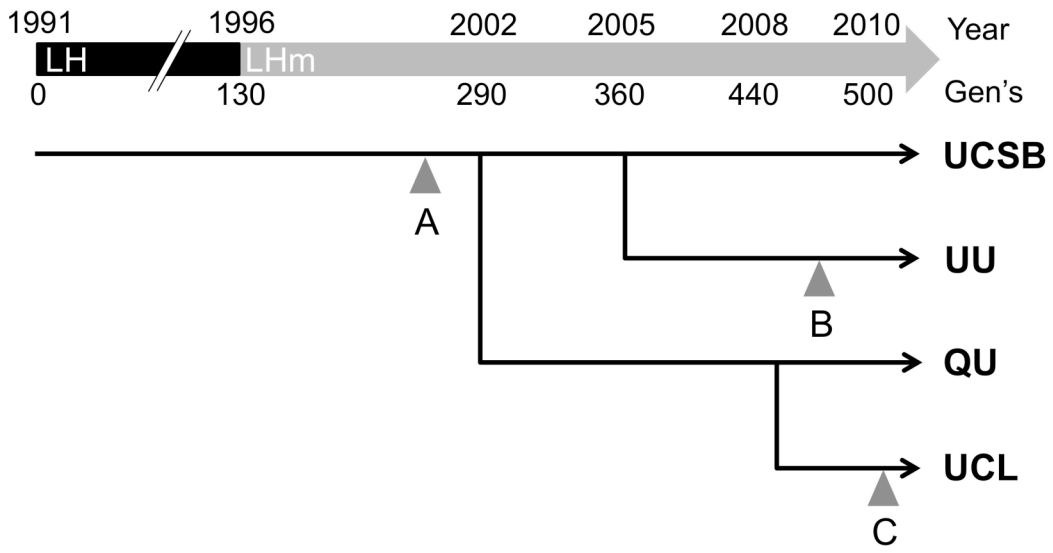


Figure A2.2 Average male and female adult fitness of 113 hemiclonal lines in the LHm-UCL population. The outlier hemiclone, represented by the filled circle in the bottom left corner, was removed from further analyses. For comparison, a group of LHm-UU hemiclones were re-tested at UCL, UK. The blue dots denote the hemiclones that, when originally assayed at Uppsala, Sweden, belonged to the male beneficial/female detrimental fitness class. The red dots show the UCL fitnesses of LHm-UU hemiclones that exhibited female beneficial/male detrimental fitness at Uppsala.

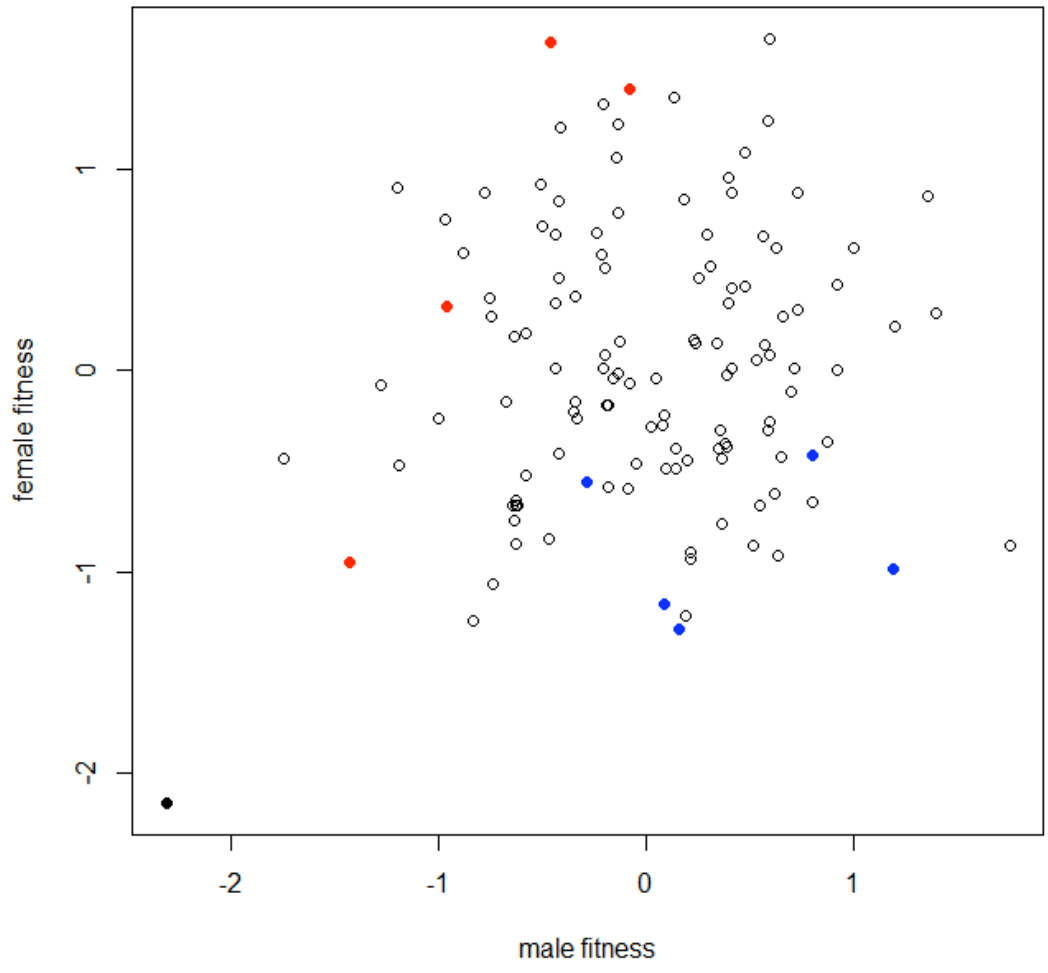


Figure A2.3 Schematic representation of F_{ST} values across microsatellite markers on arms of chromosomes X, II and III. The height of the bars indicate the value of the F_{ST} estimates for the different markers, marker identifiers are given next to the marker locations. Red dots indicate the centromeres.

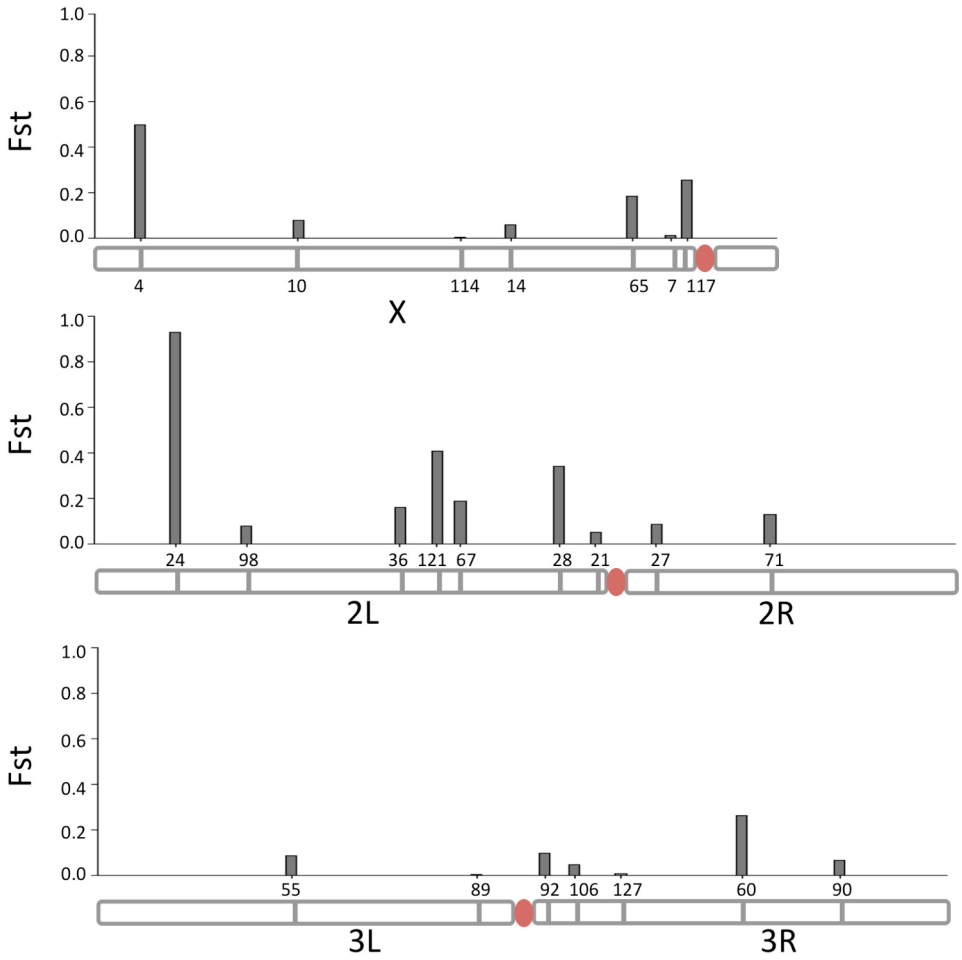


Table A2.S1 Microsatellite markers used in this study. The table provides the name, cytogenetic location (Gen. loc.), primer sequences, type of 5' primer dye (Dye), fragment length range (Length) and reference for each locus. The list is sub-divided into blocks corresponding to chromosomes X, II and III. Primers for loci 10 and 14 were modified and differ from the originally published sequence.

Locus ID	Locus name	Gen. loc.	Forward primer	Reverse primer	Dye	Length	Reference
X							
4	AF047180	1-0	tacctaggaaacccgaccc	tcttgttgcgaatttgttca	fam	270-272	Colson et al. 1999
10	DELTEX	1-17	acgcaataagttggcgta	aatcaggataatgcctaat	hex	121-131	after Schug et al. 1998
114	AC011068	1-38	agcagaagcggcaacatat	cggtaagaggctttggtgtc	fam	218-220	J. Gockel, unpublished
14	DROYP3	1-44	atccggctatttgcaatcaa	tctggcttatatagctcc	tet	191-201	after Schug et al. 1998
65	DMARIADNE	1-58	aacactgtcccccatccacat	tctgttcaactccttcggct	fam	119-133	Colson et al. 1999
7	AF017777	1-65	attagctaactccaagaacg	aatcctctcagctcagcgta	tet	164-168	Colson et al. 1999
117	AC011760	1-66	tggctttagtctggcctttg	caagccgaaaactaggca	hex	185-197	J. Gockel, unpublished
II							
24	AC004441	2-3	agaatacaaactcgattgcc	tgaaagtcaaatgctgctg	hex	162-174	I. Colson, unpublished
98	AC004373	2-12.5	aatgcgtgtgtttggatgaa	gtcccagtctcccagtgaa	fam	185-189	Colson et al. 1999
36	AC005555	2-31.5	ggttgctgggagaaagac	gccacacattcgcatctc	tet	153-173	Colson et al. 1999
121	AC005889	2-35	gcgtggctggcatatagagt	taagccccctcgtgtaattg	fam	187-191	J. Gockel, unpublished
67	DMU12269	2-39	tgggatccgtggatcatagt	attcgggaatgaggacact	fam	239-243	I. Colson, unpublished
28	AC004118	2-50.6	ccaacttgggcgagagaatt	gcttaattgcctcactgtgc	hex	288-298	Colson et al. 1999
21	AC004759	2-54.3	acagacggaaagccaaaatg	cactccgcctcgtttcttac	tet	222-225	Colson et al. 1999
27	AC006472	2-61	tcctccatgtaaagataaacgc	aactcgcaaattgcctaac	fam	263-285	Colson et al. 1999
71	AC004516	2-76	tcgtcgcccgttaatata	accgttcgtgggtcaaata	hex	252-270	Colson et al. 1999
III							
55	DMU14395	3-18	gggcagaggaaaagcactca	tcggtgagaccgtaatctg	fam	285-287	Colson et al. 1999
89	DMZ60MEX1	3-43	aaatctgttgctcatactgccc	aaccggcgaaatgttcag	hex	89-92	Goldstein and Clark 1995
92	DROPROSA	3-51	caataaccacacgcattcca	aaccacttcctgtttggcc	fam	108-122	Colson et al. 1999
106	DMTRXIII2	3-54.2	acaaaagccgaacgagaaaa	ctgctgcgttgtggctcctt	hex	276-278	S. J. Macdonald, unpublished
127	AC007647	3-58.5	ctttcctgcatctggctagg	atctgcggcgtcttctaatg	fam	189-199	J. Gockel, unpublished
60	DMU25686	3-73	cgataatttactctgtgctcc	cagctcacacaaaaggca	hex	132-150	Colson et al. 1999
90	DROROUGH	3-91.1	aagcaatgccacacaatgag	cggttatttttttccttggc	fam	72-81	I. Colson, unpublished