The genetics and evolutionary dynamics of sexually antagonistic polymorphisms in *Drosophila melanogaster*

Filip Ruzicka

A dissertation submitted in partial fulfilment

of the requirements for the degree of

Doctor of Philosophy

of

University College London

Department of Genetics, Evolution and Environment,

University College London

December 10th, 2018

I, Filip Ruzicka, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Signature……

December 10th, 2018

Abstract

The evolution of sexual dimorphism is constrained by a shared genome between males and females. This constraint can lead to 'sexual antagonism' where segregating alleles at given genetic loci have opposing fitness effects in each sex. Despite its wide taxonomic incidence, little is known about the identity, genomic location and evolutionary dynamics of sexually antagonistic polymorphisms. This is a major knowledge gap, since a better understanding of antagonistic polymorphisms can shed light on two fundamental questions: (i) how does the genome evolve to accommodate divergent and often contradictory selective pressures, and (ii) what evolutionary forces maintain genetic variation for fitness?

In this thesis, I describe the genetics and evolutionary dynamics of sexually antagonistic polymorphisms. I first highlight the limitations of previous genetic studies of sexual antagonism (Chapter 2). Specifically, I re-analyse a prominent study of antagonistic gene expression and show that inferences of antagonistic selection were driven by non-random population structure in the sample of genomes considered, rendering previous conclusions unreliable. I then present the first genome-wide association study of sex-specific fitness and sexual antagonism in a laboratory-adapted population of *D. melanogaster* (Chapter 3). I show that antagonistic variation disproportionately accumulates in coding regions but not on the X chromosome. I proceed to test whether sexually antagonistic selection maintains population genetic variation (Chapter 4), as has long been proposed but never tested. Consistent with this hypothesis, I find multiple signatures of balancing selection associated with antagonistic loci across

populations of *D. melanogaster* separated over 10,000 years, and possibly across species boundaries. Finally, I present experimental work testing whether a specific candidate gene—*fruitless*—is under antagonistic selection (Chapter 5). The results presented are consistent with balancing but not antagonistic selection.

Overall, this thesis underscores the fundamental difficulty of evolving genetic mechanisms that accommodate the divergent evolutionary interests of each sex.

Impact statement

The work described in this thesis has been widely disseminated within the academic community. It has been presented at national and international conferences, including The Centre for Ecology & Evolution's Summer Symposium 'War & Peace: The Dynamics of Evolutionary Conflict' (London, June 2016), the 16th Congress for the European Society for Evolutionary Biology (Groningen, August 2017), PopGroup 51 (Bristol, January 2018) and The Genetics Society's Autumn Symposium 'Genotype to Phenotype to Fitness' (Exeter, November 2018).

With regard to dissemination in scientific journals, a publication based on work presented in this thesis is currently in revision (see Appendix A), while two additional papers based on chapters presented in this thesis are in preparation. Future work building on the results presented here is also likely (see Chapter 6). Additionally, a recent publication (see Appendix C) and special issue on 'Local adaptation and the evolution of sex differences', published in *Phil. Trans. Roy. Soc. B.* (August 2018) was inspired by shared interest by members of the academic community in the theme of sexual antagonism and sex differences.

Finally, this work has inspired dissemination in the wider, non-academic community. An exhibition entitled 'Agonism/Antagonism' at the Grant Museum of Zoology (20th Sept. 2018 – 21st Dec. 2018) was created by Neus Torres Tamarit, an artist in residence in the Reuter laboratory.

Acknowledgements

Most of all, I would like to thank my primary supervisor, Max Reuter. His friendliness, guidance and critical comments have made this PhD an immensely enjoyable and instructive experience.

Thanks to my secondary supervisor, Kevin Fowler, for valuable discussions and comments throughout. Thanks to members of the Dros/Stalkie and associated labs—in particular, Lara, Mark and Flo—for their great company. Thanks also to students—Harvi, Olivia, Didem—for their tireless help. Thanks to some great scientists for enlightening discussions—Andrew Pomiankowski, Tim Connallon and Aida Andrés in particular.

Thanks finally to my examiners, Mark Thomas and Göran Arnqvist, for stimulating discussions during my viva, and critical suggestions for improvement.

Love to Mama and Tata and the family.

Contents

List of Figures

List of Tables

Chapter 1

General Introduction

1.1. Genetic constraints on the evolution of sexual dimorphism

Males and females are often strikingly dimorphic (Darwin 1871). In a seminal experiment, Bateman (1948) provided a key insight into the underlying cause by showing that different reproductive strategies maximise fitness in each sex. Specifically, Bateman showed that in fruit flies (*Drosophila melanogaster*), male reproductive success increases linearly with the number of matings, while in females it levels off after a limited number of matings. Trivers (1972) and later Parker (1979) generalised Bateman's findings to suggest that the production of unequally sized gametes in each sex (anisogamy) generates unequal energy costs and favours fundamentally different reproductive, ecological and social roles. Selection optimising phenotypes to these contrasting roles—sex-specific selection—is therefore widespread (Slatkin 1984; Andersson 1994; Arnqvist & Rowe 2005), and this gives rise to the contrasting morphologies and behaviours that are widely observed between the sexes in nature.

The fact that sexual dimorphism spans a dizzying array of phenotypes, from classic exaggerated traits like the peafowl's tail to genome-wide patterns of gene expression (Parsch & Ellegren 2013; Ingleby et al. 2015), seems also to indicate that genetic responses to sex-specific selection—*i.e.*, the evolution of dimorphic phenotypes—is easy.

Yet this is not necessarily so. The two sexes must respond to divergent selection pressures using *shared* genetic material. This is problematic because shared genes will tend to code for shared phenotypes, thus

generating strong positive genetic correlations between homologous traits in each sex.

The non-independence of trait expression in each sex can be readily illustrated by experiments where selection on a given trait is only applied in one sex (*i.e.* selection is applied sex-limitedly). Such experimental studies often find that the trait under selection not only responds in the selected sex, but also exhibits a correlated response in the opposite sex (Delph et al. 2004; Harano et al. 2010)—as expected if genes underlying trait expression are shared between the sexes. The degree of trait non-independence can further be quantified by estimating an 'intersexual genetic correlation' (r_{mf}), defined as:

$$r_{mf} = \frac{cov_{mf}}{\sqrt{\sigma_f^2 \sigma_m^2}}$$

with cov_{mf}, representing the additive genetic covariance between the sexes, and σ_f^2 and σ_m^2 representing additive genetic variances in females and males respectively. Experimental estimates of r_{mf} are frequently positive and often approach 1 (Cowley & Atchley 1988; Fairbairn & Roff 2006; Poissant et al. 2010; Griffin et al. 2013), again implying that most traits share a genetic basis in each sex.

These twin observations—ubiquitous positive intersexual genetic correlations and ubiquitous sex-specific selection—imply that, until genetic mechanisms evolve that allow genes to be sex-specifically expressed, the shared genome effectively *prevents* the independent evolution of phenotypes towards their sex-specific optima (Lande 1980). Another way of thinking about this problem is that each sex is a different environment for evolution

but the shared genome causes the ‘migration’ (sex) of alleles between each environment, thus constraining ‘local adaptation’ (sexual dimorphism).

This ‘genetic constraint’ on the evolution of sexual dimorphism invariably generates a mismatch between observed levels of dimorphism and levels of dimorphism favoured by selection (‘optimal’ dimorphism). In many cases, the mismatch is such that *opposing* trait values are favoured in each sex (Fig. 1.1). Such traits are known as ‘sexually antagonistic’ traits.

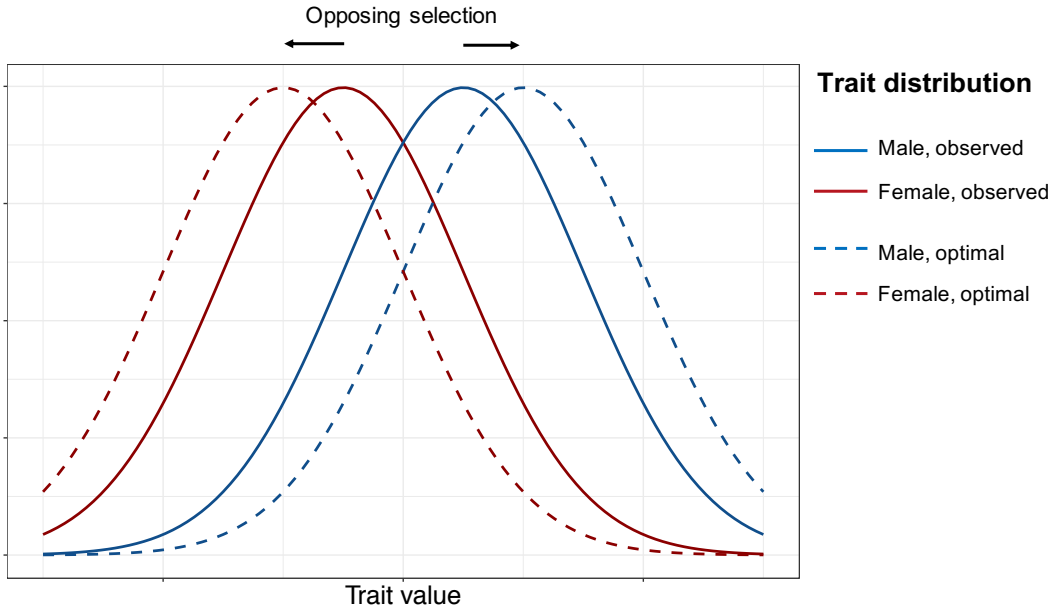


Figure 1.1. Trait distributions for a sexually antagonistic trait. Curves depict the population distribution of a trait in each sex. The trait is currently dimorphic, as made clear by divergent mean values in each sex (‘observed’, full curve). However, there is a mismatch between observed dimorphism and the level of dimorphism favoured by selection (‘optimal’, dashed curves). Each sex favours opposing trait optima (arrows) and the trait is therefore ‘sexually antagonistic’.

At the genetic level, sexually antagonistic traits are characterised by genetic loci harbouring alleles with opposing fitness effects in each sex (Bonduriansky & Chenoweth 2009; Van Doorn 2009). To understand why these loci— 'sexually antagonistic polymorphisms'—arise, consider a trait; say, human height. Consider also that both sexes are currently dimorphic but that this dimorphism is sub-optimal: selection favours taller males and smaller females (*i.e.*, height is a sexually antagonistic trait). It can be seen that any allele that increases height will be favoured in males. However, due to positive trait correlations (r_{mf}) between the sexes, this height-increasing allele will also be expressed in females and will be disfavoured in this sex. The alternative (height-decreasing) allele will, for the same reasons, be favoured in females but disfavoured in males. Consequently, each sex's inability to evolve independently towards its phenotypic optimum is invariably accompanied by the appearance of alleles with opposing fitness effects in each sex. These sexually antagonistic polymorphisms can be thought of as a genetic reflection of the constraints on the evolution of sexual dimorphism that are imposed by a shared genome.

1.2. Phenotypic evidence for sexual antagonism

Sexually antagonistic polymorphisms are expected to be common, since the two requirements for their evolution—sex-specific selection and positive trait correlations—are a pervasive feature of most sexual species. For example, evidence for sexual antagonism has been reported in seed beetles (*Callosobruchus maculatus*) (Gay et al. 2011), zebra finches (*Taeniopygia guttata*) (Price & Burley 1993), ground crickets (*Acanthoplus discoidalis*)

(Fedorka & Mousseau 2004), red deer (Foerster et al. 2007; Mainguy et al. 2009), collared flycatchers (*Ficedula albicollis*) (Brommer et al. 2007), snakes (*Vipera berus*) (Forsman 1995), lizards (*Uta stansburiana*) (Calsbeek & Sinervo 2004), plants (*Silene latifolia*) (Delph et al. 2004) and fruit flies (*D. melanogaster*) (Chippindale et al. 2001; Stewart & Rice 2018).

To establish the presence of sexually antagonistic polymorphisms, empiricists have employed a variety of approaches. Predominantly, these rely on detecting the indirect phenotypic effects of antagonistic genetic variation. The most common approach has been to search for a measurable mismatch between extant trait dimorphism and optimal trait dimorphism—more formally, to find evidence that traits exhibit strong positive intersexual trait correlations yet are subject to opposing selection pressures in each sex (Bonduriansky & Chenoweth 2009). For example, in humans, shorter families tend to have higher relative fitness through females, whereas taller families tend to have higher relative fitness through males (Stulp et al. 2012), implying that height is a sexually antagonistic trait. Similarly, in fruit flies (*D. melanogaster*), locomotory activity is highly genetically correlated between the sexes, yet there is opposing selection for this trait in each sex (Long & Rice 2007). Using this method, a whole suite of traits has been shown to be sexually antagonistic, including immunity (Vincent & Sharp 2014), body size (Merilä et al. 1997), wing shape (Abbott et al. 2010), lifespan (Berg & Maklakov 2012) and colouration (Roberts et al. 2009) (see Bondurianky and Chenoweth (2009) and Cox and Calsbeek (2009) for a review of the traits identified).

An extension of this approach has been to explicitly measure the intersexual genetic correlation for fitness (r_{mf}^{W}) across a sample of genotypes (r_{mf}^{W} is simply r_{mf} where the trait under consideration is 'fitness'). In contrast to the approach described above, which is focussed on traits, this approach establishes whether genome-wide genetic variation tends to have antagonistic fitness effects. If sexually antagonistic polymorphisms are common across the genome, genotypes with beneficial effects in males will tend to have detrimental effects in females, resulting in a negative r_{mf}^{W} (Rice & Chippindale 2001). Using this rationale, and consistent with the presence of sexual antagonism, Chippindale et al. (2001) showed that male and female fitness are negatively correlated among a sample of fly lines from a laboratory-adapted population of *D. melanogaster* flies. This result was replicated in later studies of the same population (Pischedda & Chippindale 2006; Innocenti & Morrow 2010) (but see Chapter 2). Using a half-sibling breeding design, Delcourt et al. (2009) similarly showed a negative r_{mf}^{W} among families sampled from two populations of *Drosophila serrata*, while Berger et al. (2014) detected a negative r_{mf}^{W} among isofemale lines of a Togolese population of seed beetles (*Callosobruchus maculatus*). Pedigree analyses have further reported negative r_{mf}^{W}s among populations of collared flycatchers (*Ficedula albicollis*) (Brommer et al. 2007) and red deer (*Cervus elaphus*) (Foerster et al. 2007).

A third method for establishing the presence of sexual antagonism is to apply sex-limited selection on fitness. This method, like the correlative method described above, focuses on detecting the fitness effects of genome-wide antagonistic variation. The rationale is that if sexually antagonistic

polymorphisms are common, restricting selection to a single sex should release sexually antagonistic variants favouring that sex from counter-selection in the other sex and allow them to increase in frequency. Accordingly, this selective regime should increase the fitness of the selected sex, but also cause a concomitant decrease in the fitness of the non-selected sex. Consistent with this rationale and the action of sexual antagonism, applying male-limited selection over many generations in *D. melanogaster* flies results in a fitness increase in males and a fitness decrease in females, as well as changes in various male reproductive traits (Rice 1996; Rice 1998; Prasad et al. 2007; Bedhomme et al. 2008; Abbott et al. 2010).

Taken together, the phenotypic studies conducted so far provide compelling evidence that sexual antagonism is pervasive in populations of many different organisms. Although not all studies provide such consistent evidence of sexually antagonistic variation (e.g. Morrow et al. (2008), Punzalan et al. (2014), Collet et al. (2016)), the data generally favours the view that sexual antagonism represents an important constraint to the evolution of sexual dimorphism.

1.3. Genetic evidence for sexually antagonistic polymorphisms

In contrast to the ample phenotypic evidence for sexual antagonism, little progress has been made towards describing its genetic basis. The handful of relevant studies that are available (Rowe et al. 2018) have addressed the genetics of antagonism using one of two opposing approaches. Some have studied antagonistic genetic variation on a genome-wide scale, providing

insights into the general properties of antagonistic variants. Others have identified and investigated individual antagonistic loci, giving us some examples of the precise mechanisms and traits that underlie antagonistic variation. I review these studies and their results in the following sections.

1.3.1. Genome-wide studies

Three studies have attempted to describe the genetic basis of genome-wide antagonistic variation. First, Innocenti & Morrow (2010) assayed male and female fitness of 100 *D. melanogaster* fly lines and measured gene expression among a subset of lines with extreme fitness effects spanning the 'antagonistic fitness axis'; that is, lines with highly male-beneficial female-detrimental (MB) and highly female-beneficial male-detrimental (FB) fitness effects respectively. Comparing gene expression profiles between MB and FB lines, the researchers identified ~1,300 candidate antagonistically expressed genes. Following from this work, Hill (2017) whole-genome sequenced the extreme fly lines used by Innocenti & Morrow (2010) and looked for fixed nucleotide differences between MB and FB lines, rather than gene expression differences. This follow-up study found ~6,000 variants that were fixed between MB and FB lines, and inferred that these differences represent candidate antagonistic polymorphisms.

In both cases, the researchers reported non-random functional patterns associated with their candidate genes. Innocenti & Morrow (2010) showed that candidate genes tend to be less sex-biased in expression than the genome-wide average, that candidate genes are disproportionately found on the X chromosome and that they are broadly expressed across tissues. Hill

(2017) additionally found associations between antagonistic polymorphisms and regulatory functions, including associations with key regulators of sexual differentiation in *D. melanogaster*, such as *fruitless*.

However, despite their promise, both studies suffer from a number of important limitations. First, Innocenti & Morrow's (2010) study focused on expression differences, so it could only identify correlates of the targets of antagonistic selection rather than the sequence variants themselves. For instance, a large number of genes could be found to have an antagonistic expression pattern, even though the actual target of antagonistic selection at the genetic level is a single variant in the coding sequence of a transcription factor. Second, the studies relied on a small sample of fly lines, which reduces power to detect true associations. Third, neither study corrected for relatedness or population structure among the fly lines analysed. Failure to correct for both factors can artificially decrease the number of false positives. Overall, these limitations mean that firm conclusions about the identity and functions of antagonistic polymorphisms based on these studies must currently be considered tentative (Rowe et al. 2018).

A third genome-wide study was conducted by Lucotte et al. (2016). They analysed samples from human populations comprising thousands of whole-genome sequences and looked for significant allele frequency differences between males and females ('intersexual F_{ST}'). Genetic differentiation between the sexes can arise if allelic variants cause differential mortality in males and females; polymorphisms with high intersexual F_{ST} values were therefore inferred to be under sex-specific selection. The authors identified a number of loci that were significantly differentiated

between the sexes and took them to be sexually antagonistic polymorphisms. From these polymorphisms, the authors indicated that antagonistic polymorphisms tend to be found on the X chromosome, and used this result to support classic theory which predicts that the X chromosome is a hotspot for antagonistic fitness variation (Rice 1984; Patten & Haig 2009). Compared to Innocenti & Morrow (2010) and Hill (2017), this study has two advantages: it focused on genetic polymorphisms rather than expression patterns, and it relied on much larger sample sizes. However, the method that was employed could not establish whether polymorphisms have high intersexual F_{ST} due to opposing selection in the two sexes (sexually antagonistic selection) or due to different strengths of sex-specific selection in each sex but no difference in sign (sexually concordant selection). Additionally, their design relied on differences in viability between the sexes to generate a signal. It therefore could not capture reproductive components of fitness—*i.e.*, sexual selection. Taken together, these limitations preclude firm conclusions about the identity and functions of antagonistic polymorphisms (Rowe et al. 2018).

1.3.2. Single locus studies

There have been three investigations of individual antagonistic loci. First, Roberts et al. (2009) examined the genetic basis of the orange blotch phenotype across various species of cichlid fish. The orange blotch phenotype is sexually antagonistic because it enhances female camouflage but disrupts male colour patterns that play a role in mate recognition and attraction. In this study, the researchers narrowed down the Quantitative

Trait Locus (QTL) for the phenotype to a regulatory region of the *Pax7* gene. Further examination of the region showed that it is closely linked to the female sex determining region (Roberts et al. 2009). Coupling to the female sex-determination region allows the orange blotch phenotype to be predominantly expressed in the sex which it confers most benefits to (females), effectively 'resolving' sexual antagonism. As such, this study provided strong evidence for a theorised link between sex-linkage and sexually antagonistic genes (Van Doorn & Kirkpatrick 2007).

In another study, Barson et al. (2015) mapped the genetic basis of age at maturation in salmon (*Salmo salar*). In this species, age at maturation is a sexually antagonistic trait, with males favouring earlier maturation age (owing to the benefits of decreased mortality before spawning) and females favouring later maturation age (owing to strong positive correlations between developmental time, body size and female fecundity in this sex). Analysing data across several populations, the authors found a strong association between age at maturation and variants at the *VGLL3* locus. They further showed that the early-maturation 'E' allele is dominant in males, while the late-maturation 'L' allele is dominant in females. Sex-specific dominance ensures that each sex matures closer to its phenotypic optimum, and is generally favourable for the maintenance of antagonistic fitness variation (Fry 2010). For this reason, this study represented an important first confirmation of a prominent genetic mechanism for alleviating sexual antagonism (Fry 2010; Spencer & Priest 2016).

Finally, experimental evolution among *D. melanogaster* fly populations carrying different versions of the *Cyp6g1* gene revealed that this insecticide

resistance polymorphism was sexually antagonistic (Smith et al. 2011; Hawkes et al. 2016). These studies are unique in having experimentally tested whether alternative alleles at a specific candidate antagonistic locus are associated with opposing fitness effects in each sex; as such, they represent valuable advances towards establishing that antagonistic loci identified through inferential approaches are truly causal.

Overall, these detailed investigations of single loci have been highly insightful, as they have provided the first empirical insights into the genetic mechanisms that favour the maintenance of antagonistic polymorphisms. Nevertheless, there are currently too few of these studies to permit generalisations about the genetic properties of antagonistic genes. This limitation, combined with the limitations of the genome-wide studies described previously, highlights the need for more extensive genetic investigations of sexual antagonism.

1.4. Open questions in the study of sexual antagonism

A better understanding of the genetics of sexual antagonism can shed light on a range of biological questions, including the evolution of sex determination (Van Doorn & Kirkpatrick 2007), the evolution of mating systems (Kirkpatrick & Hall 2004; Albert & Otto 2005), and the maintenance of genetic variation for fitness (Turelli & Barton 2004). An exhaustive description of the broader relation between sexual antagonism and other evolutionary processes can be found elsewhere (Bonduriansky & Chenoweth 2009; Van Doorn 2009). Here I focus on providing background to research questions that are relevant to this thesis. The questions discussed can be

loosely divided into two categories: (i) 'population genetic' questions, where the focus is on the relationship between sexually antagonistic selection, population genetic variation and quantitative genetic fitness variation, and (ii) 'genetic architecture' questions, where the focus is on the putative chromosomal location and functional effects of sexually antagonistic polymorphisms.

1.4.1. Population genetic implications of sexual antagonism

1.4.1.1. Population genetic variation

Ever since the first data on genome-wide polymorphisms became available (Lewontin & Hubby 1966), the question of which evolutionary forces determine the frequencies of genetic variants within populations has been a 'Great Obsession' among population geneticists (Gillespie 2004). This question remains the subject of intense study (Hahn 2008). Traditionally, three main mechanisms have been proposed to explain observed levels of genetic variation. The neutralist view posits that most genetic variation tends not to affect fitness, and that levels of variation are primarily determined by the equilibrium between the input of variation through new mutations and removal of variation through stochastic loss, or genetic drift ('mutation-drift balance') (Kimura 1983). Among *Drosophila* species, for example, short introns exhibit patterns of molecular variation consistent with the action of genetic drift (Parsch et al. 2010). A second model holds that genetic variants have non-negligible fitness effects and that the input of new mutations is counteracted by its removal through directional selection and thus maintained at 'mutation-selection balance'. Removal of genetic variation can

occur via the fixation of new advantageous mutations (positive selection (Maynard Smith & Haigh 1974)) or the removal of new deleterious mutations (purifying selection (Charlesworth et al. 1993)). Supporting this model, genome-wide patterns of nucleotide diversity in *D. melanogaster* are consistent with the joint action of positive and purifying selection (Elyashiv et al. 2016). In general, directional selection predicts that polymorphisms will be maintained at lower frequencies than under neutrality, with most variation made up of low-frequency deleterious alleles. Finally, a third model holds that genome-wide variation is selectively maintained under 'balancing selection' (Dobzhansky 1955). Here, selection actively maintains variants at intermediate frequencies, as opposed to the low and very low frequencies expected under neutrality and directional selection respectively. For example, the sickle-cell polymorphism in humans is maintained by balancing selection because heterozygotes are fitter than either homozygote in regions where malaria is endemic (Haldane 1949).

Although ample empirical evidence exists for each of the above mechanisms, their relative contributions to the maintenance of population genomic variation remain unclear. It is therefore essential that genomic studies clarify modes of selection operating on individual polymorphisms. With this in mind, the study of sexually antagonistic polymorphisms is highly valuable, as sexual antagonism often generates balancing selection at antagonistic loci, owing to the 'tug-of-war' between alternative alleles in each sex (Kidwell et al. 1977; Patten & Haig 2009; Connallon & Clark 2012). Antagonistic polymorphisms could therefore substantially contribute to genome-wide population genetic variation (Connallon & Clark 2014a;

Connallon & Clark 2014b). Yet despite the hypothesised relationship between sexually antagonistic selection and population genetic variation, and thus to the broader question of the evolutionary forces maintaining population genetic variation, no empirical data has so far been brought to bear on this question.

A related knowledge gap is the timescale over which sexually antagonistic selection maintains genetic variation. On the one hand, one might expect antagonistic polymorphisms to be maintained over long time periods. Theory predicts that antagonistic polymorphisms will be stabilised by slow environmental fluctuations (Connallon & Hall 2016), and fluctuations of this type are widely observed in nature, for example due to seasonal changes (Bergland et al. 2014). Furthermore, the persistence of antagonistic polymorphisms could be prolonged if evolving sex-specific expression and thus 'resolving' sexual antagonism is difficult, which could occur if the antagonistic polymorphism is located in a gene or gene region that cannot easily acquire sex-specific regulation (Stewart et al. 2010). On the other hand, antagonistic polymorphisms might be expected to be short-lived. Antagonistic polymorphisms are highly sensitive to genetic drift (Connallon & Clark 2012; Mullon et al. 2012; Hesketh et al. 2013) because the effectiveness of antagonistic selection depends on the square of sex-specific selection coefficients, $N_e(s_m^2+s_f^2)$—as such, very large selection coefficients are required overcome drift (Connallon & Clark 2012) and this effect could precipitate the fixation of alternative alleles at antagonistic loci. Alternatively, rapid environmental changes produce maladaptation that tends to align the direction of selection in the two sexes and favours the evolution of formerly

antagonistic polymorphisms under directional selection, often leading to fixation of the genetic variant that provides the greatest overall benefit across males and females (Whitlock & Agrawal 2009; Connallon & Hall 2016).

Assessing whether antagonistic polymorphisms are selectively maintained—and if so, for how long—requires two types of information: (i) data on the putative sequence targets of sexually antagonistic selection, and (ii) data on genome-wide polymorphism from independent populations which can be used to measure signatures of balancing selection (Fijarczyk & Babik 2015). Although population genomic polymorphism data is now becoming increasingly available thanks to population-wide whole-genome sequencing efforts (The 1000 Genomes Project Consortium 2010; Lack et al. 2016), data on the genetic basis of sexual antagonism remains scant.

1.4.1.2. Quantitative genetic fitness variation

Another knowledge gap concerns the nature of the evolutionary forces that maintain heritable fitness variation in natural populations (Radwan 2008). Although this question is closely linked to the causes of the maintenance of population genetic variation, it is valuable to separate the two. One reason is that the neutral theory, which is a leading hypothesis in attempting to explain observed levels of population genetic variation, is not relevant to the question of what maintains heritable fitness variation, since neutral variation—by definition—does not affect fitness. Another reason is that the methodologies and literature on the maintenance of quantitative fitness variation and genome-wide population genetic variation are often rather disparate (Charlesworth 2015).

Traditionally, fitness-relevant traits are expected to carry little heritable fitness variation. This follows from Fisher's fundamental theorem of natural selection, which states that the rate of increase in fitness is equal to its additive genetic variance (Fisher 1930). In other words, strong directional selection on fitness will quickly exhaust any heritable variation that is present in a population (Kimura 1958). Contrary to this null expectation, however, data from a number of studies in *Drosophila* (Roff & Mousseau 1987; Fowler et al. 1997; Long et al. 2009) and other species (Mousseau & Roff 1987; Messina 1993) indicates that there is substantial heritable genetic variation for fitness-relevant traits (Price & Schluter 1991; Houle 1992; Pomiankowski & Moller 1995; Merilä & Sheldon 1999). In the context of sexually selected traits, this apparent mismatch has been termed the 'lek paradox' (Borgia 1979; Kirkpatrick & Ryan 1991) and it remains largely unresolved.

To explain why fitness variation persists, several competing theories have been proposed. First, the 'genic capture' model proposes that fitness variation could be unusually polygenic and form a large enough mutational target that sufficient levels of variation can be maintained despite strong directional selection (Rowe & Houle 1996; Houle et al. 1996). Under this hypothesis, most fitness-relevant genetic variation is maintained at equilibrium between input of variation through mutation and removal through directional selection ('mutation-selection balance').

A second and broad class of hypotheses proposes that fitness variation is maintained through 'antagonistic pleiotropy'—that is, opposing selection between environments (Hedrick 2006), traits (Rose 1982; Curtsinger et al. 1994) or sexes (*i.e.*, sexual antagonism (Bonduriansky & Chenoweth 2009;

Van Doorn 2009)). Here, variation is maintained because alleles are beneficial in one environment but not another, yet migration between environments or temporal fluctuation in conditions prevents the loss of either allele. Antagonistic pleiotropy between traits similarly maintains variation because selection cannot fix variants that increase the fitness value of one trait without negative side effects on the fitness value of other traits.

One way that empirical research has evaluated the importance of competing theories has been to estimate the genetic correlation for fitness between males and females. If fitness variation is primarily determined by mutation-selection balance, as expected under the genic capture model, then genetic variants will tend to have sexually concordant fitness effects in both sexes, resulting in a positive r_{mf}^{W}. Similarly, if opposing selection between environments or traits is a primary driver of fitness variation, a positive r_{mf}^{W} would again be predicted. However, if opposing selection between the sexes (sexual antagonism) is dominant, this will tend to generate a negative r_{mf}^{W} (Bonduriansky & Chenoweth 2009). As discussed in section '1.2. Phenotypic evidence for sexual antagonism', some studies have detected a significant negative r_{mf}^{W} (Chippindale et al. 2001; Brommer et al. 2007; Delcourt et al. 2009). This provides some evidence in favour of pervasive sexual antagonism across the genome, although some estimates are derived from small samples of genotypes (e.g. N=40, Chippindale et al. (2001)) and—perhaps as a result—are not always consistent between studies (Qvarnström et al. 2006; Collet et al. 2016). More precise estimates would help clarify the nature of genetic variation for fitness.

Further clues about the forces maintaining fitness variation can be obtained by mapping the genetic basis of sex-specific fitness variation (see also '1.4.2. The genetic architecture of sexual antagonism'). For example, some theories for the maintenance of fitness variation predict an enrichment of X-linked fitness variation (Rice 1984), whereas others do not (Rowe & Houle 1996). The distribution of fitness-associated variants across the genome could therefore provide important clues as to the nature of the evolutionary forces maintaining fitness variation (Fitzpatrick 2004). Alternatively, the degree to which genetic variation is polygenic or monogenic can inform the mechanisms that maintain fitness variation, since the genic capture hypothesis relies on high polygenicity of fitness traits (Rowe & Houle 1996), whereas mechanisms of balancing selection can potentially favour fewer alleles of larger effect (Johnston et al. 2011; Barson et al. 2015). Yet without data on the genome-wide identity, distribution and functions of polymorphisms underlying sex-specific fitness, it is difficult to evaluate the strength of the evidence in support of any given mechanism for the maintenance of heritable fitness variation.

1.4.2. The genetic architecture of sexual antagonism

Uncovering the genetic basis of fitness variation has been a major goal of modern biology (Pardo-Diaz et al. 2015; Wilkinson et al. 2015), with important implications for our understanding of broader evolutionary questions (see '1.4.1.2. Quantitative genetic fitness variation'). Given the limitations of genetic data on sexual antagonism to date (see '1.3. Genetic evidence for sexually antagonistic polymorphisms'), a key research priority is

to gain a better understanding of the genetic architecture of sexual antagonism. I therefore outline a number of interesting and potentially testable predictions that have been raised in recent years with reference to sexually antagonistic loci.

1.4.2.1. The identity and location of sexually antagonistic loci

One simple question is how many independent sexually antagonistic loci exist and what are the effect sizes of their alleles? Variation for fitness is likely to depend on many traits, most of which are expected to be polygenic (Falconer & Mackay 1996). Thus, one might hypothesise that antagonistic genetic variation will consist of many loci of small effect, rather than few loci of large effect. However, this should not necessarily be taken as a given. Polymorphisms with small, balanced fitness effects in the two sexes are highly sensitive to drift, while alleles with large effects are the more likely to be maintained under balancing selection (Connallon & Clark 2014a; Connallon & Clark 2014b). Given these facts, one could instead observe a strong contribution to sexually antagonistic fitness variance from a small number of large-effect alleles (Connallon & Clark 2014a). Illustrating this, the aforementioned *VGLL3* locus (Roberts et al. 2009) explains ~40% of the variance in the sexually antagonistic 'age at maturation' trait in salmon.

Another question of interest is the distribution of sexually antagonistic polymorphisms along the genome. In an influential paper, Rice (1984) hypothesised that the X chromosome should disproportionately accumulate sexually antagonistic polymorphisms because: (i) hemizygosity of the X chromosome unmasks recessive alleles that are beneficial to the

heterogametic sex, and (ii) disproportionate transmission of the X chromosome through the homogametic sex favours dominant alleles that benefit it. In other words, any antagonistic polymorphism where the male-beneficial allele is recessive (and, by implication, the female-beneficial allele is dominant) is strongly favoured on the X chromosome.

Key evidence to support X-linkage of antagonistic polymorphisms comes from *D. melanogaster*, in which it has been shown that 97% of sexually antagonistic fitness variation is located on the X chromosome (Gibson et al., 2002)—although this value is estimated with considerable imprecision due to the small sample of fly lines used in the relevant study (N=20). More recently, Innocenti & Morrow (2010) mapped antagonistic expression patterns across the *D. melanogaster* genome, showing that the X chromosome is significantly over-represented among genes with antagonistic expression. Similarly, Lucotte et al. (2016) found an enrichment of antagonistic polymorphisms on the X chromosome in humans. In contrast to these results, Hill (2017) detected an over-representation of antagonistic polymorphisms on the autosomes in *D. melanogaster*.

More exhaustive studies are required to address this theoretical prediction. This is especially true in light of recent theory, which has challenged the straightforward predictions from Rice's (1984) model. For instance, Fry (2010) suggested that if allelic dominance often differs between the sexes, this would favour the accumulation of antagonistic polymorphisms on autosomes. Supporting this contention, the autosomal antagonistic *VGLL3* locus was shown to exhibit sex-specific dominance (Barson et al. 2015). Another factor is genetic drift, which disproportionately depletes X-

linked genetic variation due to the X chromosome's smaller effective population size (Caballero 1995). This in turn might prove detrimental to the persistence of X-linked antagonistic polymorphisms (Connallon & Clark 2012; Mullon et al. 2012). Finally, if antagonistic polymorphisms evolve under net directional selection, rather than balancing selection, the fact that recessive X-linked alleles are more visible to selection than they are on autosomes tends to favour stronger directional selection on the X (Charlesworth et al. 1987). The disproportionate depletion of X-linked variation under directional selection will then tend to favour higher mutation-selection balance equilibria of antagonistic polymorphisms on autosomes (Connallon & Clark 2012).

1.4.2.2. The biological properties of sexually antagonistic loci

Another open question concerns the biological properties of antagonistic variants. In general, sexually antagonistic polymorphisms are expected to be situated in gene regions where the evolution of sex-specific gene expression—and thus the resolution of sexual antagonism—is most difficult (Stewart et al. 2010). For example, epigenetic mechanisms such as genomic imprinting could mitigate sexual antagonism by silencing the expression of the maternally derived allele in males and the paternally derived allele in females (Day & Bonduriansky 2004), under the assumption that maternally and paternally derived alleles tend, on average, to have female- and male-beneficial effects, respectively. Alternatively, researchers have suggested that the resolution of sexual antagonism might be more difficult in protein coding than regulatory gene regions (Ellegren & Parsch 2007; Connallon &

Clark 2011b), and thus that antagonistic polymorphisms could be found more often in coding regions. This expectation arises because the resolution of a protein-coding antagonistic polymorphism is a relatively complicated process, requiring gene duplication followed by the acquisition of new sex-specific regulatory elements in each paralog. By contrast, the resolution of a regulatory antagonistic polymorphism can occur through minor modifications of existing regulatory elements, or a duplication event followed by such minor modifications—processes which are known to occur commonly (Ellegren & Parsch 2007; Williams & Carroll 2009).

Additionally, researchers have suggested that pleiotropic genes could disproportionately harbour antagonistic variation because the evolution of sex-biased or sex-limited gene expression is more difficult to achieve in such genes, owing to deleterious side-effects on the organism's fitness (Mank et al. 2008). Supporting this, Innocenti & Morrow (2010) reported that antagonistically expressed genes were widely expressed across tissues, and accordingly suggested that there was a positive relationship between the degree of gene pleiotropy and the probability than a gene harbours antagonistic polymorphisms. Beyond this, however, no firm predictions and no clear picture has emerged regarding the functions and pathways of genes harbouring antagonistic polymorphisms.

1.5. Thesis overview

1.5.1. Broad aims and methodology

In light of the knowledge gaps outlined above, the main aims of this thesis are: (i) to describe the genetic basis of sexual antagonism; (ii) to characterise

the contribution of sexual antagonism to quantitative genetic variation for fitness, and (iii) to assess the impact of sexually antagonistic selection on population genetic variation across space and time.

I use the model organism *Drosophila melanogaster*—the fruit fly—to address these questions. A number of factors inform this choice. First, fruit flies have a fully sequenced and well annotated genome (Dos Santos et al. 2015), which facilitates functional analyses. Second, experimental measurements of fitness are straightforward, given the short generation time (~14 days) and ease of manipulation of the organism. Third, the availability of a laboratory-adapted population of fruit flies—LH_M, maintained under strictly controlled conditions for ~20 years (Rice et al. 2005)—permits meaningful measurements of fitness in both sexes. Fourth, the hemiclonal design (Abbott & Morrow 2011) allows fly lines carrying an identical half-genome to be expressed as either a male or a female. This in turn permits genome-wide associations with sex-specific fitness to be conducted. Finally, there is now extensive population genomic data from worldwide populations of *D. melanogaster* (Lack et al. 2016) and sister species (Rogers et al. 2014), which provides an important platform for comparative analyses linking sexual antagonism with signatures of selection.

In the following two sections, I provide more background information on (i) the laboratory-adapted LH_M population and (ii) hemiclonal analysis, since LH_M hemiclones are the foundation and an integral component of the analyses presented throughout this thesis.

1.5.2. LH_M population

LH_M is a laboratory-adapted population of *D. melanogaster* flies. It was established in 1991 from 400 inseminated females sampled in an orchard near Modesto, California (Rice et al. 2005). Since 1996, it has been maintained at a large effective population size (~1,800 breeding adults) under a strictly controlled rearing regime. The rearing regime involves three phases: 'adult competition', 'adult oviposition' and 'juvenile competition'. In the 'adult competition' phase, newly eclosed flies (16 males and 16 females in each of 56 vials) interact for 48 hours, during which time females compete for a limiting amount (~5mg) of live yeast and males compete for matings. The subsequent 'adult oviposition' phase takes place after a transfer to fresh, unyeasted vials, where the females among the 32 adults oviposit for 18 hours to establish the next generation of the population. After the 18-hour egg-laying period, adult flies are discarded and eggs manually culled to a density of 150-200 per vial. Eggs hatch and larvae develop during the following 'juvenile competition' phase which lasts for 12 days until the emergence of a new generation of flies. Emergees are then mixed and culled to 16 males and 16 females per vial for the next round of the rearing cycle. The fact that vials are mixed at the beginning of the adult competition phase means that the population is essentially panmictic.

1.5.3. Hemiclonal analysis

Hemiclones are replicate individuals that carry an identical half-genome (chromosome X-2-3 haplotype) which is paired with a random X-2-3 or a Y-2-3 complement (females and males respectively). To generate a single hemiclonal line (see Fig. 1.2), an individual male from the base population

(LH_M) is crossed to 'clone-generator' (CG) females to create a 'target genome' (TG) male, which carries a specific X-2-3 haplotype from LH_M. By repeatedly crossing a TG male with CG females, it is possible to indefinitely maintain the initial X-2-3 LH_M haplotype because: (i) no recombination occurs during meiosis in *D. melanogaster* males, and (ii) CG females carry a Y chromosome and marked chromosomal translocations that co-segregate and can be phenotypically distinguished.

To measure the fitness of a given hemiclone, TG males are crossed with either (i) virgin LH_M females, or (ii) virgin LH_M females carrying a fused-X chromosome. These crosses produce 'hemiclonal females' and 'hemiclonal males', respectively, that carry an identical hemiclonal X-2-3 chromosome set in a background of random wild-type chromosomes. The hemiclonal design is effectively analogous to fertilising a set of clonal eggs with a random sample of sperm (Pennell & Morrow 2013). It is described in more detail elsewhere (Rice et al. 2005; Abbott & Morrow 2011).

In short, the hemiclonal design allows (i) the random sampling of genetic variation from the base LH_M population, (ii) the expression of the hemiclonal haplotype in either males or females, and (iii) the measurement of phenotypes from multiple replicate individuals for a given line.

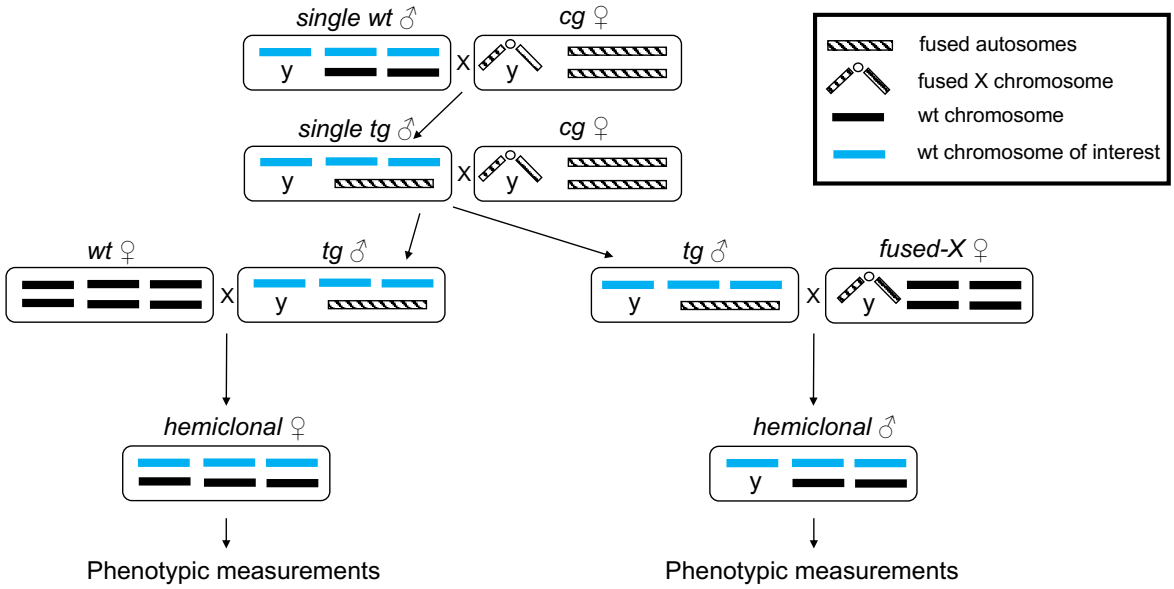


Figure 1.2. The hemiclonal amplification process. The above diagram represents the hemiclonal amplification process for a single line; multiple such lines can be created by copying this process but using a different wild-type male to initiate the first wt x cg cross. If different males from a base population (such as LH_M) are randomly sampled, each will carry a different 'wt chromosome of interest', and the set of hemiclonal lines thus created will constitute a random sample of the genetic diversity in the base population. This diagram is modified from Rice et al. (2005). wt=wild-type; tg=target genome; cg=clone-generator.

1.5.4. Thesis structure

This thesis comprises four data chapters, a discussion, and three appendices. Here I briefly describe the contents of each section.

In Chapter 2, I illustrate the limitations of current genetic studies of sexual antagonism. To do so, I present a re-analysis of two prominent genetic studies of sexual antagonism in *D. melanogaster*, that of Innocenti & Morrow (2010) and that of Hill (2017). Both studies used a nominally random

sample of LH_M hemiclones to make inferences about the genetic basis of sexual antagonism. I re-assess this assumption and use phenotypic data to show that this sample is not random. I then consider how this non-random sampling could have occurred, and I further consider whether and how this non-random sampling affects phenotypic and genetic inferences of sexually antagonistic selection made by both sets of researchers.

In Chapter 3, I use male and female fitness data from a large sample of hemiclonal fly lines from the laboratory-adapted LH_M population to present the first genome-wide association study (GWAS) of sex-specific fitness and sexual antagonism. I analyse the identity, location, and function of sex-specific, sexually concordant and sexually antagonistic polymorphisms. I consider how these data inform our understanding of the forces that maintain genetic variation for fitness. I also consider whether the patterns are in line with predictions about the location of sex-specific and sexually antagonistic variants. Finally, I consider whether the functional properties of antagonistic variants inform our understanding of the genetic factors that facilitate or hamper the evolution of sexual dimorphism.

In Chapter 4, I use information on the location of sexually antagonistic polymorphisms (derived from the GWAS data presented in Chapter 3) to consider a key question: does antagonistic selection maintain population genetic variation? To address this question, I combine data on the location of antagonism polymorphisms in LH_M with polymorphism data from publicly available genome sequences from worldwide *D. melanogaster* populations. I then test whether antagonistic loci in LH_M are associated with signatures of balancing selection across populations, which would indicate that

antagonistic selection maintains variation in these populations. As an extension of this analysis, I further assess whether antagonistic loci are associated with increased polymorphism in two sister species of *D. melanogaster*: *D. simulans* and *D. yakuba*.

In Chapter 5, I experimentally test whether a candidate polymorphic region of a highly pleiotropic and conserved gene—*fruitless*—is under balancing selection. I do so by tracking the frequency dynamics of cage populations of *D. melanogaster* initiated with extreme starting frequencies of each respective *fruitless* allele, before assessing whether allele frequencies converge to an intermediate state, as would be expected under balancing selection. I then proceed to test whether sexual antagonism is the mechanism of balancing selection by conducting sex-specific fitness assays among fly lines carrying each respective allele.

In Chapter 6, I first summarise the main findings in each chapter. I then place the findings about sexually antagonistic polymorphisms presented in this thesis in a more general context, by making links to research on the genetics and evolutionary dynamics of antagonistic pleiotropy more generally. I conclude by suggesting some promising directions for future research.

In Appendix A, I present a copy of a manuscript that is currently in revision for *Nature Ecology and Evolution*, on which I took the lead on data analysis and writing. It contains work documented in Chapters 3 and 4 and analyses jointly conducted with Mark Hill, a former PhD student in the Reuter laboratory. The manuscript describes the genetic architecture, functional

properties and evolutionary dynamics of sexually antagonistic polymorphisms.

1.5.5. Declaration

This PhD project was funded by a PhD studentship with the London NERC Doctoral Training Partnership and was performed under the supervision of Dr Max Reuter and Prof. Kevin Fowler. All work presented is mine, except where noted. Specific contributions are described in more detail at the beginning of each chapter.

Chapter 2

Cryptic population structure limits previous phenotypic and genetic inferences of sexually antagonistic selection

2.1. Declaration

All work reported in this chapter is my own, except for the following sections which were performed by collaborators and are included for context: ‘2.4.1. Sampling and phenotyping of 100 LH_M hemiclones’ (Paolo Innocenti, Edward Morrow) and ‘2.4.2. Genotyping of ‘extreme’ hemiclones’ (Paolo Innocenti, Edward Morrow, Mark Hill, Centre for Genomic Research, University of Liverpool). Analyses presented in ‘2.4.5. Re-analysis of quantitative genetic data’ build upon preparatory work (Max Reuter).

I thank Edward Morrow for sharing the sex-specific fitness data and for providing helpful information on the creation of hemiclonal lines. I thank Mark Hill for sharing the genomic data and helpful discussions.

2.2. Abstract

Quantitative genetic studies have shown that sexual antagonism is a common feature of phenotypic variation in the wild and in the laboratory, but the identification of sexually antagonistic loci has proved challenging. In a pioneering investigation, Innocenti & Morrow (2010) measured the sex-specific fitness of 100 *D. melanogaster* hemiclones from the LH_M laboratory population and reported a negative intersexual correlation for fitness (r_{mf}^W)—a hallmark of sexual antagonism. By further quantifying whole-genome gene expression among a subset of hemiclonal lines with 'extreme' sex-specific fitness effects, these authors identified ~1,300 candidate 'antagonistically expressed' genes and highlighted interesting functional associations. Nevertheless, their analysis implicitly assumed that the sample of hemiclones was a random sample from LH_M, and thus that any relationship between gene expression and fitness was causal. Here I show that the sample of 100 LH_M hemiclones used by Innocenti & Morrow (2010) is not random, and instead clusters into two distinct 'sets'. Genomic data from a sub-sample of lines from each set supports the hypothesis that both sets originate from different underlying populations. This cryptic population structure adversely affects downstream phenotypic and genotypic inferences. First, the significant negative r_{mf}^W reported by Innocenti & Morrow (2010) can no longer be supported once non-random sampling is taken into account. Second, population structure severely inflates the false positive rate among previously identified candidate loci. As such, the functional properties of these candidate genes can no longer be reliably interpreted in terms of sexually antagonistic selection.

2.3. Introduction

In the past couple of decades, quantitative genetic studies have shown that traits which are genetically correlated between the sexes often have opposing fitness optima in each sex—that is, traits are often sexually antagonistic (e.g. Long & Rice 2007; Svensson et al. 2009; Mokkonen et al. 2011; Delph et al. 2011; Tarka et al. 2014). For example, *D. melanogaster* locomotory activity and human height are subject to opposing selection pressures in each sex despite exhibiting strong positive intersexual genetic correlations (Long & Rice 2007; Stulp et al. 2012). In line with this, research has also shown that many populations harbour antagonistic fitness variation (Chippindale et al. 2001; Brommer et al. 2007; Foerster et al. 2007; Delcourt et al. 2009). Yet despite these important findings, such studies cannot shed light on the identity of antagonistic polymorphisms, their distribution along the genome, or the biological processes they typically affect. In short, they provide little understanding of the genetics of sexual antagonism.

A single study stands out in this regard (Innocenti & Morrow 2010), through its identification of genome-wide 'antagonistically expressed' genes. In this study, Innocenti & Morrow (2010) extracted a random sample of 100 hemiclonal lines from the laboratory-adapted LH_M population (see '1.5.2. LH_M population' and '1.5.3. Hemiclonal analysis') and measured the male and female fitness of each hemiclonal line. The authors first estimated the intersexual genetic correlation for fitness (r_{mf}^{W}), and reported a negative r_{mf}^{W} across the entire sample of 100 lines. A negative r_{mf}^{W} is a hallmark of extant sexual antagonism (Bonduriansky & Chenoweth 2009) and indicates that high fitness genotypes in males tend to confer low fitness when expressed in

females, and vice versa. The negative r_{mf}^{W} was accordingly interpreted as evidence for widespread extant sexual antagonism in LH_M.

The authors then took a subset of lines—5 with extreme female-beneficial male-detrimental (FB) fitness effects, 5 with extreme male-beneficial female-detrimental (MB) fitness effects, 5 with average sex-specific fitness effects—and used microarrays to measure the sex-specific, whole-body and whole-genome gene expression of individuals from each line. The authors then identified 'antagonistically expressed' genes by looking for genes where the relationship between expression level and fitness varied by sex, such that increased expression of a gene was beneficial in one sex but not the other; more formally, they looked for genes where there was a significant sex-by-fitness interaction term for expression. Doing so, the authors identified ~1,300 candidate antagonistically expressed genes across the sub-sample of 15 lines.

The authors proceeded to analyse the functional properties of the 1,300 candidate genes and highlighted some interesting characteristics. Most notably, they showed that candidate genes were: (i) less sex-biased than the genome-wide average; (ii) disproportionately X-linked; (iii) widely expressed across tissues. The lack of strong sex-biased expression among candidate genes was taken to reflect constraints on the evolution of sexually dimorphic gene expression imposed by antagonistic selection. Additionally, the enrichment of candidate genes on the X chromosome was taken to support classic theoretical predictions about preferential X-linkage of antagonistic genes (Rice 1984; Patten & Haig 2009). Finally, the wide expression of candidate genes across tissues suggested that pleiotropy is an important

factor in maintaining antagonistic genetic variation, as had been proposed by some researchers (Mank et al. 2008).

The results presented by Innocenti & Morrow (2010) constituted an importance advance in the field, and have since inspired further research. For instance, Griffin et al. (2013) found that the antagonistic genes identified by Innocenti & Morrow (2010) tend to have higher intersexual genetic correlations for gene expression (*i.e.*, expression r_{mf}), as measured in a closely related *D. melanogaster* population (Ayroles et al. 2009). This was interpreted as the shared genome constraining the evolution of sex-specific expression and thus generating a positive correlation with Innocenti & Morrow's (2010) measure of antagonistic selection. Cheng & Kirkpatrick (2016) further found that Innocenti & Morrow's (2010) antagonistic genes tend be overrepresented among genes with intermediate levels of sex-biased expression. They used this pattern to support a 'Twin Peaks' model of the relationship between antagonistic selection and sex-biased expression, where 'peaks' represent intermediate sex-biased expression and troughs represent low and high sex-biased expression respectively. Finally, the negative r_{mf}^{W} estimated from Innocenti & Morrow's (2010) sex-specific fitness data was contrasted with the absence of an analogous correlation in another population evolving in parallel. This contrast was interpreted as a rapid resolution of sexual antagonism in this latter population (Collet et al. 2016).

Yet despite its prominence, Innocenti & Morrow's (2010) analysis suffers from two potentially major shortcomings. First, their methodology relied on expression differences to map the genetic basis of sexual antagonism. This is problematic because an antagonistically expressed gene

need not itself harbour a causal antagonistic polymorphism. Instead, a causal polymorphism could be situated in a completely different gene, such as a regulator of the antagonistically expressed gene. One cannot therefore infer with certainty that an antagonistically expressed gene is the target of antagonistic selection. Second, their analysis assumed that the genotyped lines are random samples from the base population, and thus that any relationship between expression variation and fitness is causal. However, if lines from a particular fitness class also happen to be more related to each other, large numbers of the inferred 'sexually antagonistic' candidate genes will in fact be false positives. For example, two populations may evolve different sex-specific fitness due to the combined effect of neutral and phenotypically relevant changes in expression. However, unless population history is accounted for, neutral changes will be spuriously assumed to contribute to differences in fitness and show up as 'candidates'. Such confounding effects are typically accounted for in genome-wide association studies (Astle & Balding 2009; Price et al. 2010), but this is less often the case in gene expression studies—despite their potential to cause many of the same problems (Kryvokhyzha et al. 2016).

With the first shortcoming in mind—that genes with antagonistic expression are only correlates of the true underlying antagonistic polymorphisms—a follow-up study (Hill 2017) used whole-genome sequences from Innocenti & Morrow's (2010) extreme fitness lines (MB lines, N=5; FB lines, N=4) to identify putative causal antagonistic genes. This was done by comparing allelic variation among lines in each 'fitness class' (MB or FB) and searching for fixed genetic differences between the genomes of

each fitness class. A total of ~6,000 fixed single nucleotide polymorphisms (SNPs) (~2,000 genes) were detected across the extreme genomes, constituting candidates for the genetic basis of sexual antagonism. Supporting the antagonistic effects of these candidate loci, Hill (2017) showed through permutation-based approaches that candidate loci were associated with a modest false discovery rate (~25%). Furthermore, Hill (2017) found that these loci exhibited functional properties consistent with antagonistic selection. For instance, candidate SNPs were significantly clustered along the genome and had disproportionately regulatory effects. Candidate genes also tended to be less sex-biased than the genome-wide average and were disproportionately found among regulators of sexual differentiation. Finally, Hill's (2017) candidate genes overlapped significantly with those of Innocenti & Morrow (2010), implying that both studies captured a shared signal of antagonistic selection.

In this study, I re-examine the second shortcoming—that is, I assess whether the hemiclones used by Innocenti & Morrow (2010) (and subsequently by Hill (2017)) are random samples from the LH_M population. The broad aim of this work is to evaluate whether the candidate antagonistic loci identified by Innocenti & Morrow (2010) and Hill (2017) can be safely interpreted in terms of antagonistic selection. This investigation is motivated by two observations. First, the awareness that the sample of 100 hemiclones used by Innocenti & Morrow (2010) was extracted from LH_M in two sampling 'rounds' (Edward Morrow, pers. comm.). Second, the fact that all 5 MB genomes used by Hill (2017) belong to one sampling round, while 3 of the 4 FB genomes belong to another. The strong overlap between sampling round

and fitness class potentially suggests that sampling of hemiclones from LH_M has been non-random. Given that lines from each sampling round were combined and treated as a single homogeneous sample from LH_M in downstream phenotypic and genetic analyses in both studies, it is crucial to establish that sampling was in fact random. If not, the interpretation of results from downstream analyses—and, by implication, our understanding of the genetics of sexual antagonism in general—may be compromised.

I first re-analyse sex-specific fitness data from Innocenti & Morrow's (2010) full sample of 100 hemiclones in light of possible differentiation between lines from each sampling round ('sets'). In doing so, I find evidence that the two sets ('H-lines' and 'P-lines') are significantly phenotypically differentiated. I then proceed to examine the causes of differentiation between sets by comparing their whole-genome sequences, using the subset of lines that had been sequenced (Hill 2017). I next evaluate whether population structure between sets affects the quantitative genetic inferences made by Innocenti & Morrow (2010). Finally, I consider whether differentiation between sets affects the false discovery rate among the ~6,000 candidate loci identified by Hill (2017). Overall, I evaluate whether our current understanding of sexually antagonistic genetic variation in LH_M is reliable.

2.4. Methods

2.4.1. Sampling and phenotyping of 100 LH_M hemiclones

Innocenti & Morrow (2010) extracted 100 hemiclonal lines from the outbred, laboratory-adapted LH_M population in the spring/summer of 2007. For a detailed description of the hemiclonal extraction process and the LH_M population, see sections '1.5.2. LH_M population' and '1.5.3. Hemiclonal analysis'. The sampling was performed in two rounds, with an initial 21 hemiclonal lines sampled in May 2007 and an additional 79 sampled between May and September 2007 (Edward Morrow, pers. comm.). The initial 21 lines were labelled with the prefix 'H' ('H1', 'H2', etc.), while the 79 additional lines were labelled with the prefix 'P' ('P1', 'P2', etc.). I henceforth refer to the two sample groups of H- and P-lines as 'sets'. Given the small number of generations between each sampling event, Innocenti & Morrow (2010) combined fitness data from H- and P-lines and treated the combined data as a single, random and representative sample of hemiclones from LH_M.

Fitness assays across the full sample of 100 lines were conducted in the autumn of 2007 (Edward Morrow, pers. comm.). Male fitness was measured as competitive fertilisation success and female fitness as competitive fecundity. These fitness assays were run to closely match the rearing conditions experienced by males and females in the base population. They therefore capture adult lifetime reproductive fitness in both sexes. More details about fitness assays can be found in the original study.

For clarity, Figure 2.1 summarises the phenotyping and genotyping procedures performed among the full sample of 100 hemiclones.

2.4.2. Genotyping of 'extreme' hemiclones

Innocenti & Morrow (2010) used a subset (N=15) of the full sample of 100 lines to measure gene expression. 5 of these lines had extreme male-beneficial, female detrimental (MB) fitness effects, 5 had extreme female-beneficial, male detrimental (FB) fitness effects and 5 had intermediate fitness effects. In a follow-up study, Hill (2017) whole-genome sequenced a subset (N=9) of these 15 extreme lines consisting of all 5 extreme MB lines ('H9', 'H10', 'H12', 'H13', 'H14') and 4 of the 5 extreme FB lines ('H7', 'P7', 'P9', 'P22'). One FB line was lost prior to sequencing and intermediate fitness lines were not sequenced.

Each hemiclonal genome was sequenced in three different genotypic constellations, in females where the hemiclonal genome was complemented with the genome of the strain used to generate the *D. melanogaster* reference sequence (Bloomington *Drosophila* Stock Center no. 2057), in females where the hemiclonal genome was complemented with chromosomes from the inbred line Canton-S, and in males where the hemiclonal genome was complemented with chromosomes from the 'clone generator' stock used for creating and maintaining hemiclonal lines. DNA was extracted and sequenced at high coverage (~30X) from flies of each of these genotypes. Comparisons across the three complements allowed for the hemiclonal half-genome to be assigned as reference/alternate with confidence. After filtering out lowly covered reads (call rate<10 in any of the libraries), a total of 1,052,882 high-quality SNPs were detected across the nine hemiclonal genomes. Further details on sample preparation, sequencing and the SNP calling pipeline can be found in Hill (2017).

2.4.3. Phenotypic characteristics of H- and P-lines

I compared sex-specific fitness measurements between H- and P-lines to test the assumption that both are random samples from a common LH_M fitness distribution. Sex-specific fitness measurements from the full sample of 100 lines (across both sets) were first Box-Cox transformed to be normally distributed, then scaled and centred. I then modelled sex-specific fitness variation as a function of set (H-line vs. P-line) and tested for a significant effect of set using Analysis of Variance (ANOVA).

2.4.4. Genetic characteristics of H- and P-line genomes

If phenotypic differences exist between sets, these differences should be detectable at the genetic level. I therefore considered whether H- and P-lines are genetically differentiated. I did so by examining kinship relationships among the subset of H-lines (N=6) and P-lines (N=3) that were whole-genome sequenced by Hill (2017).

First, I examined the phylogenetic topology among the nine genomes by constructing an unrooted neighbour-joining tree (using the estimation of Saitou and Nei (1987)) from whole-genome SNPs across the nine lines. Distance matrix estimation and tree construction was done using *ape* and *phyclust* packages (Paradis et al. 2004). Bootstrapping (100 times) was used to assess confidence in each branch node.

Second, I conducted an *ADMIXTURE* analysis (Alexander et al. 2009) while specifying the number of population clusters (K) as K=2. Under the hypothesis that lines from each set belong to two distinct genetic clades, this

analysis is expected to assign the sub-sample of H-line genomes to a genetic ‘H-clade’ and the sub-sample of P-line genomes to a genetic ‘P-clade’.

Third, I evaluated how ‘typical’ the allelic variation in a given genome was relative to its parent population, LH_M. For instance, if a given line is not randomly sampled from LH_M, this would manifest as an excess of SNP variants along that genome that are rare in LH_M. I therefore estimated, for each line in turn, the LH_M population frequency of the variants observed across the line’s genome. LH_M population allele frequencies were derived from 203 LH_M whole-genome sequenced hemiclones extracted in 2012 (Gilks et al. 2016).

Fourth, I placed genomic variation among the nine genomes within the broader context of North American *D. melanogaster* population genetic variation. I did so by performing a Principal Component Analysis (PCA) which included SNPs from the nine genomes, LH_M, and 205 whole genomes from a wild North American outgroup population (RAL) sequenced as part of the *Drosophila* Genetic Reference Panel (Mackay et al. 2012; Lack et al. 2015). As input for the PCA, I used high-quality overlapping positions (call rate=100% among the nine lines; call rate>95% in LH_M and RAL). These were further LD-pruned prior to PC estimation (no two SNPs with r^2>0.2 within 1Kb). LD-pruning and PC estimation was performed in *LDAK* (Speed et al. 2012).

Finally, I examined whether differentiation between each line and LH_M is genome-wide or localised to a small region of the genome. To do so, I calculated mean F_{ST} between (i) the H-lines and LH_M; (ii) the P-lines and

LH_M, in non-overlapping windows along the genome (100 windows per chromosome arm).

2.4.5. Re-analysis of quantitative genetic data

I considered whether Innocenti & Morrow's (2010) quantitative genetic inferences—heritability and r_{mf}^W estimates—were affected by potential differentiation between the two sets of hemiclonal samples. To do so, I standardised sex-specific fitness data for the H- and P-lines in turn; that is, I repeated the transformations of the fitness data described above (Box-Cox transformation, scaling, centring) separately within each set, in order to remove any difference between each line's position and spread along the male and female axes. I then combined the 'set-standardised' fitness measures, and from these combined data estimated sex-specific heritabilities (h_f^2, h_m^2) and r_{mf}^W.

To estimate h_f^2, h_m^2 and r_{mf}^W, I fitted a model with *MCMCglmm* (Hadfield 2010) such that $Y_{ijk} = X_{ij} + \varepsilon_{ijk}$, where Y_{ijk} is the fitness of individual k of genotype j and sex i, X_{ij} is the sex-specific random effect of genotype j in sex i, and ε_{ijk} describes the sex-, genotype- and individual-specific residual. X_{ij} follows a bivariate normal distribution $X_{ij} \sim N(0, G)$, where

$$G = \begin{pmatrix} \sigma_{G,f}^2 & cov_{G,fm} \\ cov_{G,fm} & \sigma_{G,m}^2 \end{pmatrix}$$

is the genetic variance-covariance matrix, with $\sigma_{G,f}^2$ and $\sigma_{G,m}^2$ being female- and male- specific genetic variances, and $cov_{G,fm}$ is the intersexual

genetic covariance. ε_{ijk} follows a normal distribution $\varepsilon_{ijk} \sim N(0, \sigma^2_{R,i})$, where $\sigma^2_{R,i}$ denotes the sex-specific residual variance specific to sex i.

Heritabilities for each sex i were estimated as $h^2_i = \frac{2\sigma^2_{G,i}}{\sigma^2_{G,i}+\sigma^2_{R,i}}$ where i denotes sex (the factor 2 is added because individuals share half their genetic material in the hemiclonal design (Rice et al. 2005)). The intersexual genetic correlation was estimated as $r^W_{mf} = \frac{cov_{G,fm}}{\sqrt{\sigma^2_{G,f}}\sqrt{\sigma^2_{G,m}}}$.

h^2_f, h^2_m and r^W_{mf} were calculated for each iteration of the Monte Carlo Markov chain (100,000 iterations, 25,000 burn-in, thinning parameter=50) and point estimates obtained by taking the mean across all iterations. 95% credible intervals were obtained using the function *HPDintervals*.

2.4.6. Re-analysis of the false discovery rate among candidate loci

Hill (2017) identified ~6,000 candidate antagonistic SNPs by looking for fixed differences between MB (N=5) and FB (N=4) lines. He then used a permutation test to estimate the false discovery rate. In this permutation test, fitness class labels (MB or FB) were permuted among the nine genomes and the number of fixed differences ('pseudo-candidate SNPs') recalculated on each permutation. The false discovery rate was then calculated as: mean number of pseudo-candidate SNPs across all permutations / number of candidate SNPs in the unpermuted data.

To calculate a false discovery rate that takes into account potential genomic differentiation between H- and P-lines, I permuted fitness class labels (MB or FB) as done by Hill (2017) but ensuring that labels were not permuted between sets. As P-line genomes only contain FB genotypes (see

‘2.4.2. Genotyping of ‘extreme’ hemiclones’), this is equivalent to only permuting labels among the 6 H-line genomes (which contain 5 MB and 1 FB genotype). In this way, any population structure been the two clades is maintained but the relationship between genotype and phenotype is partially uncoupled. I then calculated the false discovery rate in the same way as above using only the five permutations where clade structure is maintained.

2.4.7. Statistical software

All statistical analyses were performed in *RStudio* (RStudio Team 2015).

2.5. Results

2.5.1. Phenotypic differentiation between H- and P-lines

I re-analysed the sex-specific fitness data from the full sample of 100 hemiclones to investigate whether H-lines (N=21) and P-lines (N=79) are random samples from LH_M. I found that H-lines have significantly elevated male fitness relative to P-lines ($F_{1,98}$=14.10, P<0.001, Fig. 2.2A), while also exhibiting significantly reduced female fitness ($F_{1,98}$=6.54, P=0.012, Fig. 2.2A). As such, both sets cannot be seen as two independent random samples from the same underlying fitness distribution of the LH_M population.

2.5.2. Genetic differentiation among extreme H- and P-line genomes

I evaluated whether phenotypic differentiation between H- and P-lines is mirrored at the genetic level. I did so by considering genetic relationships among the nine genomes sequenced by Hill (2017), 6 of which belong to H-lines and 3 of which to P-lines. Examination of the neighbour-joining tree derived from the nine whole-genome sequences suggests that the nine genomes indeed cluster into a genetic H-clade and P-clade (Fig. 2.3A), with strong bootstrap support for each node. Furthermore, an *ADMIXTURE* analysis with K=2 supports the assignment of H- and P-line genomes into two clusters that correspond to a genetic H- and P-clade (Fig. 2.3B).

The distinction between the two clades was corroborated by comparing the LH_M allele frequencies found among each of the nine genomes (Fig. 2.3C). This showed that alleles segregating among P-clade genomes tend to be found at high frequency in LH_M; by contrast, there is a noticeable dearth

of common LH_M alleles among H-clade genomes. Additionally, relative to the P-clade, H-clade genomes harbour an excess of alleles that are rare within LH_M. The combined effects of tree structure and allele frequency bias can be visually illustrated using a PCA, consisting of the nine genomes, 203 LH_M genomes and 205 genomes from a wild North American outgroup population (RAL) (Fig. 2.3D). As expected, the P-clade genomes cluster within LH_M. However, owing to the excess of rare LH_M alleles and the dearth of common LH_M alleles found among H-clade genomes, these genomes clearly fall outside the distribution of LH_M in genotypic space.

I next considered whether between-clade differentiation is genome-wide or localised to a small region of the genome. Looking at F_{ST} between each respective clade and LH_M along the *D. melanogaster* genomes indicated that elevated differentiation between the H-clade and LH_M is observed throughout the genome (Fig. 2.4). It is not restricted to a small region of the genome, as might have been expected if the differentiation between clades had been caused by the presence of an inversion in the genomes of one clade but not the other.

2.5.3. Effect of population structure on previous phenotypic inferences

I investigated whether the phenotypic and genetic differentiation between H- and P-lines had a knock-on effect on the quantitative genetic results reported by Innocenti & Morrow (2010). Specifically, I tested whether estimates of h_f^2, h_m^2 and r_{mf}^W were impacted by differentiation between sets.

After set-standardising fitness measurements to account for differences in mean sex-specific fitness between sets, I found that estimates of female heritability (h_f^2=0.61, 95% CI 0.41;0.83) and male heritability (h_m^2=0.06, 95% CI 8.3x10^{-8};0.15) are broadly consistent with Innocenti & Morrow's (2010) estimates (h_f^2=0.63, h_m^2=0.12). The higher female than male fitness heritability is primarily driven by higher additive genetic variance for female fitness ($\sigma_{G,f}^2$=0.32, $\sigma_{G,m}^2$=0.03), rather than higher residual variance for male fitness ($\sigma_{R,f}^2$=0.70, $\sigma_{R,m}^2$=0.98).

While these heritability estimates are consistent with previous inferences, the significant negative r_{mf}^W reported by Innocenti & Morrow (2010) (r_{mf}^W=–0.52, 95% CI –0.87;–0.10) is no longer detectable when fitness data are standardised by set (r_{mf}^W=–0.37, 95% CI –0.92;0.26, Fig. 2.2B). This implies that the negative genetic correlation that was previously detected had been at least partly driven by the differences in average male and female fitness between the two sets.

2.5.4. Effect of population structure on previous genetic inferences

Previous genetic inferences of sexually antagonistic selection assume that the sample of extreme MB and FB lines comes from a homogeneous underlying population. However, analyses presented here (Fig. 2.3) indicate that all 5 MB lines belong to a genetic H-clade, while 3 of the 4 FB lines belong to a genetic P-clade. I therefore evaluated the consequences of clade

structure among the nine lines on the false discovery rate of antagonistic candidate loci inferred by Hill (2017).

In Hill's (2017) approach, candidate SNPs were identified by looking for fixed genetic differences between the 5 MB lines and the 4 FB lines. The false discovery rate was then determined by permuting the fitness class labels among the nine lines (N=125 permutations). Whereas the true phenotypic labels resulted in 6,275 fixed differences between FB and MB lines, a mean of ~1,627 fixed differences was found among the permuted fitness classes (false discovery rate=25.9%, Fig. 2.5). Given clade structure among MB and FB lines, a more realistic false discovery rate was obtained by maintaining clade structure when performing permutations (*i.e.*, by ignoring all permutations where H-clade and P-clade labels are swapped). Doing so across the five possible permutations where this condition is met, a mean of 5,333 fixed differences was detected (range=4,316-6,306), resulting in a false discovery rate of 85.0% (Fig. 2.5).

2.6. Discussion

In an influential study, Innocenti & Morrow (2010) used a sample of 100 LH_M hemiclones to examine the phenotypic and genetic properties of sexual antagonism. Here I re-examined their data and show that this sample of LH_M hemiclones is phenotypically and genetically differentiated into two 'sets'. The degree of differentiation between sets is substantial and casts doubt on previous findings. Notably, accounting for population structure no longer supports a negative r_{mf}^{W} in LH_M and also substantially inflates the false discovery rate among previously identified candidate antagonistic loci. I discuss these findings and consider the limitations of our understanding of sexually antagonistic genetic variation in LH_M.

When attempting to map the genetic basis of phenotypic variation, a key assumption is that associated loci should reflect a causal effect on the phenotype, not simply associations due to genome-wide close relatedness between individuals with similar phenotypes (Astle & Balding 2009; Price et al. 2010). Innocenti & Morrow (2010) made inferences about the quantitative genetic properties and genetic basis of sexual antagonism using a sample of 100 LH_M hemiclones. Because they sampled hemiclones from a well-mixed laboratory-adapted population, they considered that their sample was homogeneous with respect to ancestry. Accordingly, they did not correct for population structure in their sample.

Here, by re-examining their phenotypic data, I found evidence that the sample of 100 LH_M hemiclones clusters into two sets which correspond to two different 'rounds' of hemiclonal extraction from LH_M (Edward Morrow, pers. comm.). These two sets—H- and P-lines—are significantly

phenotypically differentiated; specifically, H-lines have higher male fitness and lower female fitness than P-lines. Using genomic data from a subset of nine H- and P-lines (H-lines, N=6; P-lines, N=3), I further showed that H- and P-lines cluster into distinct H- and P- 'clades'. In other words, differentiation at the phenotypic level is mirrored by differentiation at the genetic level. This genetic clustering is unlikely to have occurred by chance and suggests that genetic differentiation among the H- and P-lines detected among the nine genomes extends to the broader H- and P-sets of samples.

Additional analyses highlighted three further characteristics of genomic differentiation between the two clades. First, the degree of genomic differentiation is marked. Comparing genomic variation among H- and P-clade genomes to 203 genomes from their parent population LH_M reveals that H-clade genomes are sufficiently divergent from LH_M to form a distinct clade. Second, the differentiated status of the H-clade is driven by strong allele frequency bias relative to the P-clade and the LH_M population. Thus, H-lines harbour an excess of alleles that are rare in LH_M and a dearth of alleles that are common in LH_M. Finally, differentiation between H- and P-lines is not localised to specific regions of the genome; it is genome-wide in its extent.

These results raise an obvious question: why are H-lines, a nominally random sample from LH_M, highly genetically differentiated from P-lines (and LH_M in general)? The data presented here do not suggest an obvious answer. Nevertheless, one can at least evaluate the plausibility of different scenarios in light of the patterns observed.

A first possibility is that one set of lines harbours a structural variant (e.g. an inversion) that is not found among the other set. This hypothesis

would predict a localised peak of F_{ST} near the structural variant when comparing F_{ST} between the H-clade and LH_M. However, this prediction is inconsistent with the fact that F_{ST} levels are relatively uniform across the genome between the H-clade and LH_M.

A second possibility is that contamination of foreign genetic material occurred during the amplification of H-lines, but not P-lines. Since H-lines were extracted from LH_M a few weeks before P-lines, contamination of H-lines prior to the establishment of P-lines could plausibly result in genetic differentiation. This scenario is supported by the fact that H-clade genomes harbour an excess of alleles that are rare in LH_M; these rare alleles could correspond to foreign genetic material. However, this scenario predicts that genetic differentiation between the H-clade and LH_M will be heterogeneous, such that genomic regions containing foreign genetic material will show up as peaks of high F_{ST}, whereas genomic regions containing LH_M-derived variation will show up as troughs of low F_{ST}. Again, the relative homogeneity of elevated F_{ST} between the H-clade and LH_M runs counter to this hypothesis.

A third possible explanation for the observed differentiation is a bottleneck between the establishment of both sets of lines. This scenario is supported by several lines of evidence. First, a bottleneck predicts relatively homogeneous effects on genome-wide differentiation, as is observed. Second, a bottleneck predicts that P-lines should exhibit greater inbreeding depression, owing to increased homozygosity and the unmasking of recessive alleles with deleterious effects. In line with this, exploratory analyses have indicated that P-lines have stronger inbreeding depression for

survival than H-lines when expressed in a homozygous state (Max Reuter, pers. comm.). Third, a bottleneck scenario predicts that the H-clade will carry an excess of alleles that are rare in LH_M relative to the P-clade (these rare alleles would tend to be lost among the P-clade during the bottleneck). This condition is also met.

Yet despite these suggestive lines of evidence, a bottleneck scenario is difficult to reconcile with the observation that the H-clade carries a dearth of alleles that are currently frequent within LH_M. These frequent alleles should be retained during the bottleneck and observed at similar frequencies among H- and P-clades, yet they are rarer among the H-clade. Additionally, a bottleneck could be expected to increase male fitness because it will disproportionately purge low-frequency alleles situated on the X chromosome (Caballero 1995; Pool & Nielsen 2007). Since recessive X-linked alleles are often found at low frequencies and disproportionately expressed in hemizygous males (James 1973; Cowley & Atchley 1988; Wayne et al. 2007), the removal of low-frequency variants would tend to disproportionately benefit males. Yet in opposition to this prediction, P-lines have decreased male fitness. Thus, both observations make a bottleneck scenario unlikely.

A final scenario to consider is that H-lines have been erroneously sampled from a different, locally adapted population of *D. melanogaster* altogether, whereas P-lines have been correctly sampled from LH_M. Although there is no direct evidence for erroneous sampling—nor is it particularly likely in the context of a highly controlled laboratory environment—it is worth noting that the patterns observed cannot exclude this scenario. For example,

erroneous sampling could explain: (i) why the differentiation between the H-clade and LH_M is homogeneous across the genome, (ii) why the H-clade exhibits an excess of rare LH_M alleles, (iii) why the H-clade exhibits a dearth of common LH_M alleles, and (iv) why the two sets exhibit average differences in sex-specific fitness. Overall, a firm conclusion regarding the causes of genetic and phenotypic differentiation between Innocenti & Morrow's (2010) hemiclones remains unclear.

Taking into account phenotypic differentiation between sets does not markedly change Innocenti & Morrow's (2010) estimates of sex-specific fitness heritability. Although male fitness heritability is somewhat lower than reported previously (h_m^2=0.12, Innocenti & Morrow (2010)) and in other studies in LH_M (h_m^2=0.29 (Pischedda & Chippindale 2006); h_m^2=0.41 (Collet et al. 2016)), female fitness heritability is consistent with previous estimates (h_f^2=0.63, Innocenti & Morrow (2010)) and other studies in LH_M (h_f^2=0.57 (Long et al. 2009); h_f^2=0.53 (Pischedda & Chippindale 2006); h_f^2=0.58 (Collet et al. 2016)). Higher female than male heritability is primarily driven by the higher additive variance component in females, not higher residual variance component in males (Kruuk et al. 2000; Merilä & Sheldon 2000; McCleery et al. 2004). A simple explanation for these results is that stronger sexual selection in males than females (Bateman 1948) disproportionately depletes male additive fitness variation (Fisher 1930). This explanation is consistent with studies in other species (Merilä & Sheldon 2000; Pettay et al. 2005; Brommer et al. 2007; Foerster et al. 2007; Teplitsky et al. 2009), which have similarly found higher female than male additive genetic variances (but see McCleery et al. 2004; Kosova et al. 2010).

While estimates of heritability are not markedly affected by taking into account phenotypic differentiation between sample sets, the same is not true for the estimate of the intersexual genetic correlation for fitness. While Innocenti & Morrow (2010) inferred a significantly negative r_{mf}^{W} based on the uncorrected data, I showed that the revised r_{mf}^{W} point estimate, while still negative, is not significantly different from zero when sample set is taken into account. Since a negative r_{mf}^{W} represents strong evidence for extant sexual antagonism, this result indicates that the fitness effects of genome-wide genetic variants in LH_M are less antagonistic than previously estimated. Although the absence of a significant correlation for fitness between the sexes could plausibly indicate that antagonistic variation is entirely absent (e.g. if traits are able to affect each sex independently), this is very unlikely given that genetic trait correlations between the sexes are invariably high (Poissant et al. 2010; Griffin et al. 2013) and thus that the target size for mutations with sex-limited effects is restricted. Instead, a more plausible model to explain the absence of a significant positive or negative r_{mf}^{W} is that antagonistic and concordant variation jointly contribute to overall sex-specific fitness variation, and that their opposing effects on r_{mf}^{W} cancel each other out.

The absence of a significantly negative r_{mf}^{W} reported here fits a broader pattern of declining antagonism in LH_M. Whereas estimates reported in previous studies in LH_M (Chippindale et al. 2001; Gibson et al. 2002; Pischedda & Chippindale 2006) were consistently negative, recent research has tended to report less negative (and sometimes positive) r_{mf}^{W} estimates. For instance, Long et al. (2009) found a correlation of 0.015 between the

adult fitness of sons and their mothers in LH_M using a cross-generational parent-offspring design. Similarly, Collet et al. (2016) measured fitness of ~120 hemiclones in a replicate population of LH_M that has been evolving in parallel under identical conditions to those of the original population and also reported a positive r_{mf}^W estimate of 0.21. Finally, the estimate of r_{mf}^W derived from 202 LH_M genomes (Chapter 3) is also positive.

This trend towards a less negative r_{mf}^W is puzzling, because previous theoretical (Whitlock & Agrawal 2009) and empirical research (Long et al. 2012; Berger et al. 2014) tends to favour the opposing view—namely, that prolonged adaptation to stable environmental conditions exacerbates antagonism. This effect arises because alleles with sexually concordant effects that are maladaptive in the current environment are purged, leaving behind the genetic variation that is maintained due to its antagonistic effects (Whitlock & Agrawal 2009). Given the mismatch with this theoretical expectation, several possible explanations for this trend can be put forward. One possible explanation is that earlier estimates of r_{mf}^W were incorrectly estimated due to relatively small sample sizes (e.g. N=40 in Chippindale et al. (2001) and N=20 in Gibson et al. (2002)) or publication bias, and thus that r_{mf}^W has always been non-negative. However, evidence from larger studies where sex-limited evolution has been applied and antagonistic fitness effects detected seems to corroborate these earlier estimates (Rice 1998; Prasad et al. 2007; Bedhomme et al. 2008). A more plausible explanation for the observed trend is genetic drift. LH_M has been maintained at a modest effective population size of ~1,800 individuals for hundreds of generations. A gradual shift from negative to non-negative r_{mf}^W can be expected if drift has

strong effects on antagonistic variation, which both theoretical (Connallon & Clark 2012; Mullon et al. 2012) and empirical (Hesketh et al. 2013) research suggests it does. Similarly, if antagonistic variation disproportionately accumulates on the X chromosome, as predicted by some theory (Rice 1984; Patten & Haig 2009) and some experiments (Gibson et al. 2002), the disproportionate depletion of X-linked genetic variation under drift (Caballero 1995) could further exacerbate the depletion of antagonistic variation (Mullon et al. 2012). A final explanation for the transition from negative to non-negative r_{mf}^{W} is that genetic changes have occurred and brought the sexes closer to their sex-specific fitness optima during LH_M's laboratory maintenance, thus resolving sexual antagonism in LH_M and generating a non-negative r_{mf}^{W} (Collet et al. 2016). However, this explanation is generally at odds with evidence for long-term persistence of antagonistic polymorphisms presented in Chapter 4.

In addition to affecting the estimates of quantitative genetic parameters, population structure among the sample of extreme hemiclones inflates the false positive rate among candidate antagonistic SNPs that have been identified based on these lines. Accordingly, the false discovery rate estimated from a re-analysis of Hill's (2017) data is much higher when accounting for population structure than without. A re-analysis of Innocenti & Morrow's (2010) gene expression data, although not presented here, is likely to yield similar results. This is because patterns of gene expression, like patterns of nucleotide variation, are primarily governed by drift and purifying selection, and so tend to reflect the population history of the sample of individuals from which they originate (Khaitovich et al. 2004; Khaitovich et al.

2005; Whitehead & Crawford 2006; Storey et al. 2007). Failure to account for phylogenetic relationships between individuals will cause spurious associations between gene expression and phenotypic measurements, just as it does in genome-wide association studies (Astle & Balding 2009). An illustrative example of such inflation was recently described among populations of the shepherd's purse *Capsella bursa-pastoris*. In this species, previously detected signals of local adaptation were shown to be entirely attributable to phylogenetic relationships between individuals used in the study (Kryvokhyzha et al. 2016).

The high rate of false positives among candidate antagonistic SNPs and genes identified by Hill (2017) and Innocenti & Morrow (2010) hampers the interpretation of their functional properties in terms of sexually antagonistic selection. The non-random functional patterns reported in both studies may instead reflect the non-random properties of SNPs that are differentiated between H- and P-clades. For example, enrichment of candidate genes on the X chromosome (Innocenti & Morrow 2010) could be expected after divergence between H- and P-clades, either because drift is more pronounced on the X chromosome or because the rate of adaptive evolution on the X chromosome can often be faster than on autosomes (Charlesworth et al. 1987; Caballero 1995). Enrichment of candidate SNPs in regulatory regions, as found by Hill (2017), could similarly be expected if H- and P-lines are adapted to different local environments. Gene regulatory regions often disproportionately contribute to local adaptation between populations (Wittkopp & Kalay 2011; Fraser 2013), so an enrichment of candidate SNPs in such regions is unsurprising. For example, F_{ST}-outliers

between LH_M and a replicate of the same population evolving under identical conditions are disproportionately associated with regulatory functions, despite the close relatedness of these two populations and the similar environmental conditions they are exposed to (Collet et al. 2016). Finally, non-random functional enrichments between populations can arise if one of the two populations is subject to a bottleneck (Pavlidis et al. 2012), as might have occurred between H- and P-lines.

In summary, cryptic population structure casts doubt on previous phenotypic and genetic inferences of sexually antagonistic selection. Without the ability to distinguish population structure from antagonistic fitness effects, the functional interpretations of candidate loci presented in both studies should be treated with caution. More generally, this work emphasizes the importance of correcting for population structure in studies which aim to map the genetic basis of phenotypic variation. Finally, this study motivates future research to more finely resolve the genetic properties of sexually antagonistic SNPs and genes (see Chapter 3).

2.7. Figures

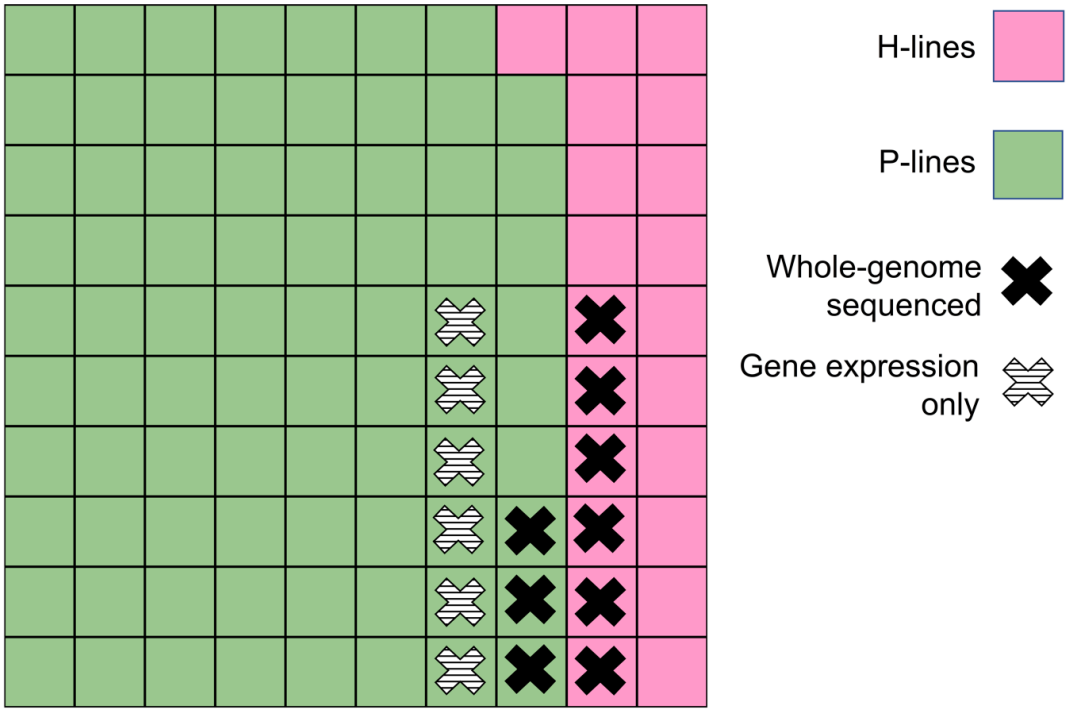


Figure 2.1. Summary of phenotyping and genotyping procedures among the full sample of hemiclones. Each square represents a hemiclonal line and colours represent the set (H- or P-line) to which each line belongs to. Black crosses denote lines for which both whole-genome sequence data (Hill 2017) and gene expression data (Innocenti & Morrow 2010) exists; striped crosses denote lines for which only gene expression data exists; no crosses indicate that only phenotypic data exists for that particular line.

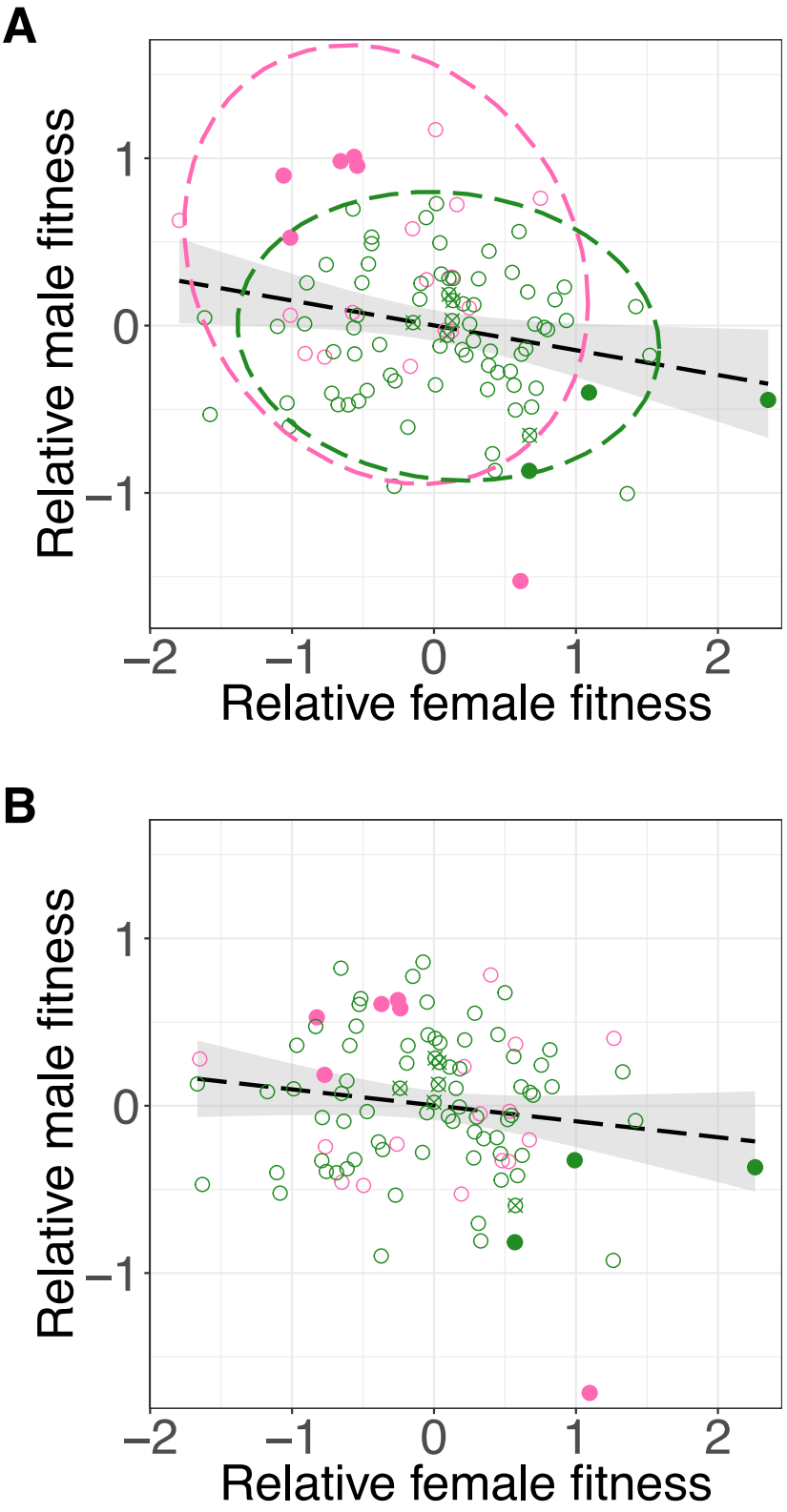
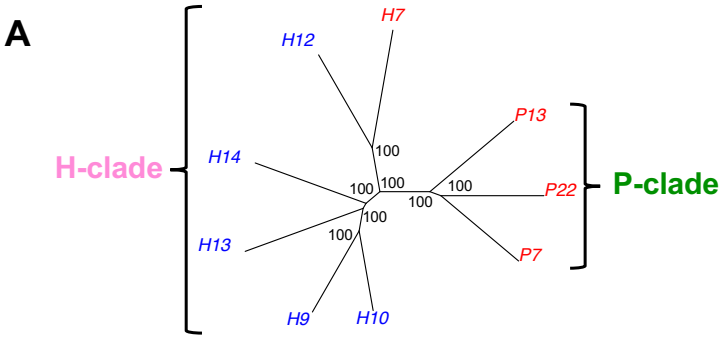
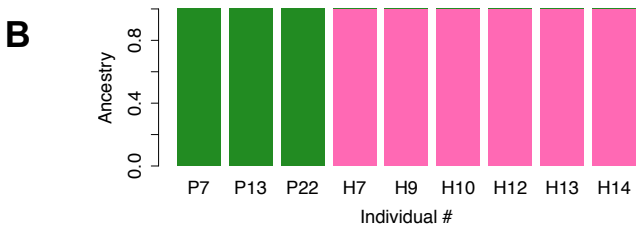
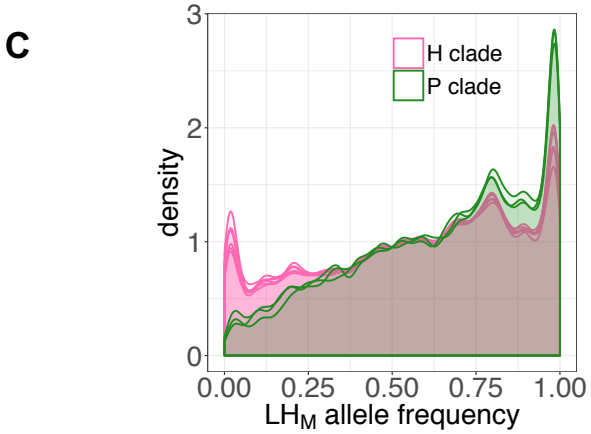

A
Relative male fitness
1
0
−1
−2
−1
0
1
2
Relative female fitness
B
Relative male fitness
1
0
−1
−2
−1
0
1
2
Relative female fitness

Figure 2.2. Sex-specific fitness among the full sample of hemiclones, before and after accounting for differentiation between sets. A. Dots represent (normalised, scaled and centred) male and female fitnesses for each of the 100 hemiclones used by Innocenti & Morrow (2010), split by line type (H-line=pink, P-line=green) with 95% confidence ellipses. Full circles denote extreme lines whole-genome sequenced by Hill (2017); open circles with crosses denote lines measured for gene expression by Innocenti & Morrow (2010) but not sequenced by Hill (2017); open circles denote lines for which no sequence data exists. To visually illustrate the significantly negative r_{mf}^{W} previously reported in this sample, a fitted regression line (with 95% confidence intervals) is presented. **B.** Re-analysis similar to that above for A., except that sex-specific fitnesses have been standardised within each set to account for population structure between H- and P-lines.


A
H-clade
P-clade
H12
H7
P13
H14
P22
H13
P7
H9
H10
100


B
Ancestry
0.0
0.4
0.8
P7
P13
P22
H7
H9
H10
H12
H13
H14
Individual #


C
H clade
P clade
density
0
1
2
3
0.00
0.25
0.50
0.75
1.00
LHM allele frequency

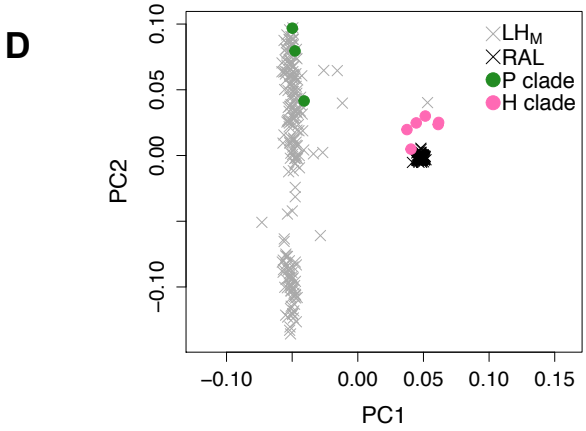

D
LHM
RAL
P clade
H clade
PC2
−0.10
0.00
0.05
0.10
−0.10
0.00
0.05
0.10
0.15
PC1

Figure 2.3. Genetic structure between H- and P-clades. A. Neighbour-joining tree constructed from whole-genome SNPs across the nine genomes sequenced by Hill (2017). Colours denote the fitness effects of each hemiclonal genome (blue=MB, red=FB). **B.** Individual ancestry proportions for each of the nine hemiclonal genomes estimated from *ADMIXTURE* analysis with K=2. **C.** LH_M population allele frequency of alleles found among each of the nine hemiclonal genomes (H-clade=pink, P-clade=green). **D.** PCA plot of the nine hemiclonal genomes (H-clade=pink, P-clade=green), 203 LH_M genomes (grey) and 205 genomes from RAL (black).

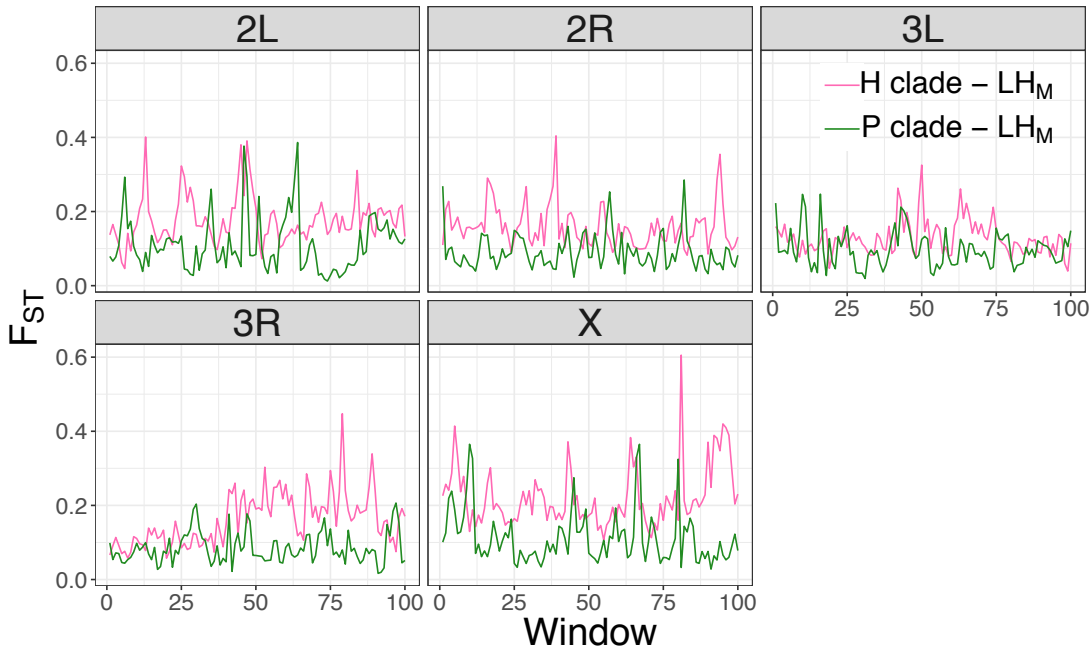


Figure 2.4. Genetic differentiation between clades and LH_M. Mean F_{ST} between clade (H/P) and LH_M, in 100 non-overlapping windows along the five major *D. melanogaster* chromosome arms.

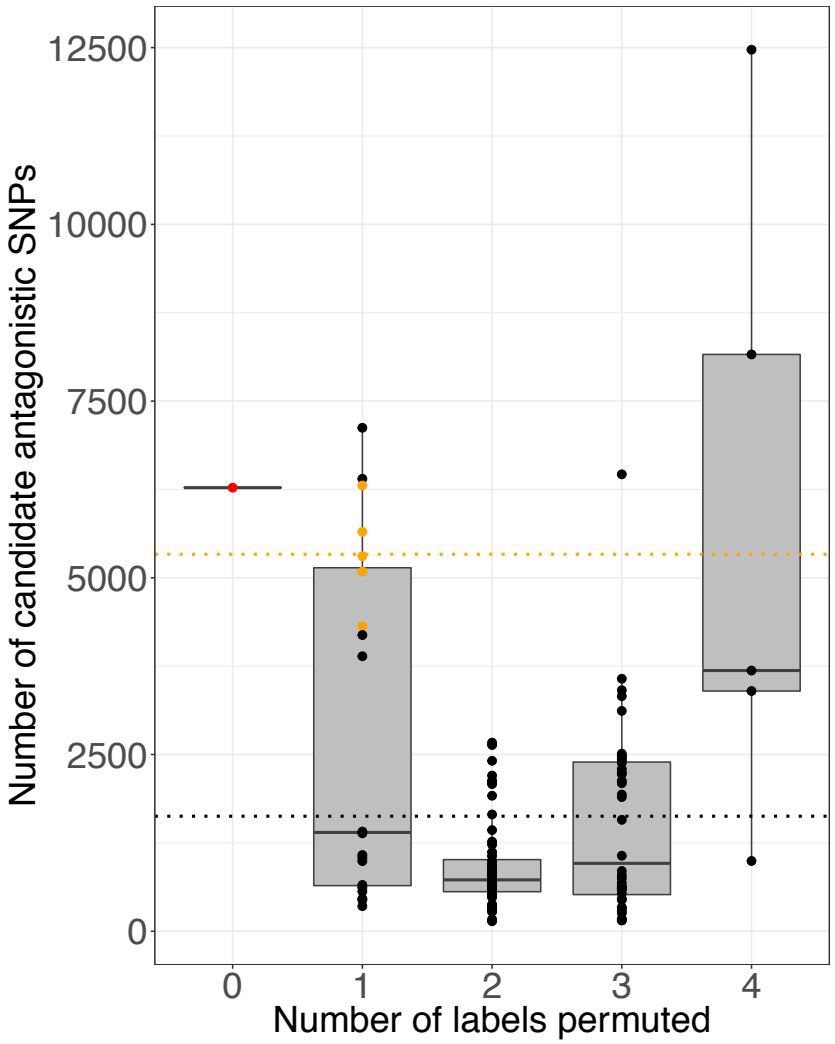


Figure 2.5. False discovery rate among previous candidate SNPs before and after accounting for population structure between clades. The number of candidate antagonistic SNPs, identified by Hill (2017) and defined as the number of fixed differences between the five MB and the four FB genomes (red dot), is shown. All 125 possible permutations of fitness class labels are presented as a function of the number of labels (MB/FB) that are swapped. Black dots denote the number of pseudo-candidates for each individual permutation (black dashed line denotes the mean number of pseudo-candidates across all permutations). Orange dots (and line) denote the number (and mean) of pseudo-candidates for the subset of five permutations in which labels are not swapped between clades, in order to maintain clade structure.

Chapter 3

Maintaining genetic variation for fitness: insights from a genome-wide association study of sex-specific fitness and sexual antagonism

3.1. Declaration

All work reported in this chapter is my own, except for the following sections which were performed by collaborators and are included for context: ‘3.4.1. Sex-specific fitness measurements of LH_M hemiclones’ (Tanya Pennell, Ilona Flis and Edward Morrow), ‘3.4.2. Sexually antagonistic and concordant phenotypes’ (Max Reuter) and ‘3.4.3. ‘Genotyping’ (Gilks et al. 2016). Some complementary analyses were performed by Mark Hill and are not presented here; they can be found in ‘Appendix A’.

I thank Edward Morrow for sharing the sex-specific fitness data which motivated this study, Doug Speed for helpful discussions about implementation of the *LDAK* software, and Tim Connallon for helpful discussions and comments.

3.2. Abstract

Fitness and fitness-relevant traits are under strong directional selection, yet multiple studies across taxa have reported ample heritable variation for fitness. The evolutionary forces maintaining it remain unclear. To clarify the causes of fitness variation, it is crucial to better understand the nature and identity of genome-wide genetic variation for fitness, yet few studies have done so. Here I combine whole-genome sequence data from 202 hemiclonal *D. melanogaster* fly lines with replicated male and female fitness measurements to conduct a genome-wide association study of sex-specific fitness. I show that sex-specific fitness is highly heritable but I find few confidently associated loci, implying that sex-specific fitness variation consists of many loci of small effect, as opposed to few loci of large effect. In line with previous work, I find that the X chromosome contributes disproportionately to male fitness variation. I decompose sex-specific fitness variation into 'sexually concordant' and 'sexually antagonistic' components and find that both components contribute equally to sex-specific fitness variation, resulting in no net fitness correlation between the sexes. However, contrary to predictions from classic theory, the genomic distribution of antagonistic variants is not significantly enriched on the X chromosome. Functional analyses also reveal that antagonistic polymorphisms are enriched in coding regions. This indicates that the evolution of sex-specific proteins is a limiting step in the resolution of sexual antagonism and thus that coding regions could play a crucial role in maintaining antagonistic fitness variation.

3.3. Introduction

Fitness and fitness-relevant traits are by definition under strong directional selection and should quickly deplete the genetic variation that fuels their evolution (Fisher 1930; Merilä & Sheldon 1999; Ellegren & Sheldon 2008). Yet contrary to this expectation, quantitative genetic studies have repeatedly estimated appreciable additive genetic variances for fitness and fitness-relevant traits (Mousseau & Roff 1987; Roff & Mousseau 1987; Houle 1992; Messina 1993; Pomiankowski & Moller 1995; Fowler et al. 1997). The puzzling mismatch between low expected and high observed fitness variances is often called the 'lek paradox' in reference to sexually selected traits (Borgia 1979; Kirkpatrick & Ryan 1991), but the mismatch applies more generally to other fitness-relevant traits (Charlesworth 2015).

Two broad classes of hypotheses have been put forward to solve this puzzle. One hypothesis is that fitness is influenced by many traits, which are together so polygenic that they 'capture' a large enough mutational target to explain observed levels of fitness variation. Under this 'genic capture' model (Rowe & Houle 1996), genetic variation for fitness primarily evolves under directional selection and is maintained by the balance between the generation of new variation by mutation and the removal of deleterious variants by directional selection ('mutation-selection balance'). The genic capture model also implies that alleles are beneficial or deleterious regardless of sex—in other words, that genetic variation tends to have either sex-limited or sexually concordant fitness effects (Connallon 2010).

A second class of hypotheses for the maintenance of fitness variation is antagonistic pleiotropy, which—in its broadest sense—encompasses

opposing selection between environments, traits or sexes (Radwan 2008). Opposing selection between environments (Levene 1953; Hedrick 2006), arises due to the migration of alleles between locally adapted patches where each respective allele is favoured. Opposing selection between traits arises because the beneficial effects of an allele on one trait are traded off against its detrimental effects on a different, correlated trait (Rose 1982; Curtsinger et al. 1994). Finally, opposing selection between the sexes (sexual antagonism) arises because the shared genome prevents each sex from evolving optimal sex-specific phenotypes (Bonduriansky & Chenoweth 2009; Van Doorn 2009). In all three cases, genetic variation is expected to primarily be maintained under balancing selection. Nevertheless, while sexual antagonism predicts that genetic variants will have sexually antagonistic fitness effects, opposing selection between environments or traits predicts that variants will tend to have sex-limited or sexually concordant fitness effects.

The relative importance of these various models toward the maintenance of fitness variation remains unclear (Connallon 2010), despite the fact that individual traits exhibiting patterns of selection consistent with each model have been described (Brooks 2000; Cotton et al. 2004; Bonduriansky & Chenoweth 2009). Going forward, it is essential that empirical studies characterise the relative contributions of each evolutionary force towards the maintenance of fitness variation (Radwan 2008; Charlesworth 2015).

Some progress towards this goal has been made using quantitative genetic approaches. For example, estimates of additive fitness variance for

viability estimated from diallel crossing designs in *D. melanogaster* reveal much higher variances than expected under mutation-selection balance alone (Kusakabe & Mukai 1984), consistent with an important role for balanced polymorphisms (Charlesworth 2015). Similarly, responses to artificial selection in outbred versus inbred lines suggest an important role for balanced polymorphisms (Kelly & Willis 2001; Charlesworth et al. 2007). Quantitative genetic studies have also provided insights into the forces maintaining genetic variation through the estimation of the intersexual genetic correlation for fitness (r_{mf}^{W}) (Rice & Chippindale 2001), which quantifies the extent to which genetic variants have sexually antagonistic (negative r_{mf}^{W}) or sexually concordant (positive r_{mf}^{W}) effects. Since a negative r_{mf}^{W} is only predicted when genetic variants have sexually antagonistic effects, and since variants with sexually antagonistic effects will often evolve under balancing selection (Connallon & Clark 2012), these experiments can provide suggestive evidence that antagonistic balancing selection plays a key role in the maintenance of fitness variation. The fact that such studies have yielded r_{mf}^{W} estimates that are rarely positive and often negative (Chippindale et al. 2001; Foerster et al. 2007; Delcourt et al. 2009; Brommer et al. 2007) implies that antagonistic balancing selection is indeed influential.

In addition to quantitative genetic approaches, genome-wide association studies (GWAS) can clarify the mechanisms that maintain genetic variation. GWAS take advantage of past recombination events throughout the genome to narrow down the genetic loci underlying a phenotype of interest (Bush & Moore 2012; Pardo-Diaz et al. 2015; Wilkinson et al. 2015). Specifically, a regression of the phenotype onto each allele is

performed, and loci with low statistical P-values are taken to be putatively causal for the phenotype under consideration (including fitness).

The GWAS approach can clarify the mechanisms that maintain genetic variation in a number of ways. First, the GWAS design allows the genetic architecture of fitness variants to be determined, which is informative because different models predict different genetic architectures of fitness variation. Under genic capture model, for instance, genetic variation for fitness is expected to consist of many alleles of small effect, whereas models of antagonistic pleiotropy generally predict fewer alleles of large effect (Connallon & Clark 2014b). Illustrating the insights that can be gained, a recent GWAS of a sexually selected trait (horn size) among Soay sheep (*Ovis ares*) uncovered a large effect polymorphism underpinning variation in this fitness-relevant trait (Johnston et al. 2011). Further investigation revealed that individuals with larger horns had increased reproductive success but suffered antagonistically pleiotropic viability costs (Johnston et al. 2013). Thus, this study provided supporting evidence in favour of the view that fitness variation is maintained by few balanced loci of large effect.

GWAS can also clarify the mechanisms maintaining fitness variation in a different way: by allowing the relative contributions of different chromosomal regions to fitness variation to be compared (Reinhold 1998; Fitzpatrick 2004). Because different modes of selection predict different patterns of fitness variation along the genome, uncovering the location of fitness variation can prove informative. For instance, the genic capture model and sexual antagonism generally predict different contributions of the X chromosome and autosomes to fitness variation. If directional selection

underpins most fitness variation and alleles are maintained at mutation-selection balance—as expected under the genic capture model—autosomes should carry more genetic variation for fitness (Haldane 1937; Hedrick and & Parker 1997). This expectation arises because the X chromosome is hemizygous in males, making recessive X-linked alleles more visible to selection whenever they are expressed in this sex (Charlesworth et al. 1987), and lowering the mutation-selection balance equilibrium frequency on this chromosome relative to autosomes. Conversely, if fitness variation is mostly maintained by sexually antagonistic selection, the X chromosome should be a hotspot for genetic fitness variation (Rice 1984; Patten & Haig 2009). Under sexual antagonism, an X-linked male-beneficial recessive allele is almost exclusively expressed in males (due to the hemizygosity of the X chromosome), while its dominant female-beneficial counterpart is disproportionately expressed in females (due to the preferential transmission of the X chromosome among females). Since each allele is preferentially expressed in the sex it confers benefits to, the invasion of these alleles is favoured on the X. Models predict that antagonistic polymorphisms will accumulate on the X chromosome across a wide range of dominance parameters when selection coefficients are small ($s<0.1$) (Patten & Haig 2009).

It should be noted that these predictions are not as clear-cut as made out above. For instance, various processes can shift the expectation for antagonistic polymorphisms towards autosomes, including the extent of sex-specific dominance, epistasis between loci, genetic drift, selection on multiple alleles and the extent of assortative mating (Fry 2010; Patten et al. 2010;

Úbeda et al. 2011; Arnqvist 2011; Mullon et al. 2012; Arnqvist et al. 2014). In addition, there may be a mismatch between the amount of variation expected at the sequence level, and the amount that is detectable in a fitness screen. For instance, even if the X chromosomes and autosomes harbour proportional amounts of sequence variation, there will be increased fitness variance associated with X-linked alleles in males because X-linked variation is not averaged across two chromosomes as it is in females (James 1973; Cowley & Atchley 1988; Connallon 2010; Reinhold & Engqvist 2013). Nevertheless, with these caveats in mind, an examination of the distribution of fitness variation between chromosomes can prove informative.

Finally, GWAS of fitness can also shed light on specific genetic features that facilitate the maintenance of genetic variation for fitness. For example, the probability of observing sexually antagonistic polymorphisms will in part depend on how easily sexual antagonism is 'resolved' (*i.e.*, how easily each allele is expressed in the sex it benefits). Some researchers have indicated that resolving antagonistic polymorphisms in coding regions will be difficult because this will require two unlikely events: a duplication event and the acquisition of a new sex-specific regulatory element that allows the sex-specific expression of each paralog (Ellegren & Parsch 2007; Stewart et al. 2010; Connallon & Clark 2011b). By contrast, if antagonistic polymorphisms are situated in regulatory regions, minor modifications of existing regulatory elements can resolve sexual antagonism with relative ease (Williams & Carroll 2009; Stewart et al. 2010). This theory predicts that antagonistic polymorphisms will tend to be situated in coding regions. Yet without

knowledge of the location of antagonistic polymorphisms, it is difficult to evaluate whether these predictions are true.

The above discussion makes clear that a better understanding of the genetic basis of fitness variation is highly valuable. However, previous studies that have been conducted with this goal in mind suffer from a number of shortcomings (Wilkinson et al. 2015). First, quantitative genetic studies often rely on small samples sizes and are consequently associated with large uncertainties (Chippindale et al. 2001; Gibson et al. 2002). Second, efforts to map genome-wide fitness have often focussed on a specific sexually selected trait, rather than broader measurements of fitness (Reinhold 1998; Fitzpatrick 2004; Mackay 2010; Johnston et al. 2011; Randall et al. 2013). Third, studies that have managed to measure fitness more broadly are often problematic for other reasons. For example, the *D. melanogaster* hemiclones used to identify sex-specific and sexually antagonistic genes (Innocenti & Morrow 2010) focussed on gene expression differences, which are not necessarily indicative of the underlying causal polymorphisms. Similarly, a study of sex-specific selection in humans could not distinguish between variants with sexually concordant and sexually antagonistic effects (Lucotte et al. 2016).

To address these limitations, uncover the genetic basis of sex-specific fitness and clarify the causes of its maintenance, I perform a GWAS of sex-specific fitness in *D. melanogaster*. To overcome the difficulty of associating a given genotype with simultaneous measurements of male and female fitness, I make use of the hemiclonal design (Abbott & Morrow 2011), which enables a single half-genome (‘hemiclone’) to be expressed in either sex

(see ‘1.5.3. Hemiclonal analysis’). I then combine sex-specific fitness measurements from ~200 lines with previously published whole-genome sequences from these same lines (Gilks et al. 2016) to perform a GWAS of sex-specific fitness.

This design has a number of beneficial features. First, the sex-specific fitness measurements are meaningful because they mimic the rearing conditions experienced by the outbred, laboratory-adapted population (LH_M) from which the hemiclones are extracted. Second, the hemiclonal design allows a given hemiclonal genome to be measured across multiple replicate individuals of both sexes. Replicated phenotypic measurements decrease the environmental variance associated with measuring fitness, thus increasing study power. Third, the GWAS approach means that narrow regions of the genome can be associated with phenotypic measurements, which represents a substantial improvement in terms of precision relative to previous quantitative trait locus (QTL) designs (Mackay 2010).

Using GWAS data for male and female fitness, I ask two broad questions. First, what evolutionary forces maintain sex-specific fitness variation in this laboratory-adapted *D. melanogaster* population? To answer this question, I perform a range of analyses: (i) I decompose sex-specific fitness variation into sexually antagonistic and sexually concordant components, and estimate the genetic correlation between male and female fitness from this sample of hemiclones; (ii) I compare the distribution of sex-specific fitness variation (and its components) on autosomes and the X chromosome; and (iii) I examine the effect size, gene functions and variant effects associated with sex-specific fitness variation (and its components).

Second, what genetic mechanisms facilitate the maintenance of fitness variation? I address this question by examining the effect sizes, variant effects and functions associated with genetic variation for sex-specific fitness.

3.4. Methods

3.4.1. Sex-specific fitness measurements of LH_M hemiclones

In order to sample genetic variation from the laboratory-adapted LH_M population, 223 hemiclonal lines were created in 2012, ~545 generations since LH_M was established (Gilks et al. 2016). For more details on the rearing regime of the laboratory-adapted LH_M population and hemiclonal analysis, see sections '1.5.2. LH_M population' and '1.5.3 Hemiclonal analysis'.

Male and female lifetime adult fitness was measured for each of the 223 hemiclonal lines. The protocol followed previous studies (Chippindale et al. 2001; Rice et al. 2005; Innocenti & Morrow 2010; Collet et al. 2016) and mimics the rearing regime experienced by LH_M flies. In both assays, a competitor stock homozygous for the recessive, eye-colour mutation *brown* (*bw)* was used. This competitor stock is maintained by following an identical rearing regime to that of the base LH_M population and enables hemiclone x *bw* progeny (wild-type eyes) to be distinguished from *bw* x *bw* progeny (brown eyes).

Male fitness was measured as competitive fertilisation success. 'Adult competition' vials containing 5 hemiclonal males from a given line, 10 competitor *bw* males and 15 virgin *bw* females were set up. After two days, each *bw* female was isolated in an individual 'oviposition' vial containing no additional yeast and left to oviposit for 18 hours. On day 12 post egg-laying, progeny were scored for eye colour. Male fitness was calculated as the proportion of offspring sired by the 5 hemiclonal males (those with wildtype eye-colour), combining progeny data from the 15 oviposition vials. This

assay was repeated a further 5 times in a blocked design; estimates for each line were therefore based on fitness measurements from 25 hemiclonal males.

Female fitness was measured as competitive fecundity. Adult competition vials containing 5 virgin hemiclonal females from a given line, 10 competitor *bw* females and 15 *bw* males were set up. Two days later, the 5 hemiclonal females were isolated into individual oviposition vials and left to oviposit for 18 hours. These vials were immediately chilled at 4°C and fecundity measured as the number of eggs laid per female. This assay was replicated a further 5 times in a blocked design; each line estimate therefore measured the fitness of 25 hemiclonal females. Note that each block contains fitness measurements from each line; between-line variance attributable to block effects is therefore minimised.

Fitness data was subjected to quality control and pre-processing in preparation for quantitative genetic and association analysis. Male fitness data from competition vials where not all 5 focal males were present at the assay were excluded from further analysis. Similarly, female oviposition vials where the female died or where fewer than 2 eggs were present (indicating partial sterility or failure to mate) were also excluded. Proportion data from the male assays and count data from the female assays were then Box-Cox transformed, scaled, and centred around zero within each block. Data from each block was averaged to obtain one fitness estimate for each line and sex.

3.4.2. Sexually antagonistic and concordant phenotypes

In order to identify candidate sexually antagonistic and sexually concordant SNPs, a linear transformation of male and female fitness values for each hemiclonal line was undertaken (see Fig. 3.2. for illustration). Specifically, in a two-dimensional space where each axis represents scaled and centred sex-specific fitness values, a sexually antagonistic phenotype ('antagonistic index') and a sexually concordant phenotype ('concordant index') were defined through 45° clockwise rotation of the coordinate system of the male and female fitness plane. The antagonistic axis defined in this way quantifies the position of individual lines on a continuum ranging from extremely male-beneficial, female-detrimental (low antagonistic index), through to extremely female-beneficial, male-detrimental (high antagonistic index) effects. Similarly, the sexually concordant axis measures whether genotypes are generally detrimental across the two sexes (low concordant index) or generally beneficial across the two sexes (high concordant index). The approach used for mapping both indices is analogous to the frequently used transformation of weight and height into a Body Mass Index (BMI) in order to quantify obesity in humans.

3.4.3. Genotyping

DNA extraction, whole-genome sequencing and SNP calling pipelines are described in more detail in Gilks et al. (2016). Briefly, DNA was extracted from a female heterozygous for the hemiclonal genome and a genome complement derived from the sequenced reference stock (Bloomington *Drosophila* Stock Center no. 2057). Samples were sequenced on the HiSeq 2500 (Illumina) platform. The median coverage across all samples was 31X.

Single nucleotide polymorphisms (SNPs) were called using the BWA-Picard-GATK pipeline. Burrow-Wheels aligner (Li et al. 2009) *mem* was used to map cleaned sequence reads to *D. melanogaster* genome assembly release 6. Fine-mapping was performed with *Stampy* (Lunter & Goodson 2011) and *Genome Analysis Tool-Kit* (*GATK*) v3.2.2 (Van der Auwera et al. 2013). SNPs were called using *HaplotypeCaller* (GATK v3.4-0) relative to the BDGP+ISO1/dm6 assembly (Dos Santos et al. 2015). Reads with base quality<20 were omitted, as were reads with a stand call and emit value<31. Hard-filtering was applied using the following thresholds: Quality-by-Depth>2, strand bias<50, mapping quality>58, mapping quality rank sum>-7.0 and read position rank sum>-5.0. Indels and non-biallelic SNPs were removed prior to further analysis. Note that three of the 223 phenotyped samples did not yield sufficiently high-quality sequence data and were removed from further analysis. 220 hemiclonal remained for further quality control.

3.4.4. Quality control of genotypes

Further quality filtering steps were applied in *vcftools* (Danecek et al. 2011) and *PLINK* (Purcell et al. 2007). First, sites with depth <10 and genotype quality <30 were removed. Second, only individuals with >85% non-missing positions were retained. Third, positions with poor genotype information across all retained individuals (<95% call rate) were discarded. Finally, given the relatively small sample size of the dataset as a whole and the low power of an association test for rare variants, only common variants (MAF>0.05) were retained for further analysis. From an initial dataset of 220 hemiclones

containing 1,312,336 SNPs, a quality-filtered dataset of 765,980 SNPs from 203 individuals remained at this stage.

To detect any genotypic outliers, I examined LH_M's population structure using principal components analysis (PCA). Overlapping SNP positions from the 203 LH_M genomes and from an outgroup population (RAL) consisting of 205 North American whole-genomes sequenced as part of the *Drosophila* Genetic Reference Panel (Mackay et al. 2012; Huang et al. 2014; Lack et al. 2015) were used as input to construct a genetic similarity matrix. This set of SNPs was pruned for linkage disequilibrium (LD) such that no two SNPs with r^2>0.2 within 10Kb remained. The top 5 PC axes explaining the most variation were extracted in *LDAK* ('Linkage-Disequilibrium Adjusted Kinships' (Speed et al. 2012)) and inspected in pairwise plots. After removal of one outlier, the final dataset used for association analysis contained 202 individuals and 765,764 SNPs.

3.4.5. Heritability analyses

I estimated the SNP heritability (h^2_{SNP}) of all four phenotypes in *LDAK* (Speed et al. 2012). This approach uses Restricted Maximum Likelihood (REML) to fit a linear mixed model that expresses the vector of phenotypes Y as a function genome-wide SNP genotypes, treated as random effects:

$$Y \sim N(0, \sigma^2_{SNP} K + \sigma^2_e I)$$

Y is a vector of phenotypes, K the kinship matrix, σ^2_{SNP} the additive genetic variance, I an individual identity matrix and σ^2_e the residual variance. From REML variance components, SNP heritability (h^2_{SNP}) can then be estimated as: $h^2_{SNP} = \sigma^2_{SNP}/(\sigma^2_{SNP} + \sigma^2_e)$.

LDAK corrects for local linkage when calculating SNP heritabilities to avoid inflation of h^2_{SNP} in clusters of linked sites that otherwise arises because several SNPs tag the same causal polymorphism. SNPs are weighted inversely proportional to their local linkage, such that SNPs in high LD contribute less to h^2_{SNP} than SNPs in low LD. *LDAK* also allows one to set the parameter α that determines how SNPs are weighted by their MAF (as MAF$^{\alpha}$) when calculating the kinship matrix K. I used the default of α=–0.25 which provides a steeper relationship between MAF and h^2_{SNP} than the value of -1 that is frequently used in studies on humans (in other words, SNPs with higher MAF are assumed to contribute more to h^2_{SNP} than SNPs with low MAF). This model has been shown to substantially improve heritability estimates across a wide range of traits (Speed et al. 2017). Note that running these same analyses without MAF weighting and without LD-weighting (*i.e.*, by applying the GCTA model (Yang et al. 2011)) produces very similar results (not shown).

Empirical significance was assessed by permuting phenotype labels 1,000 times, re-calculating h^2_{SNP} on each permutation as above, and counting the number of permuted estimates which exceeded the observed.

To estimate the intersexual genetic correlation for fitness, r^W_{mf}, I used *GEMMA* (Zhou & Stephens 2014), which implements a bivariate linear mixed model that is identical in form to the univariate model described above, except that Y is a two-column matrix of male and female fitness. This software was used because *LDAK* does not implement a bivariate analysis. Note also that the *GEMMA* model sets α=–1 and does not LD-weight SNPs.

3.4.6. Genome-wide association mapping

For each of the four phenotypes, I performed a GWAS by applying a linear mixed model to test the effect of allelic variants at each SNP on the phenotypic values. The linear mixed model approach includes a kinship matrix as a random effect to account for the heritable portion of genetic variation attributable to genome-wide genetic similarity between individuals. This approach has been shown to effectively control the false positive rate and increase power to detect true associations in samples with moderate degrees of population structure and close relatedness, such as LH_M (Astle & Balding 2009; Price et al. 2010).

The linear mixed model can be expressed as:

$$Y = \beta X + g + e$$

where

$$Var(g) = N(0, \sigma^2_{SNP} K)$$

$$Var(e) = N(0, \sigma^2_e I)$$

Y, σ^2_{SNP}, K, σ^2_e, I are defined as above, while X defines the genotype at the SNP being tested and β denotes the fixed effect coefficient (effect size) associated with the allele.

GWAS were implemented in *LDAK* (settings as above) and a Wald χ^2 test was used to generate P-values for each position. To confirm that the linear mixed model effectively controlled for genetic confounding, I estimated the genomic inflation factor (λ_{median}; calculated as median χ^2_{obs} / median χ^2_{exp}) using *GenABEL* (Aulchenko et al. 2007). λ_{median} quantifies the extent of inflation due to relatedness and population structure; a value close to 1 suggests that genetic confounding has been well controlled.

3.4.7. Defining candidate SNPs and genes

I corrected for multiple testing using a False Discovery Rate (FDR) approach, where each SNP was assigned a probability of being a false positive (Benjamini & Hochberg 1995), and converted P-values into Q-values. The false discovery rate approach achieves a balance between false positives and false negatives that can be considered appropriate for characterising the general properties of many variants (Bergland et al. 2014). I defined candidate SNPs as sites with Q-values<0.3 and non-candidate SNPs as sites with Q-values>0.3.

For phenotypes where candidate SNPs (those with Q-values<0.3) were present, I estimated the number of independent associated loci through LD clumping in *PLINK*. LD clumping takes the first candidate SNP along the genome as an 'index SNP' and clusters neighbouring SNPs (*i.e.*, those within a specified distance and LD threshold) around the index SNP, forming one 'clump' that is approximately independent from other clumps. An LD (r^2) threshold of 0.4 and a distance threshold of 10Kb were specified in this instance.

For functional analyses, a gene-based association test in *LDAK* was also performed. This test partitions h^2_{SNP} into genes via REML and computes a P-value using a likelihood ratio test, while correcting for local relatedness using the SNPs in each gene. Gene start/end coordinates were downloaded from the UCSC genome browser and extended by 5Kb up- and 5Kb downstream to include potential regulatory regions (Ensembl default). P-values for each gene were obtained using options '–calc-genes-reml',

'ignore-weights YES' and α=–0.25, after which P-values were transformed into Q-values (this Q-value was used as input for GO enrichment tests).

3.4.8. Genomic distribution of fitness variation

To examine the relative contribution of autosomes and the X chromosome to genetic variation for a given phenotype, I used two complementary methods. First, I partitioned the genome into X chromosome and autosome subsets and calculated h^2_{SNP} via REML in *LDAK* each subset in turn (settings as above). The observed proportion of h^2_{SNP} contributed by each compartment was then compared to the expected proportion (*i.e.*, the fraction of LD-weighted predictors belonging to each compartment) using a two-sample Z-test.

As a complementary approach, and when possible (*i.e.*, when SNPs with Q-value<0.3 were present), I also compared the proportion of candidate SNPs mapping to the autosomes and X chromosome to the proportion of all SNPs mapping to autosomes and X chromosome. Significant over- or under-representation was assessed using a χ^2 test.

3.4.9. Functional analyses of candidate loci

To assess the relative contribution of functional categories to fitness, I first mapped all SNPs to functional categories using the Variant Effect Predictor (McLaren et al. 2010). Total h^2_{SNP} for a given phenotype was then partitioned into functional categories, and the observed proportion of h^2_{SNP} contributed by each category estimated using REML in *LDAK* (settings as above). I then used a permutation test to compare observed and expected h^2_{SNP} for each

functional subset and determine statistical significance. Here I shifted genome-wide annotations to a random starting point along a 'circular genome', which breaks the relationship between each SNP and its annotation while preserving the order of annotations and their associated LD structure (Cabrera et al. 2012). h^2_{SNP} was re-calculated via REML for each of 1,000 permuted datasets and two-tailed P-values determined as the sum of permuted estimates with more extreme absolute values than the observed.

I also investigated the functional properties of genes associated with the phenotypes of interest. To do so, I performed a Gene Ontology (GO) analysis in *PANTHER* (Protein Analysis Through Evolutionary Relationships) *v.13.1* (Mi et al. 2017), using the statistical enrichment tool. I used Q-values derived from *LDAK*'s gene-based association test as input. A Wilcoxon Rank-Sum test was then applied to generate P-values for each GO term by comparing the Q-value ranks of genes mapping to a particular category to Q-value ranks from all genes. FDR correction was applied to the P-values from this test.

3.4.10. Statistical software

All statistical analyses were performed in *RStudio* (RStudio Team 2015).

3.5. Results

3.5.1. LH_M population structure

Whole-genome data from 202 quality-filtered LH_M genomes showed that this population forms a well differentiated cluster when compared to a known outgroup population of *D. melanogaster* flies from the USA (RAL) (Fig. 3.1A). While some pairs of individuals show relatively high degrees of relatedness, the 202 LH_M hemiclonal lines generally have low genome-wide kinship (Fig. 3.1B). This indicates that the lines represent a good sample of the overall genetic diversity within the population.

In accordance with the modest effective population size of this population, there are relatively high background levels of LD (r^2~0.4) at distances (1Kb) where LD is negligible in wild populations (Mackay et al. 2012; Pool et al. 2012). As expected from its smaller effective population size, the X chromosome displays higher levels of LD than autosomes (Fig. 3.1C).

3.5.2. Male and female heritabilities

Figure 3.2 presents standardised male and female fitness data from each of 223 hemiclonal lines. Using the 202 quality-filtered genomes to estimate SNP heritability, I found that both male and female fitness have appreciable heritabilities. Female fitness heritability is higher than male fitness heritability (female h^2_{SNP}=0.59, SD 0.13, P<0.001; male h^2_{SNP}=0.29; SD 0.16; P=0.007). This difference is driven primarily by the additive variance component (female σ^2_{SNP}=0.23, SD 0.07; male σ^2_{SNP}=0.08, SD 0.05), while environmental variances are comparable between the sexes (female σ^2_e=0.16, SD 0.05;

male σ_e^2=0.19, SD 0.04). The antagonistic and concordant components of sex-specific fitness also have appreciable heritabilities (antagonistic index h_{SNP}^2=0.51, SD 0.15, P=0.001; concordant index h_{SNP}^2=0.43, SD 0.16, P=0.002).

3.5.3. Genome-wide association studies of sex-specific fitness and its components

Figure 3.3A,B presents P-values from a mixed model association analysis of female and male fitness respectively. The genomic inflation factors for both phenotypes are close to one, suggesting that the mixed model approach employed controls well for relatedness and population structure (female λ_{median}=1.07; male λ_{median}=1.00).

For female fitness, the most significant association P-value is 4.22 x 10^{-6}. This value does not reach the Bonferroni-corrected threshold (6.52 x 10^{-8}). The minimum Q-value value was 0.36. For male fitness, the most significant association P-value is 4.01 x 10^{-6} and the minimum Q-value is 0.19. There are 248 SNPs with Q-values<0.3 (31 LD-independent clusters) which represent candidates for the genetic basis of male fitness.

Figure 3.3C,D presents P-values from an association analysis of the antagonistic and concordant index, respectively. These phenotypes also show little evidence of inflation due to relatedness or population structure (antagonistic λ_{median}=0.97; concordant λ_{median}=1.06). The most significant association P-value for the antagonistic index is 1.27 x 10^{-6}, with a minimum Q-value value of 0.26. 2,372 SNPs have Q-values<0.3, representing 2,372 candidate antagonistic SNPs (226 LD-independent clusters). For the

concordant index, the most significant association P-value is 1.49 x 10^{-5}, and the minimum Q-value value is 0.78.

3.5.4. Relative contributions of antagonistic and concordant components to sex-specific fitness variation

Estimating the intersexual genetic correlation for fitness (r_{mf}^{W}) permits an estimation of the relative contributions of antagonistic and concordant genetic variation to sex-specific fitness variation. Using bivariate REML to estimate r_{mf}^{W} reveals that male and female fitnesses are uncorrelated (r_{mf}^{W}=0.03, SE 0.28).

3.5.5. Genomic distribution of sex-specific fitness variation

I partitioned SNP heritability into chromosomal compartments to test whether genetic variants associated with male and female fitness are disproportionately associated with the X chromosomes or autosomes. This analysis showed that there is less female h_{SNP}^{2} than expected on the X chromosome (h_{SNP}^{2} observed=0.05, h_{SNP}^{2} expected=0.20; X-linked enrichment=0.25, SD 0.63), while there is more male h_{SNP}^{2} than expected on the X chromosome (h_{SNP}^{2} observed=0.55, h_{SNP}^{2} expected=0.20; X-linked enrichment=2.72, SD 1.24). In both cases, however, the difference is not statistically significant (Fig. 3.4A,B; females: two-sample Z test, P=0.293; males: two-sample Z test, P=0.085). For male fitness, where individually significant SNPs (Q-value<0.3) were detected, it was additionally possible to compare the number of candidate to non-candidate SNPs on the X chromosome and autosomes. This analysis revealed significant enrichment

of candidate male SNPs on the X chromosome (χ_1^2=1409.6, observed=222, expected=29.7, X-enrichment=7.47, P<0.001).

Partitioning SNP heritability of the antagonistic and concordant indices by chromosomal compartment, I found that neither the antagonistic index (Fig. 3.4C; h_{SNP}^2 observed=0.39, h_{SNP}^2 expected=0.20; X-linked enrichment=1.92, SD 0.89; Two-sample Z test, P=0.274) nor the concordant index (Fig. 3.4D; h_{SNP}^2 observed=0.16, h_{SNP}^2 expected=0.20; X-linked enrichment=0.77, SD 0.92; two-sample Z test, P=0.789) are enriched on either compartment. There is also no significant enrichment of the 2,372 candidate antagonistic SNPs on the X chromosome; in fact, there is a trend towards antagonistic SNPs being disproportionately situated on autosomes (χ_1^2=3.67, observed=253, expected=283.7, X-enrichment=0.89, two-tailed P=0.055)

3.5.6. Functional characteristics of sex-specific fitness variation

By partitioning SNP heritability into functional categories, I tested whether genetic variants associated with sex-specific fitness are disproportionately associated with specific functions. Doing so, I found that female h_{SNP}^2 is enriched for missense variants and depauperate for synonymous variants (Fig. 3.5). Female fitness genes are under-represented (Q-value<0.05) for the protein class ‘esterase’, and weakly (Q-value<0.3) associated with a number of molecular functions (Table 3.1).

There is no significant enrichment for male h_{SNP}^2 among any variant effect category (Fig. 3.5). Male fitness genes are weakly (Q-value<0.3)

associated with two biological processes: ‘hormone metabolic process’ and ‘regulation of hormone levels’ (Table 3.1); no other enrichments are found.

Focussing on antagonistic and concordant indices, I found that antagonistic h^2_{SNP} is enriched among missense variants (Fig. 3.5). Antagonistic genes are weakly associated with various molecular functions (Table 3.1). There is no significant enrichment for concordant h^2_{SNP} among any variant effect category (Fig. 3.5). Concordant genes are under-represented for the biological processes ‘somitogenesis’, ‘somite development’, ‘chordate embryonic development’ and the molecular function ‘glucuronosyltransferase activity’ (Table 3.1).

3.6. Discussion

I conducted a genome-wide association study of sex-specific fitness across 202 hemiclones derived from a laboratory-adapted population of *D. melanogaster* flies. I found that genetic variation in sex-specific fitness is explained by the joint contributions of polymorphisms with sexually antagonistic and sexually concordant effects. While there is no bias of antagonistic variants on the X chromosome, functional analyses showed that coding variation contributes disproportionately to the sexually antagonistic component of fitness variance. I discuss these findings in light of previous theory about the maintenance of fitness variation, the genetic architecture of fitness variants and the evolution of sexual dimorphism.

Numerous studies have shown that wild and laboratory populations across taxa exhibit appreciable levels of additive genetic variation for fitness (Mousseau & Roff 1987; Roff & Mousseau 1987; Messina 1993; Pomiankowski & Moller 1995; Fowler et al. 1997; Kruuk et al. 2000; McCleery et al. 2004). Consistent with this, analysis of 202 *D. melanogaster* hemiclonal genomes uncovers high SNP heritability for adult fitness in both sexes. Similar heritabilities are estimated based on these same data using phenotypic values alone (see Appendix A; female h^2_{SNP}=0.42; male h^2_{SNP}=0.16). The estimates are also in line with previous studies in this population (Pischedda & Chippindale 2006; Long et al. 2009; Collet et al. 2016), including re-analyses of Innocenti & Morrow's (2010) phenotypic data presented in Chapter 2.

To clarify the causes of high fitness heritability, I estimated the intersexual genetic correlation for fitness (r^W_{mf}). I found that r^W_{mf} is not

significantly different from zero. The lack of a positive or negative r_{mf}^{W} has two possible interpretations: (i) alleles tend to have sex-limited effects, and thus no genetic correlation arises because male and female fitness values are determined by two independent sets of loci; and (ii) alleles simultaneously affect the fitness of both sexes, but some variants have sexually antagonistic fitness effects (reducing r_{mf}^{W}) while others have sexually concordant fitness effects (increasing r_{mf}^{W}), and thus no net genetic correlation arises either.

There are two reasons to favour the latter explanation. First, it is known from previous studies that strong positive intersexual correlations are observed for most traits in *Drosophila* (Cowley & Atchley 1988; Ayroles et al. 2009; Poissant et al. 2010; Griffin et al. 2013) and that most genes in *Drosophila* do not have sex-limited expression (Connallon & Clark 2011a; Parsch & Ellegren 2013; Ingleby et al. 2015). The target size for mutations with sex-limited effects is therefore restricted—an observation that is at odds with the appreciable heritabilities estimated for sex-specific fitness and their antagonistic and concordant components in this study. Second, comparison of Q-values for the antagonistic and concordant index showed that antagonistic variants are enriched for low Q-values relative to variants along the concordant index, where the minimum Q-value is 0.78. This discrepancy is difficult to reconcile with a predominant role for sex-limited variation, which would predict symmetrical effect sizes along both axes. It instead suggests that antagonistic and concordant components harbour variation maintained through fundamentally different processes (e.g. balancing selection and mutation-selection balance respectively). In short, the absence of a

significant intersexual fitness correlation can be best interpreted as the result of shared contributions from (i) variants with sexually antagonistic effects and (ii) variants with sexually concordant effects, which are most likely evolving under directional selection and are maintained at mutation-selection balance.

I next examined the genomic distribution of fitness variation. Classic theory predicts that antagonistic variation will be enriched on the X chromosome (Rice 1984) while concordant variation evolving under mutation-selection balance will be enriched on the autosomes (Haldane 1937). However, I find no evidence for enrichment of antagonistic h^2_{SNP} on the X chromosome, nor enrichment of concordant h^2_{SNP} on autosomes.

Several factors could explain the absence of a chromosomal effect on antagonistic and concordant h^2_{SNP}. First, the absence of X-enrichment of antagonistic h^2_{SNP} could be due to genetic drift. Genetic drift disproportionately affects the X chromosome due to its smaller effective population size (Caballero 1995) and reduces the effectiveness of antagonistic selection on the X (Mullon et al. 2012). This could disproportionately deplete X-linked antagonistic variation, counteracting the X-enrichment expected based on classic theory (where genetic drift is not taken into account). Analyses of LD show that drift indeed plays an important role in LH_M: not only is background LD in LH_M generally higher than observed in wild populations, it is also appreciably higher on the X chromosome than autosomes. Additionally—and as discussed in Chapter 2—there is a revealing contrast between early studies of fitness in LH_M, which reported negative r^W_{mf}s (Chippindale et al. 2001) and disproportionate contributions of the X chromosome to negative r^W_{mf}s (Gibson et al. 2002), and more recent

studies (including this one) where no such patterns are detected (Long et al. 2009; Collet et al. 2016; Morrow et al. 2008). These changes are consistent with drift gradually depleting X-linked antagonistic variation during the prolonged maintenance of LH_M at modest effective population sizes (Gilks et al. 2016; Rice et al. 2005).

Aside from drift, a number of additional processes could shift the conditions for the maintenance of antagonistic polymorphisms and explain the lack of X-enrichment of antagonistic h^2_{SNP}. For example, if allelic dominance differs between the sexes, the autosomes become a more favourable environment for the accumulation of antagonistic polymorphisms because the sex-averaged fitness of heterozygotes is higher than the sex-averaged fitness of either homozygote (Fry 2010; Spencer & Priest 2016). Such a mechanism has been recently shown to maintain the antagonistic *VGLL3* polymorphism in salmon (Barson et al. 2015). Alternatively, linkage between multiple antagonistic alleles (Patten et al. 2010; Úbeda et al. 2011), assortative mating based on fitness (Arnqvist 2011), or epistasis between loci (Arnqvist et al. 2014) can increase the likelihood that autosomal polymorphisms are maintained. While there is no direct evidence for any of these mechanisms affecting individual antagonistic loci, linkage (Slatkin 2008), assortative mating (Crespi 1989) and epistasis (Huang et al. 2012) are commonly occurring processes that could plausibly contribute to the absence of X-enrichment of antagonistic h^2_{SNP}.

Finally, the absence of autosomal enrichment of concordant h^2_{SNP} could be due to a biased screen for X-linked and autosomal fitness effects, which favours the detection of X-linked effects in males. The X-linked bias in males

arises because allelic effects are not averaged out across two copies on the X in males as they are in females (James 1973; Cowley & Atchley 1988; Wayne et al. 2007). X-linked variation is effectively 'magnified' in a screen for fitness (Reinhold & Engqvist 2013) and potentially counteracts the lower underlying X-linked heterozygosity expected under mutation-selection balance. Supporting this interpretation, the data presented here shows that candidate SNPs for male fitness are disproportionately situated on the X chromosome. Additionally, enrichment of male h^2_{SNP} on the X chromosome—although not statistically significant—is the largest in terms of effect size among the four phenotypes studied here.

Overall, the simple prediction that antagonistic polymorphisms will be predominantly X-linked and concordant polymorphisms predominantly autosomal is not borne out in this population. Nevertheless, variation underlying male fitness does appear to be disproportionately X-linked, as expected from theory (James 1973; Cowley & Atchley 1988; Wayne et al. 2007). Overall, comparisons of X-linked and autosomal fitness variation highlight the importance of considering other evolutionary processes (such as drift) and features of experimental design (such as the increased visibility of the X chromosome in a fitness screen) in making conclusions about the identity and distribution of sex-specific fitness variants.

Aside from the evolutionary forces maintaining fitness variation, a key knowledge gap concerns the genetic architecture these variants. Functional analyses presented here emphasize two main features. First, fitness variation is highly polygenic. This is made clear by the fact that SNP heritabilities for all four phenotypes are high, yet few SNPs reach low Q-

values (none below 0.19). Such a pattern is expected if many loci of small effect, rather than few loci of large effect, together contribute to fitness variation. High polygenicity is a pre-requisite for the genic capture model (Rowe & Houle 1996). It is therefore no surprise that there are no detectable large effect concordant polymorphisms (indeed, the lowest Q-value for the concordant index is 0.78 despite the appreciable heritability of this component).

It is perhaps more surprising to find no large effect antagonistic loci, given that large effect loci are easiest to maintain (Connallon & Clark 2012) and given empirical evidence for large effect loci in Soay sheep (Johnston et al. 2011) and salmon (*Salmo salar*) (Barson et al. 2015) (in studies with similar experimental designs and sample sizes). One possible explanation for the absence of large effect antagonistic loci is genetic drift in LH_M, which could have acted to purge these loci during the population's laboratory maintenance. Additionally, the loci described by Johnston et al. (2011) and Barson et al. (2015) are only 'large effect' with respect to the specific component of fitness measured (horn size and age at maturation, respectively), not necessarily with respect to total fitness. This narrow definition of fitness could bias the effect size estimate upward. Finally, if most traits under sexually antagonistic selection are highly polygenic, selection coefficients associated with individual antagonistic loci are likely to be very small, thus reducing the probability that each individual polymorphism is maintained under balancing selection (Turelli & Barton 2004).

A second functional pattern highlighted in this study is the association between antagonistic variation and missense effects. Complementary

analyses presented elsewhere confirm this association (see Appendix A). Given the low levels of genetic variation typically observed in coding regions (McDonald & Kreitman 1991), this pattern may seem surprising. However, it can be understood if one considers the process required to resolve sexual antagonism in coding and regulatory regions respectively. In coding regions, the resolution of sexual antagonism requires a two-step process in which a gene is first duplicated and the two paralogs subsequently gain regulatory sequences that permit sex-specific expression. In regulatory regions, resolving sexual antagonism can be achieved by the re-purposing of existing *cis*-regulatory elements, such as the addition of a new binding site (Stewart et al. 2010). Minor modifications of existing regulatory elements are widely observed and may therefore be relatively easy to achieve (Williams & Carroll 2009), but a duplication event followed by the evolution of a new sex-specific regulatory element is likely to require much longer waiting times. This simple scenario predicts that coding regions will accumulate antagonistic polymorphisms whereas regulatory regions will more rapidly resolve conflicts—an expectation that is in line with patterns seen here. Coding regions therefore appear to play a key role in the accumulation of antagonistic polymorphisms and the maintenance of fitness variation more generally.

In summary, I performed the first genome-wide association study of sex-specific fitness and its antagonistic and concordant components. The results presented here reveal that antagonistic variation is disproportionately associated with coding variation. This association, and the difficult process of antagonistic resolution it suggests, motivates further work to consider the

relationship between antagonistic variation and molecular patterns of balancing selection (see Chapter 4).

3.7. Figures and Tables

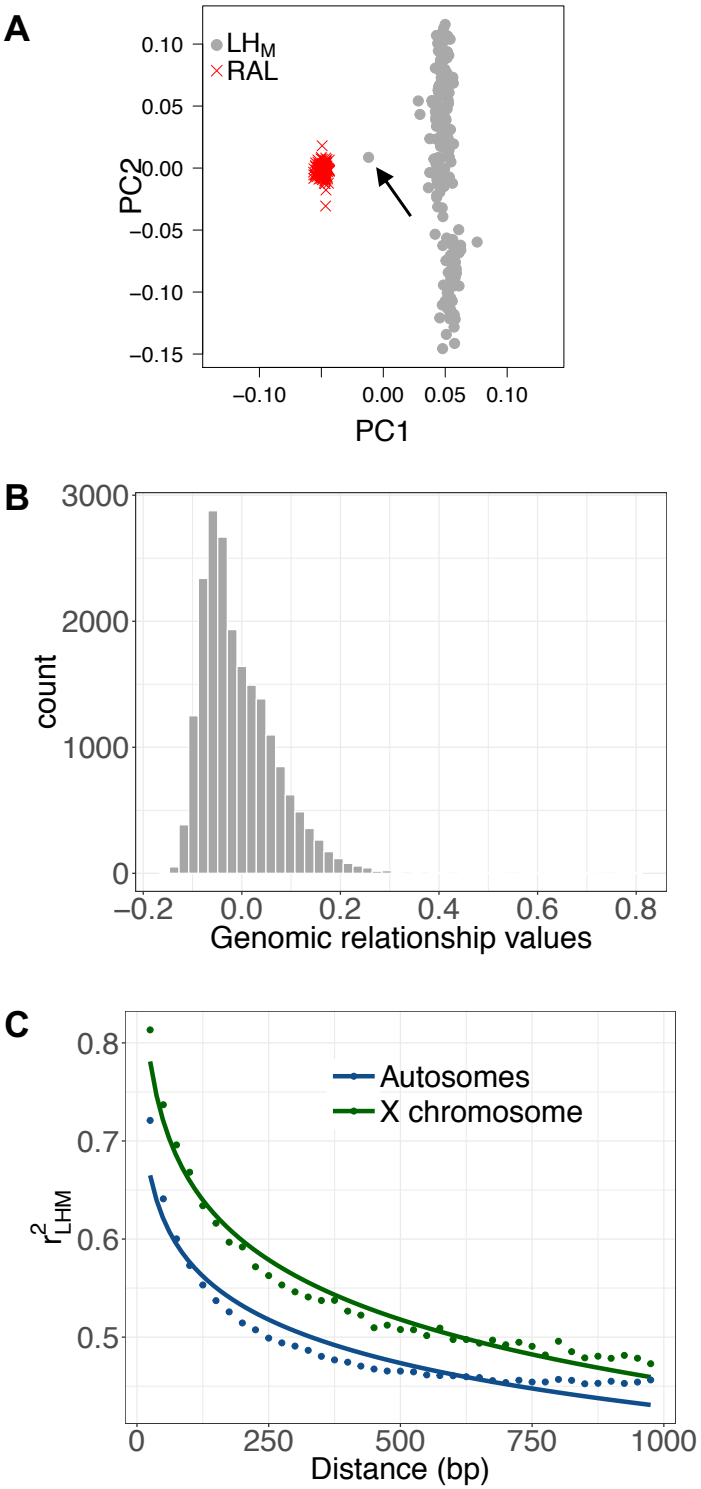


Figure 3.1. LH_M population structure. A. Scatter plot of the first and second principal components of a principal component analysis (PCA) constructed from SNPs present among 203 LH_M genomes (grey) and 205 genomes from the RAL (red), an outgroup population. Principal components were computed from common (MAF>0.05), LD-pruned (no two SNPs with r^2>0.2 within 10Kb) and high quality (site-level call rate>95%) sites only. One notable outlier individual (black arrow) was removed prior to association mapping. **B.** Histogram of off-diagonal genomic relationship values between the 202 LH_M individuals retained for association testing. **C.** Linkage disequilibrium (r^2) between pairs of SNPs within 1Kb of each other, split by chromosomal compartment (autosomal or X-linked). Points represent mean LD in 25bp bins; curves represent a fitted declining exponential relationship between LD and distance.

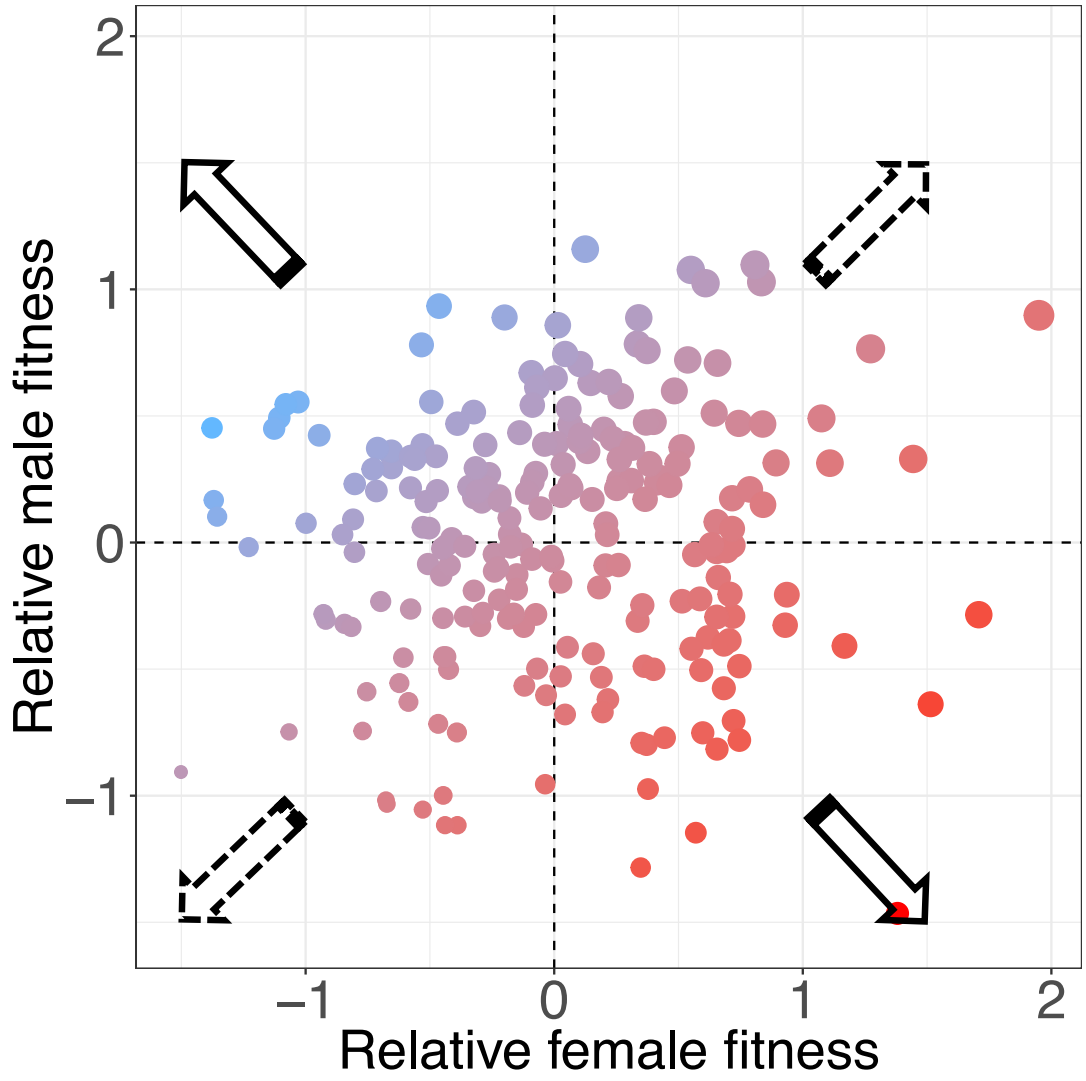

Relative male fitness
2
1
0
−1
−1
0
1
2
Relative female fitness

Figure 3.2. Sex-specific fitness measurements. Estimates of male and female relative lifetime reproductive fitness estimates for 223 *D. melanogaster* hemiclonal lines. Fitness measures were Box-Cox transformed, scaled and centred. Colours denote each hemiclone's 'antagonistic index', *i.e.* its position along a spectrum (full black arrows) ranging from male-beneficial, female-detrimental fitness effects (blue) to female-beneficial, male-detrimental effects (red). Sizes denote each hemiclone's 'concordant index', *i.e.* its position along a spectrum (dashed black arrows) ranging from male-detrimental, female-detrimental fitness effects (small) to male-beneficial, female-beneficial effects (large).

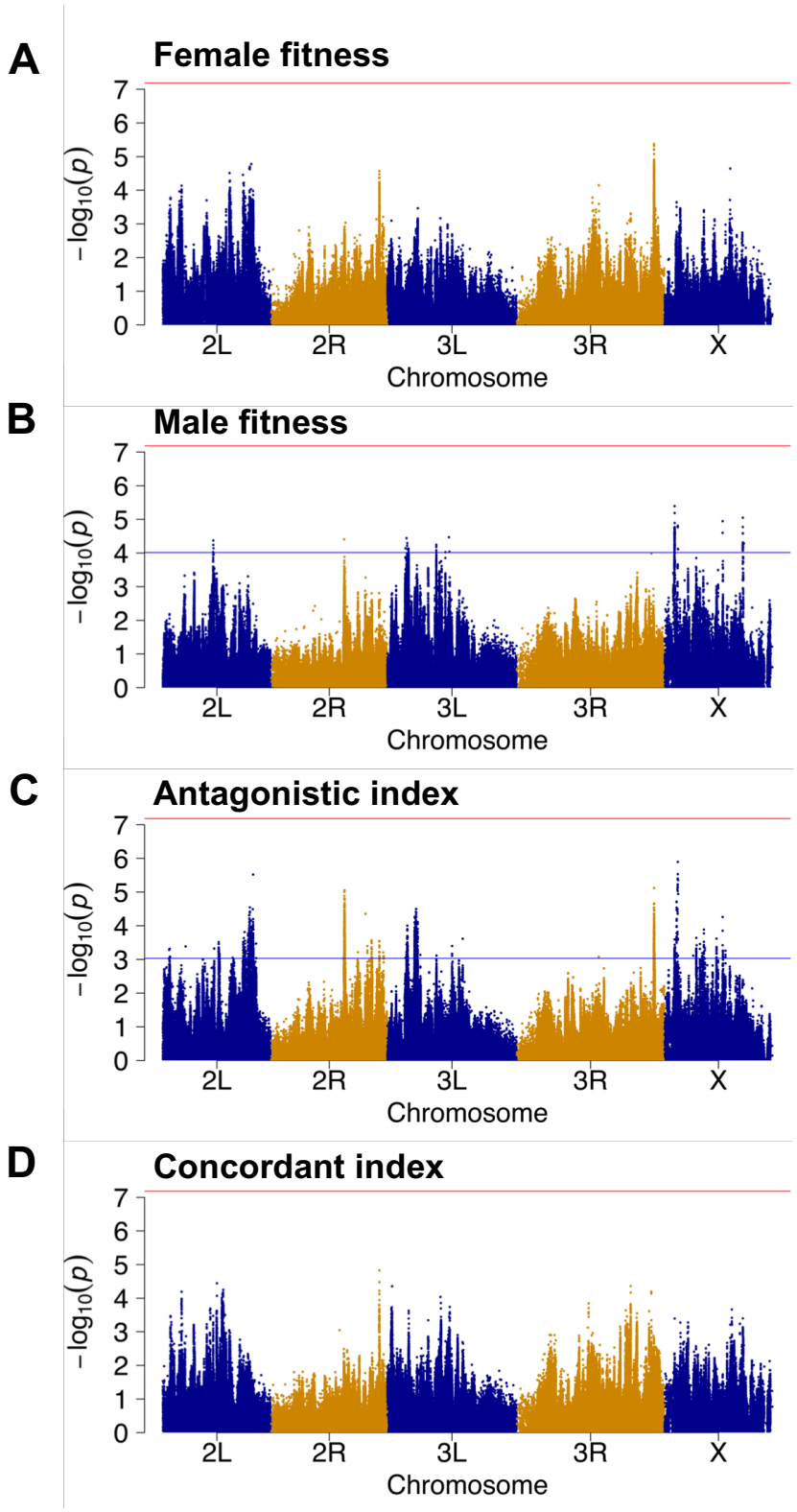

A
Female fitness
B
Male fitness
C
Antagonistic index
D
Concordant index
−log10(p)
0
1
2
3
4
5
6
7
2L
2R
3L
3R
X
Chromosome

Figure 3.3. Genome-wide association studies. Association of each SNP with each of the four phenotypes under study, along the five major chromosome arms (coloured) of the *D. melanogaster* genome. $-log_{10}$(P) values from a Wald χ^2 association test of each predictor against female fitness (**A**), male fitness (**B**), the antagonistic index (**C**) and the concordant index (**D**) are plotted. For each phenotype, I implemented the association in *LDAK* under a linear mixed model, using a genetic similarity matrix as a random effect to correct for relatedness and population structure between hemiclonal lines. Red lines denote a Bonferroni significance threshold; blue lines represent a 30% FDR threshold.

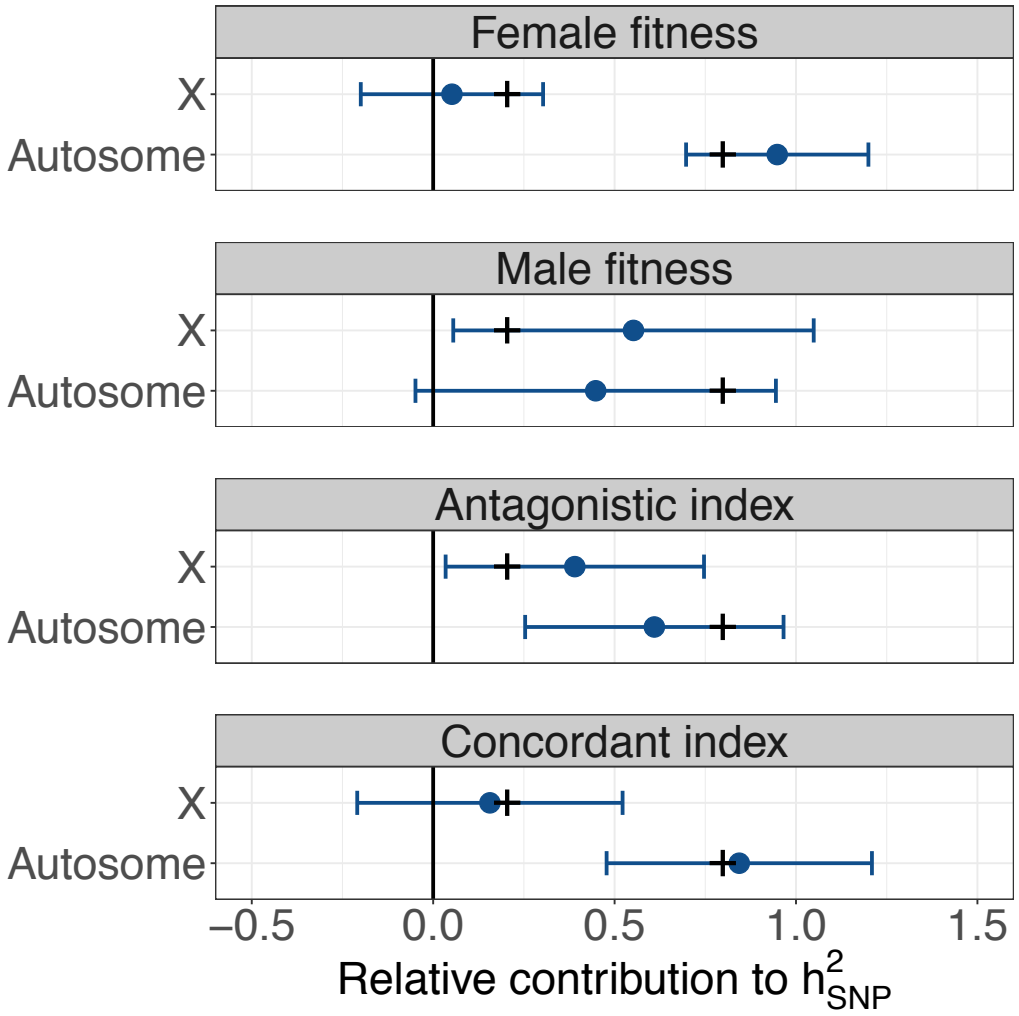


Figure 3.4. Chromosomal distribution of fitness variation. Relative contribution of different chromosomal compartments (X chromosome and autosomes) to each phenotype's SNP heritability (h^2_{SNP}). Dots represent point estimates for a given compartment's contribution to total h^2_{SNP} computed by Restricted Maximum Likelihood (REML) under the LDAK model (blue: P>0.05). Black bars represent 95% confidence intervals around these estimates (calculated as the h^2_{SNP} estimated for a given partition ±1.96SD). Black notches represent the expected contribution of each chromosomal compartment to total h^2_{SNP}, based on the fraction of SNP predictors in each compartment.

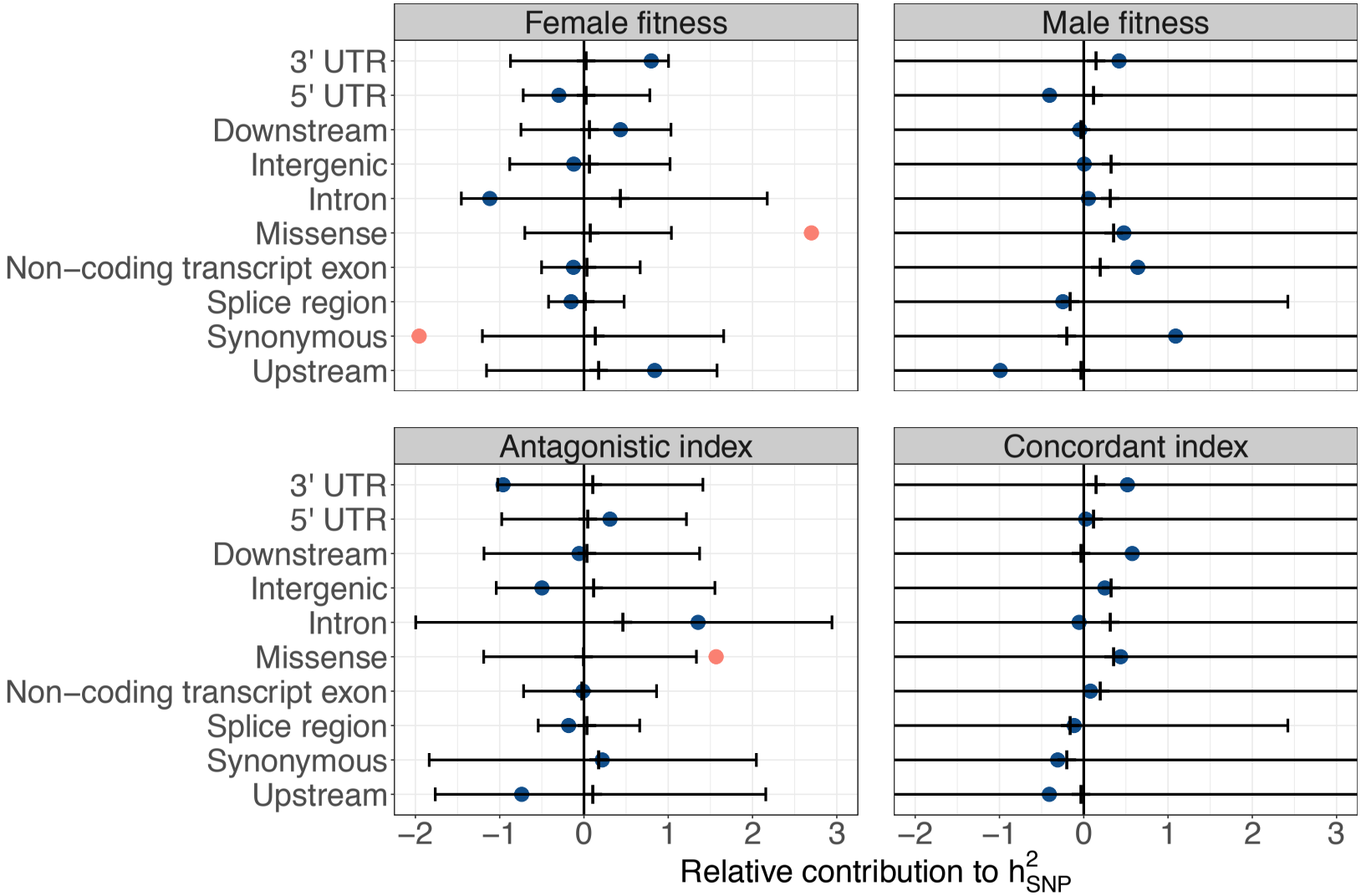

Female fitness
Male fitness
Antagonistic index
Concordant index
3' UTR
5' UTR
Downstream
Intergenic
Intron
Missense
Non−coding transcript exon
Splice region
Synonymous
Upstream
−2
−1
0
1
2
3
Relative contribution to h^2_{SNP}

Figure 3.5. Functional annotations of fitness variants. Relative contribution of different functional categories to each phenotype's SNP heritability (h^2_{SNP}). Blue dots represent point estimates computed by Restricted Maximum Likelihood (REML) under the LDAK model. Mean expected contribution of each genomic partition to phenotypic h^2_{SNP} (black notch) and 95% confidence intervals (black bars) were computed through 1000 'circular permutations' of annotation categories along the genome (see Methods).

Table 3.1. Gene functions associated with sex-specific fitness. Shown are enriched Gene Ontology (GO) and PANTHER Classifications for genes associated with each phenotype.

PHENOTYPE	GO TERM	DESCRIPTION	OVER/ UNDER	Q-VALUE
FEMALE FITNESS	PC00097	esterase	–	0.023
	GO:0004857	enzyme inhibitor activity	–	0.091
	GO:0015020	glucuronosyltransferase activity	–	0.091
	GO:0030414	peptidase inhibitor activity	–	0.091
	GO:0004866	endopeptidase inhibitor activity	–	0.114
	PC00090	defense/immunity protein	–	0.122
	PC00100	extracellular matrix glycoprotein	+	0.133
	GO:0061135	endopeptidase regulator activity	–	0.165
	GO:0004867	serine-type endopeptidase inhibitor activity	–	0.165
	GO:0033038	bitter taste receptor activity	–	0.165
	GO:0032450	maltose alpha-glucosidase activity	–	0.165
	GO:0003700	DNA binding transcription factor activity	–	0.179
	GO:0004556	alpha-amylase activity	–	0.186
	GO:0061134	peptidase regulator activity	–	0.186

MALE FITNESS	GO:0042445	hormone metabolic process	–	0.219
	GO:0010817	regulation of hormone levels	–	0.219
ANTAG-ONISTIC INDEX	GO:0016706	oxidoreductase activity	+	0.199
	GO:0004656	procollagen-proline 4-dioxygenase activity	+	0.199
	GO:0019798	procollagen-proline dioxygenase activity	+	0.199
	GO:0031543	peptidyl-proline dioxygenase activity	+	0.199
	GO:0051213	dioxygenase activity	+	0.199
	GO:0031545	peptidyl-proline 4-dioxygenase activity	+	0.236
	GO:0031386	protein tag	–	0.236
	GO:0031418	L-ascorbic acid binding	+	0.261
CONCOR-DANT INDEX	GO:0015020	glucuronosyltransferase activity	–	0.029
	GO:0001756	somitogenesis	–	0.031
	GO:0061053	somite development	–	0.031
	GO:0043009	chordate embryonic development	–	0.031

Chapter 4

Longstanding signatures of balancing selection associated with sexually antagonistic polymorphisms in *D. melanogaster*

4.1. Declaration

All work reported in this chapter is my own. I thank Aida Andrés for helpful discussions regarding analyses of balancing selection.

4.2. Abstract

Theory indicates that sexually antagonistic selection, where opposing alleles at given loci are selected in each sex, can elevate heterozygosity at the affected loci, thereby contributing to the maintenance of genetic variation. However, empirical research has neither investigated whether antagonistic loci are associated with signatures of balancing selection, nor the timescale over which antagonistic selection maintains polymorphisms. Here I present the first detailed analysis of the evolutionary dynamics of sexually antagonistic polymorphisms. I combine data on the identity of antagonistic loci from a genome-wide association study of sexual antagonism (Chapter 3) with publicly available polymorphism data from worldwide *D. melanogaster* populations to test whether antagonistic loci are associated with signatures of balancing selection. I find that genome-wide antagonistic loci bear multiple hallmarks of balancing selection: elevated minor allele frequencies, elevated regional polymorphism and reduced population differentiation. Furthermore, the molecular genetic effects of antagonistic selection span the *D. melanogaster* distribution range and are clearly detectable in populations that are separated from LH_M by over 10,000 years. There is even some indication that antagonistic selection operates in *D. simulans*, a sister species that is separated from *D. melanogaster* by ~1.5 million years. Overall, these results indicate that the constraints to sexual dimorphism reflected by antagonistic polymorphisms are highly evolutionarily persistent. This work supplements other recent research indicating that balanced polymorphisms may be more common than previously thought.

4.3. Introduction

Genetic polymorphisms are abundant in natural populations across taxa (Lewontin & Hubby 1966; The 1000 Genomes Project Consortium 2010; Pool et al. 2012). This observation is generally interpreted according to one of three models: (i) neutral evolution, under which extant levels of variation are primarily maintained at an equilibrium between input of new mutations and stochastic loss (Kimura 1983), (ii) directional selection, under which variation is maintained at an equilibrium between input of new mutations and loss through positive selection (Maynard Smith & Haigh 1974) or purifying selection (Charlesworth et al. 1993) (iii) balancing selection, under which selection acts to maintain rather than remove genetic variation at given loci (Dobzhansky 1955).

Balancing selection can occur through a variety of mechanisms (Maynard Smith 1998). With overdominance, a polymorphism is maintained because heterozygotes have higher fitness than either homozygote (Johnston et al. 2013; Hedrick 2012). This is seen in the case of the human sickle-cell polymorphism, where heterozygotes enjoy a viability benefit in geographical regions where malaria is endemic (Allison 1954; Haldane 1949). Under negative frequency-dependence, a balanced polymorphism arises because each allele is favoured when rare and disfavoured when common and thus neither is lost. This is seen at the major histocompatibility complex in vertebrates (Aguilar et al. 2004) and plant self-incompatibility loci (Vekemans & Slatkin 1994). Under spatially or temporally heterogeneous selection, each allele is favoured in a different environment and neither is lost due to migration between environments (Levene 1953; Hedrick 2006). This is

seen among seasonally fluctuating polymorphisms in *D. melanogaster* (Bergland et al. 2014).

Despite these examples, it remains unclear to what extent balancing selection contributes to standing genetic variation across the genome (Leffler et al. 2012). In contrast to directional selection and neutral evolution, whose genome-wide effects have been amply documented (Parsch et al. 2010; Elyashiv et al. 2016), a consensus has emerged that balancing selection plays a relatively minor role towards the maintenance of molecular genetic variation (Asthana et al. 2005; Bubb et al. 2006; Hedrick 2012). This view is based on the fact that balanced loci identified through 'candidate gene' approaches are few (Charlesworth 2006), and that genome-wide scans for balancing selection have similarly only detected a clear signal of selection in a handful of loci (Andrés et al. 2009; Leffler et al. 2013; Croze et al. 2017).

Yet despite this, there are good reasons to think that balancing selection could play an important role in maintaining molecular genetic variation. First, quantitative genetic studies have detected considerable additive genetic variance for fitness, implying that polymorphisms are not neutral. Furthermore, the estimated levels of variance are much higher than predicted under mutation-selection balance alone, indicating that genetic variation for fitness has not evolved solely under directional selection (Kelly & Willis 2001; Charlesworth et al. 2007; Charlesworth 2015). Taken together, these results suggest that balanced polymorphisms make an important contribution to standing genetic variation. Second, the dearth of balanced polymorphisms detected by genome-wide selection scans might in part be artefactual; scans for balancing selection are highly conservative because

they do not consider molecular patterns that could be confused with positive selection (e.g. extended linkage disequilibrium (LD), partial selective sweeps) or demographic effects (Fijarczyk & Babik 2015). The stringent steps required to avoid detecting other processes then limit these scans to focussing on ancient polymorphisms (Leffler et al. 2013) and are likely to cause them to miss a substantial portion of balanced loci. Finally, molecular population genetic studies have neglected sexually antagonistic selection, under which alleles at given loci have opposing fitness effects in each sex (Kidwell et al. 1977; Patten & Haig 2009). Given the well documented quantitative genetic fitness effects of sexual antagonism (Chippindale et al. 2001; Foerster et al. 2007; Brommer et al. 2007; Prasad et al. 2007), it is essential to evaluate the role of sexually antagonistic selection in maintaining molecular genetic variation if conclusions are to be drawn about the importance of balancing selection more generally.

Sexually antagonistic selection, like other mechanisms of balancing selection, is predicted to elevate heterozygosity at the affected loci owing to opposing selection pressures on the two alleles in each sex (Kidwell et al. 1977; Patten & Haig 2009). Antagonistic polymorphisms are particularly likely to have elevated heterozygosity when sex-specific selection coefficients are large and similar between the sexes; under these conditions, antagonistic selection is truly ‘balancing’ because each allele has a stable, non-zero equilibrium frequency (Connallon & Clark 2012). But elevated heterozygosity can also arise when selection coefficients are small or uneven; that is, when antagonistic alleles evolve under net directional selection. In this second scenario, although antagonistic alleles are expected to fix eventually,

antagonistic selection slows the process of allelic loss considerably and elevates heterozygosity relative to a polymorphism evolving under non-antagonistic directional selection (Connallon & Clark 2012).

Models show that the conditions permitting the evolution of antagonistic polymorphisms throughout the genome are highly permissive. Assuming small selection coefficients, positive intersexual trait correlations (r_{mf}) and modest intersexual fitness correlations (r_{mf}^{W})—three conditions which are likely to be widely met (Cox & Calsbeek 2009; Poissant et al. 2010)—it can be shown that most new beneficial mutations will be sexually antagonistic rather than sexually concordant (Connallon & Clark 2014b; Connallon & Clark 2014a). Thus, antagonistic polymorphisms should be common throughout the genome.

Given the above, it could be assumed that detecting molecular signatures of antagonistic selection will be straightforward. However, a couple of important caveats are worth mentioning. First, the effectiveness of antagonistic selection is relatively weak; that is, antagonistic polymorphisms are strongly sensitive to genetic drift (Connallon & Clark 2012). For example, in the case of an antagonistic allele with additive effects, the efficacy of antagonistic balancing selection is proportional to the square of selection coefficients, as $N_e(s_m^2+s_f^2)$—where s_m and s_f are sex-specific selection coefficients. By contrast, the efficacy of selection for an allele evolving under non-antagonistic directional selection is linear in s, N_es. Because selection coefficients are typically small (and their squares therefore very small), antagonistic selection requires unusually strong selection coefficients (and/or very large population sizes) to be effective. For this reason, antagonistic

polymorphisms might often evolve as effectively neutral and generate weak molecular signatures of balancing selection (Connallon & Clark 2013). Second, the elevation in heterozygosity at antagonistic loci is only necessarily observed when compared to directionally selected non-antagonistic loci (Connallon & Clark 2012). It is possible for antagonistic polymorphisms to have lower heterozygosity than neutral polymorphisms, when equilibrium allele frequencies are close to 0 or 1, where alleles are prone to stochastic loss or fixation (Mullon et al. 2012). This effect would also tend to dampen elevations in heterozygosity associated with antagonistic loci throughout the genome.

Bearing these caveats in mind, one can make a general prediction that antagonistic loci will exhibit population genetic signatures of balancing selection. For example, balancing selection causes individual polymorphisms to exhibit elevated heterozygosity because low frequencies of each respective allele are unstable. Candidate antagonistic polymorphisms should therefore exhibit elevated minor allele frequencies. In addition, the genealogies at balanced and linked neutral sites are expected to exhibit deep coalescence times; therefore, statistics—like Tajima's D—that quantify the excess in common polymorphisms resulting from long internal branches near an antagonistic polymorphism, should take on elevated values (Tajima 1989). Finally, antagonistic polymorphisms should exhibit low levels of population differentiation. This effect arises because the effective migration rate near a balanced polymorphism is increased compared to a neutrally evolving gene, so that population differentiation via drift is decelerated (Schierup et al. 2000).

But despite these simple predictions, no empirical studies have established whether antagonistic loci display signatures of balancing selection—let alone whether such patterns occur on a genome-wide scale or how they vary in space and time. While three genetic loci have previously been shown to be under sexually antagonistic selection—the *Pax7* locus in cichlid fish (Roberts et al. 2009), the *VGLL3* locus in salmon (Barson et al. 2015) and the *Cyp6g1* locus in *D. melanogaster* (Smith et al. 2011; Hawkes et al. 2016)—no broader examination of population genetic patterns at these loci was performed. Furthermore, the focus on isolated loci and the absence of genome-wide studies of antagonism have so far precluded generalisations about antagonistic selection as a general mechanism for generating balancing selection. As a consequence, the population genetic effects of genome-wide antagonistic alleles await detailed investigation.

Here I address this knowledge gap and characterise the population genetic properties of antagonistic polymorphisms across populations of *D. melanogaster*. I do so by combining two main data sources. First, I use data on the genome-wide identity of antagonistic alleles, derived from a genome-wide association study of a sexually antagonistic phenotype among 202 individuals from the laboratory-adapted LH_M population of *D. melanogaster* flies (presented in Chapter 3). Second, I use polymorphism data from three worldwide populations of *D. melanogaster* flies (part of the '*Drosophila* Genome Nexus' (Lack et al. 2016)) to quantify patterns of balancing selection. I then combine these two datasets to ask whether antagonistic polymorphisms, as identified in LH_M, are under balancing selection in independent populations.

The general approach for testing whether antagonistic sites are under balancing selection is to compare their population genetic properties—levels of heterozygosity, regional polymorphism and population differentiation—to those of non-antagonistic sites across the independent populations in the *Drosophila* Genome Nexus. If antagonistic sites are more balanced than non-antagonistic sites in a given population sample, then antagonistic selection is inferred to be acting there. This approach considers all antagonistic sites as a whole, rather than identifying specific polymorphisms under selection (Berg & Coop 2014). The approach can be considered analogous to studies which have established a link between height-associated SNPs and observed height variation in contemporary human populations (Turchin et al. 2012). Furthermore—and importantly—the analysis pipeline includes stringent steps to account for various potential confounders, such as population structure, non-random properties of genetic variation across the genome, and differences in ascertainment between antagonistic and non-antagonistic sites due to the process of GWAS discovery (see Methods).

An additional interest of this study is to establish the timescale over which antagonistic polymorphisms are selectively maintained. With this in mind, a beneficial aspect of the *Drosophila* Genome Nexus dataset is that the populations contained within it span a range of divergence times from LH_M. For instance, while the Raleigh (RAL) population sampled in North Carolina is separated from the Californian LH_M population by approximately 150 years, the Zambian (ZI) population sampled in *D. melanogaster*'s ancestral African range is separated from LH_M by over 10,000 years (Duchen

et al. 2013). By analysing the relationship between antagonism and polymorphism in populations of varied evolutionary distances from LH_M, it becomes possible to evaluate the timescale over which antagonistic selection operates. If the association between antagonistic loci and patterns of balancing selection is global, then selection is inferred to be ancient; if it is local, antagonistic selection is inferred to be transient. To extend this analysis, I also analyse polymorphism data from two sister species—*D. simulans* and *D. yakuba*—and ask whether antagonistic polymorphisms are trans-specific.

4.4. Methods

4.4.1. Candidate antagonistic SNPs and regions

To categorise single nucleotide polymorphisms (SNPs) and regions as either antagonistic or non-antagonistic, I used data from a genome-wide association study of a sexually antagonistic phenotype (the 'antagonistic index', presented in Chapter 3) across 202 *D. melanogaster* hemiclones from the laboratory-adapted LH_M population.

For analyses that rely on individual polymorphisms (see '4.4.4. SNP-based analyses of the allele frequency spectrum'), I used two approaches to compare antagonistic and non-antagonistic sites. The first was to compare the properties of candidate antagonistic SNPs (N=2,372) to non-antagonistic SNPs (N=763,392), as defined based on a False Discovery Rate (FDR) cut-off, and as described in more detail in Chapter 3. A second and complementary approach did not rely on the binary antagonistic/non-antagonistic classification. Rather, the raw P-values for each SNP were used as a continuous measure of association with the antagonistic phenotype. P-values were then correlated with measures of balancing selection, under the hypothesis that SNPs that are more significantly associated with the antagonistic index will tend to show stronger signals of balancing selection.

I also performed analyses that considered larger genomic regions (windows). Such analyses have two advantages: they allow an analysis of sequence-based (rather than SNP-based) measures of balancing selection, such as Tajima's D, and they alleviate the multiple testing burden by collapsing many correlated SNPs into a single window. This in turn permits a more stringently defined set of antagonistic windows to be defined (Q-

value<0.1). To define antagonistic windows, I implemented LDAK's 'set-based' association test, which calculates window-wide antagonistic SNP-heritability (h^2_{SNP}) via restricted maximum likelihood (options using '–calc-genes-reml', 'ignore-weights YES' and α=–0.25; see Chapter 3 for details), corrects for local relatedness using the variants in each window, and computes a P-value using a likelihood ratio test (LRT). The sets considered were 1000bp windows defined according to *D. melanogaster* Reference 5 genome coordinates, and subsequently converted to Release 6 coordinates using the liftOver tool (Hinrichs et al. 2006). This was a necessary step, as publicly available polymorphism data that formed the starting point of these analyses was mapped to Release 5 of the *D. melanogaster* genome, whereas the GWAS data was mapped to Release 6. I then calculated window-based Q-values from the LRT P-values, and defined antagonistic windows as those with Q-value<0.1 (Q-value≥0.1 for non-antagonistic windows).

4.4.2. Comparative population genomic data

To analyse population genomic data outside the LH_M population, I used publicly available whole-genome sequences from three wild *D. melanogaster* populations (part of the '*Drosophila* Genome Nexus' (Lack et al. 2015; Lack et al. 2016)). One population is non-African, sampled in Raleigh, USA (RAL; N=205) as part of the *Drosophila* Genetic Reference Panel (Mackay et al. 2012) and derived from a recent (~150 years (Duchen et al. 2013)) migration out of Africa. The two remaining populations come from *D. melanogaster*'s ancestral distribution range in sub-Saharan Africa: Zambia (ZI; N=197) and

South Africa (SA; N=118). Both populations diverged from LH_M at least 10,000 years ago (Duchen et al. 2013). The ZI population is both the largest sample of African *D. melanogaster* available and—based on its higher heterozygosity compared to other African population samples (Pool et al. 2012)—the most ancestral population of *D. melanogaster* known. The geographical location of each population and further information can be found in Lack et al. (2016).

All genome sequences were derived from high-coverage sequencing (~20X or above) of either inbred lines or haploid embryos (see www.johnpool.net/genomes.html for further details). Note also that the SA population defined here combines data from two sub-populations ('SD' and 'SP' in the *Drosophila* Genome Nexus), which have negligible levels of population differentiation (F_{ST}=–0.002) (Lack et al. 2016).

All genome sequences were downloaded as FASTA files from the *Drosophila* Genome Nexus website (www.johnpool.net/genomes.html). These files had undergone standardised alignment and quality filtering steps before being made available. I further quality-filtered for admixture and identity-by-descent by applying scripts provided on the Genome Nexus website. I then used *snp-sites* (Page et al. 2016) to call SNPs and convert the multiple sequence alignments to VCF format. Allele frequencies in each population were calculated using *vcftools* (Danecek et al. 2011), and allele frequencies further filtered by excluding tri-allelic and poorly covered sites (call rate<20) within each population.

4.4.3. Testing for sexually antagonistic balancing selection

The general approach for testing whether antagonistic sites are under balancing selection was to compare their population genetic properties to non-antagonistic sites across independent populations in the *Drosophila* Genome Nexus. I chose not to interpret differences between antagonistic and non-antagonistic SNPs in LH_M because the probability of detecting a SNP as antagonistic in the GWAS design is not independent of polymorphism. As such, it is impossible to determine whether elevated polymorphism at antagonistic sites in LH_M is the result of ascertainment bias due to GWAS discovery, or balancing selection.

Examining signatures of balancing selection in independent populations constitutes a much more robust test for balancing selection because: (i) allele frequencies in a given population become increasingly uncorrelated to LH_M as the evolutionary distance between that population and LH_M increase, thus reducing the effect of ascertainment bias; (ii) it is possible, in SNP-based analyses, to frequency-match allele frequencies between antagonistic and non-antagonistic SNPs in LH_M, which eliminates ascertainment bias entirely and permits a like-for-like comparison between the two classes of SNP. Note, however, that this 'frequency-matching' approach is somewhat conservative because it assumes that allele frequency differences between classes of SNP in LH_M is purely an artefact of statistical testing rather than a consequence of selection.

4.4.4. SNP-based analyses of the allele frequency spectrum

I looked for increased minor allele frequencies (MAFs) at antagonistic relative to non-antagonistic sites, the pattern that is expected under the

hypothesis that antagonistic sites are subject to balancing selection (Tajima 1989; Fijarczyk & Babik 2015). To do so, I first LD-pruned the LH_M dataset by clumping (in *PLINK*) to avoid pseudo-replication due to correlations between SNPs. For antagonistic sites, 226 of the 2,372 antagonistic SNPs were obtained by choosing the most significant antagonistic SNP as an index SNPs and clustering neighbouring (within 10Kb and with $r^2>0.4$) antagonistic SNPs around the index SNP. For non-antagonistic sites, an identical procedure was followed but using non-antagonistic SNPs as index SNPs. Pruning in this manner reduced the original dataset of 765,764 SNPs to 36,319 'LD-independent' SNPs.

For each of these 36,319 SNPs, I estimated MAFs in the three independent populations. I assigned MAF=0 to sites which were monomorphic in a comparison population, or sites which were polymorphic for variants other than those segregating at that site in LH_M.

I then compared MAF between LD-independent antagonistic and non-antagonistic SNPs in each comparison population. This was done using a Monte Carlo approach where, 1000 times, 226 randomly drawn non-antagonistic 'control' SNPs were frequency-matched to the 226 antagonistic SNPs. The matching procedure first corrected LH_M MAF for 'linked selection' by taking the residuals of a regression of LH_M MAF on estimates of linked selection (Elyashiv et al. 2016). Estimates of linked selection quantify local recombination rates and proximity to functional sequences in *D. melanogaster*, thereby accounting for factors that affect polymorphism along the genome, such as background selection and selective sweeps. Sets of 226 non-antagonistic SNPs were then drawn to match the residual ('linked

selection' corrected) LH_M MAFs of the 226 antagonistic SNPs and, for each set of non-antagonistic sites, the mean MAF in the comparison population was calculated. The 1000 values generated in this way provided a null distribution of mean MAFs for non-antagonistic sites in each population. Empirical P-values for deviations in polymorphism between antagonistic and non-antagonistic sites were then calculated by comparing, in each population, the mean MAF of the 226 antagonistic SNPs to the null MAF distribution.

A second analysis used the same LD-independent dataset but considered the whole spectrum of P-values, rather than a binary split of SNPs into antagonistic/non-antagonistic categories. To this end, the 36,319 SNPs were binned in two dimensions, by residual LH_M MAF (20 quantiles) and P-values (100 quantiles). One SNP was then drawn from each of these MAF/P-value bins (2,000 SNPs in total), MAF was recorded in the comparison population of interest, and MAF was correlated with P-values of the associated SNPs in LH_M using a Spearman's rank correlation. Under the hypothesis of antagonism-mediated balancing selection, SNPs with low P-values should tend to have higher MAFs in the population under consideration than SNPs with high P-values, resulting in a negative correlation between P-value and MAF.

4.4.5. Window-based analyses of the allele frequency spectrum

In addition to the SNP-based analyses, I performed genome-wide sliding windows analyses (1,000bp windows, 500bp step size) to investigate regional signatures of polymorphism. Tajima's D, which contrasts SNP

polymorphism (nucleotide diversity, π) and SNP abundance (Watterson's estimator, θ_W), was compared for windows defined as antagonistic (Q-value<0.1) or non-antagonistic (Q-value≥0.1) from the window-based association analysis. Under the hypothesis that antagonism generates balancing selection, values of Tajima's D are expected to be elevated in antagonistic windows. Values of Tajima's D were calculated for each comparison population using *PopGenome* (Pfeifer et al. 2014), using all SNPs present in a given population (regardless of whether it was present in LH_M or not). As in SNP-based analyses, estimates of linked selection (estimated in 1,000bp windows) were incorporated as controls by calculating the residuals of a regression of Tajima's D on estimates of linked selection. Since estimates of linked selection were not available for windows on the X chromosome, estimates of recombination rate were used on this chromosome instead (Comeron et al. 2012). A generalised linear model (GLM), assuming Gaussian error structure, was then used to compare residual Tajima's D between antagonistic and non-antagonistic windows.

4.4.6. Analyses of population differentiation

I tested for an association between antagonistic SNPs and signatures of reduced population differentiation. Measures such as F_{ST} are often considered problematic because they do not correct for the dependency of this statistic on local levels of polymorphism (Cruickshank & Hahn 2014). However, the availability of genome-wide estimates of linked selection in *D. melanogaster* permitted the incorporation of this confounding variable explicitly (Elyashiv et al. 2016). I therefore estimated F_{ST} over 1,000bp

windows, using *PopGenome* (Pfeifer et al. 2014), correcting F_{ST} for linked selection in a way analogous to that used for Tajima's D (see '4.4.5. Window-based analyses of the allele frequency spectrum'). I calculated residual F_{ST} for all three pairwise combination of populations. Since the distribution of F_{ST} values is not normally distributed, residual F_{ST} between antagonistic and non-antagonistic windows was statistically contrasted using Wilcoxon Rank-Sum tests.

4.4.7. Analyses of balancing selection in *D. simulans* and *D. yakuba*

I also tested whether antagonistic SNPs were associated with signatures of balancing selection in two closely related species which are part of the *melanogaster* species group: *D. simulans* and *D. yakuba*. The divergence time between *D. melanogaster* and *D. simulans* is estimated at ~1.5 million years, while *D. melanogaster* and *D. yakuba* are estimated to be ~3 million years apart (Obbard et al. 2012). To examine polymorphism in these two species, I obtained high-coverage whole-genome sequences from twenty individuals of each species (Rogers et al. 2014, http://www.molpopgen.org/). These genome sequences were originally sampled in Madagascar and Kenya (*D. simulans*), and Cameroon and Kenya (*D. yakuba*), in both species' ancestral distribution ranges. I aligned whole-genome sequences from each individual to the *D. melanogaster* Reference 5 genome using Mauve (Darling et al. 2010) and then converted *D. melanogaster* Reference 5 coordinates to Reference 6 coordinates using the liftOver tool (Hinrichs et al. 2006).

To test whether antagonistic SNPs were associated with signatures of balancing selection in each species, I performed two complementary analyses. First, I asked whether antagonistic SNPs in LH_M were more likely to be trans-specific (*i.e.*, found across species) than non-antagonistic SNPs. I modelled trans-specific status as a binomial variable (1='is trans-specific and allele identities match'; 0='is not trans-specific, or is trans-specific but allelic identities do not match'). I then performed a logistic regression with antagonistic status (Q-value<0.3) in LH_M as an independent variable, and MAF in LH_M included as a covariate. I also repeated the same analysis using raw P-value as the independent variable rather than a binary antagonistic/non-antagonistic classification. Second, I tested whether MAF in the species of interest was higher at antagonistic than non-antagonistic sites (MAF=0 was assigned to sites which were not trans-specific, or which were trans-specific but allele identities differed between species). Antagonistic/non-antagonistic MAF was compared using a Wilcoxon Rank-Sum test. I also correlated MAF with raw P-value using a Spearman's rank correlation.

In both analyses, I only considered sites which were well-covered (depth=20) in the non-*melanogaster* datasets. Additionally, owing to the small number of trans-specific polymorphisms as a whole, and thus the low power of this analysis, I used the full LH_M SNP dataset rather than an LD-pruned dataset (as done in '4.4.4. SNP-based analyses of the allele frequency spectrum').

4.4.8. Statistical software

All statistical analyses were performed in *RStudio* (RStudio Team 2015).

4.5. Results

4.5.1. Sexually antagonistic selection in a derived, non-African population

I first assessed whether antagonistic sites exhibit signatures of balancing selection in a non-African *D. melanogaster* population (RAL). Owing to the close relationship between this population and the source population LH_M, there has been little time for selection to generate noticeable differences in MAF between the two classes of SNP (antagonistic vs. non-antagonistic) when accounting for MAF between classes within LH_M. SNP-based analyses of the allele frequency spectrum therefore have relatively low power. Consistent with this, antagonistic SNPs have non-significantly elevated MAFs relative to non-antagonistic SNPs in RAL (P=0.322, Fig. 4.1A,B). Nevertheless, when considering P-values to quantify the strength of association with the sexually antagonistic phenotype, a significant negative correlation between P-values and MAF was detected in RAL (ρ=−0.06; P=0.044, Fig. 4.1C), indicating that SNPs that are more strongly associated with the antagonistic phenotype in LH_M have higher MAF in RAL, once their MAF in LH_M is accounted for.

Additionally, strong signals of balancing selection were observed when considering alternative tests. Window-based analyses showed that antagonistic windows have significantly elevated Tajima's D relative to non-antagonistic windows in RAL ($F_{1,115477}$=224.63, P<0.001; Fig. 4.1D).

4.5.2. Sexually antagonistic selection in two ancestral, African populations

I assessed whether antagonistic loci exhibit detectable signatures of balancing selection in two African populations of *D. melanogaster*, ZI and SA.

In support of the idea that antagonistic selection operates in African populations, I found that antagonistic SNPs have consistently higher MAFs than non-antagonistic SNPs across both African population samples (ZI: P=0.024, Fig. 4.2A,B; SA: P=0.001, Fig. 4.3A,B). Additionally, significant negative correlations were detected between association P-values in LH_M and MAF in both African populations (ZI: $\rho=-0.07$; P=0.001, Fig. 4.2C; SA: $\rho=-0.06$; P=0.003, Fig. 4.3C). Significant negative correlations across the whole distribution of P-values show that the significant difference in MAF between antagonistic and non-antagonistic sites does not depend on the Q-value cut-off chosen to define both classes of SNP.

SNP-level elevations in MAF were corroborated by further analyses. Regional analyses of polymorphism indicated that antagonistic windows have elevated Tajima's D in both African population samples (ZI: $F_{1,116099}=60.63$, P<0.001, Fig. 4.2D; SA: $F_{1,110954}=4.24$, P=0.039, Fig. 4.3D). Finally, all three pairwise combinations of populations exhibit lower F_{ST} at antagonistic windows relative to non-antagonistic windows (RAL-SA: W=5.10 x 10^7, P<0.001; RAL-ZI: W=5.07 x 10^7, P<0.001; SA-ZI: W=5.53 x 10^7, P<0.001; Fig. 4.4).

4.5.3. Sexually antagonistic selection in *D. simulans* and *D. yakuba*

I found no clear evidence that antagonistic SNPs are more likely to be trans-specific ($F_{1,643121}$=1.59, P=0.206; Fig 4.5A) or have higher MAFs (W=6.7 x 10^8, P=0.201, Fig 4.5B) in *D. simulans* when comparing antagonistic vs. non-antagonistic SNPs as a binary category. However, when looking at the relationship between GWAS P-value and polymorphism in *D. simulans*, I found a significant negative relationship between the P-value and both the likelihood of being trans-specific ($F_{1,643121}$=8.17, P=0.004, Fig 4.5A) and MAF in *D. simulans* (ρ=−3.5 x 10^{-3}, P=0.004). In other words, SNPs that are more closely associated with the antagonistic phenotype in LH_M are more likely to be conserved in *D. simulans* and show elevated MAF in this species.

Performing the same analyses in *D. yakuba*, I found no evidence that antagonistic SNPs are more likely to be trans-specific ($F_{1,494815}$=1.97, P=0.159; Fig 4.5C) or have higher MAFs (W=4.26 x 10^8, P=0.125, Fig 4.5C) when comparing antagonistic vs. non-antagonistic SNPs as a binary category. I also found no significant relationship between GWAS P-value and the likelihood of being trans-specific ($F_{1,643121}$=2.53, P=0.111, Fig 4.5D) or MAF (ρ=1.96 x 10^{-3}, P=0.168).

4.6. Discussion

I combined data from a genome-wide association study of sexual antagonism (Chapter 3) with publicly available polymorphism data from three worldwide populations of *D. melanogaster* and two closely related species to analyse the molecular population genetic effects of sexually antagonistic selection. I detected multiple signatures of balancing selection at antagonistic loci across populations from *D. melanogaster*'s distribution range, indicating that the heritable phenotypic variation in sex-specific fitness that can be generated by sexual antagonism is mirrored by a signal of increased polymorphism at the underlying genetic loci. These signatures of balancing selection span long evolutionary timescales (>10,000 years), and, to some extent, species boundaries, demonstrating persistent genetic effects of sexual antagonism. In the following, I discuss the implications of these results for our understanding of sexually antagonistic selection in particular and balanced polymorphisms in general.

Theory predicts that sex-specific selection and positive intersexual trait correlations (Cox & Calsbeek 2009; Poissant et al. 2010) generate permissive conditions for the evolution of sexually antagonistic polymorphisms across the genome (Connallon & Clark 2014b; Connallon & Clark 2014a). Theory also predicts that antagonistic loci will tend to have elevated heterozygosity compared to non-antagonistic loci (Connallon & Clark 2012; Mullon et al. 2012). Yet until now, no empirical link between sexually antagonistic selection and genome-wide patterns of molecular genetic variation had been established.

This study demonstrates that sexually antagonistic loci identified in a laboratory-adapted North American population of *D. melanogaster* are associated with multiple signatures of balancing selection across wild population samples from *D. melanogaster*'s worldwide distribution range. Compared to non-antagonistic loci, antagonistic loci bear the hallmarks of balancing selection: elevated minor allele frequencies, elevated regional polymorphism, and reduced population differentiation. These signatures of balancing selection are found firstly in a North American population sample (RAL). This population, like LH_M, is descended from a very recent (~150 years, (Duchen et al. 2013)) colonisation of North America from the species' ancestral African distribution range. This result indicates—perhaps unsurprisingly—that antagonistic balancing selection exerts detectable molecular genetic effects in the short time span since LH_M and RAL diverged.

More remarkably, antagonistic selection also generates clear signatures of balancing selection in populations sampled in *D. melanogaster*'s ancestral African range (ZI and SA). These ancestral populations are separated from LH_M by at least 10,000 years, or hundreds of thousands of generations (Duchen et al. 2013). As such, the population genetic effects of antagonistic balancing selection are highly evolutionarily persistent. Further illustrating this point, I found some indication that sites associated with the antagonistic phenotype in LH_M are more likely to be polymorphic in a sample of 20 individuals from *D. melanogaster*'s sister species, *D. simulans*. While the effect size is small and larger samples of individuals will be needed to confirm this association, this result potentially

indicates that sexually antagonistic selection can maintain polymorphisms across species boundaries, *i.e.* over timescales that predate the divergence time between both species ~1.5 million years ago (Obbard et al. 2012).

The relationship between antagonistic loci and signatures of balancing selection is robust to various processes that could generate similar patterns: (i) demographic differences between populations, which were controlled for by focussing on the antagonistic vs. non-antagonistic contrast within each population (in all but F_{ST}-based analyses); (ii) population structure and relatedness within LH_M, which was controlled for by incorporating a kinship matrix when performing the GWAS (Chapter 3); (iii) genome-wide variation in recombination rate and proximity to functional sequences, which was accounted for by incorporating a genome-wide map of linked selection; (iv) pseudo-replication between closely linked SNPs, which was addressed by collapsing many SNPs into a single window (in window-based analyses) or LD-pruning SNPs (in SNP-based analyses); (v) ascertainment bias due to the process of GWAS discovery, which was addressed by matching antagonistic and non-antagonistic sites by their MAFs in LH_M (in SNP-based analyses).

One exception to the consistent pattern of balancing selection found across *D. melanogaster* populations is a weak elevation of antagonistic MAF relative to non-antagonistic MAF in the RAL population. However, this weak signal can be explained by the low statistical power of this comparison, which arises because antagonistic and non-antagonistic MAFs are standardised within LH_M and because LH_M and RAL share a very recent common ancestor. These two factors mean that antagonistic selection has 'little time'

to generate differences between the two classes of site. In line with this interpretation, less conservative statistical tests (*i.e.*, analyses of Tajima's D and F_{ST}) indicate that antagonistic selection operates in RAL. Furthermore, the alternative interpretation of this result—that antagonistic selection is acting within LH_M and Africa, but not among the closely related RAL population—appears implausible.

The fact that antagonistic loci are associated with signatures of balancing selection in evolutionarily distant African populations of *D. melanogaster*—and potentially *D. simulans*—is remarkable. It indicates that adaptive conflict between males and females generates a stable constraint to the evolution of dimorphism that is unaffected by the continuous adaptation of populations to environmental conditions they encountered during their colonisation of the globe (Li & Stephan 2006; Pool et al. 2012; Duchen et al. 2013).

Several factors could help explain why sexual antagonism generates such a longstanding effect on population genetic variation. First, *D. melanogaster* populations are very large, with effective population sizes on the order of 10^6 or above (Langley et al. 2012). Previous theory indicates that the effectiveness of antagonistic selection—*i.e.*, its ability to withstand genetic drift—is low compared directional selection (Connallon & Clark 2012; Mullon et al. 2012), and that antagonistic polymorphisms should therefore generate relatively weak population genetic signatures of balancing selection (Connallon & Clark 2012; Connallon & Clark 2013). The large effective population sizes found in *D. melanogaster* likely offset this and help generate the clear patterns of balancing selection observed.

Second, strong linkage between antagonistic sites with parallel fitness effects (*i.e.*, linkage between multiple male-beneficial alleles, or linkage between multiple female-beneficial alleles) could contribute to long-term persistence of antagonism. Theory has previously shown that linkage between multiple causal sites increases the strength of selection across the antagonistic haplotype as a whole, which in turn favours its stable maintenance over time (Úbeda et al. 2011). However, this hypothesis is difficult to test, since without knowledge of the causal site(s), it is impossible to differentiate between a scenario where multiple causal antagonistic sites facilitate long-term selection or selection at a single antagonistic site generates linkage with nearby neutral sites. Additionally, long-term balancing selection is expected to reduce levels of linkage disequilibrium in nearby regions, as the long genealogies generated by balanced polymorphisms provide additional opportunities for recombination to generate haplotypic diversity and thus reduce LD (DeGiorgio et al. 2014). Adding this to the fact that LD will differ between antagonistic and non-antagonistic sites simply due to the process of GWAS discovery, patterns of LD near antagonistic sites are therefore difficult to interpret.

A third factor that could explain the long-term persistence of antagonistic polymorphisms is the interaction between sexual antagonism and fluctuating selection. A recent model has examined the relationship between the extent and speed of environmental fluctuations and the persistence of antagonistic polymorphisms (Connallon & Hall 2016), concluding that slow and continuously fluctuating environmental conditions facilitate the long-term persistence of antagonistic variation. By contrast,

rapid environmental changes tend to align the direction of selection in the two sexes and favour fixation. Ample evidence for local adaptation in *D. melanogaster* populations to seasonal (Bergland et al. 2014), altitudinal (Pool & Aquadro 2007), and latitudinal (Zhao et al. 2015) environments indicates that fluctuating selection is widespread. The pervasiveness of slow fluctuating selection in *D. melanogaster* could plausibly aid the long-term persistence of antagonistic polymorphisms.

A final important factor to consider with regard to the maintenance of antagonism is the propensity of such polymorphisms to undergo 'resolution'—*i.e.*, for each allele to be preferentially expressed in the sex it benefits. 'Resolution' is a broad term that encompasses many genetic events, some of which facilitate and some of which hamper the maintenance of antagonistic polymorphisms. On the one hand, some mechanisms of resolution could favour the maintenance of antagonistic polymorphisms and help explain their persistence. For example, it is plausible that many antagonistic polymorphisms uncovered in LH_M are situated on autosomes (see Chapter 3) because they are sex-specifically dominant (*i.e.*, the male-beneficial allele is dominant in males and the female-beneficial allele dominant in females). If so, each allele is preferentially expressed in the sex it benefits, there is no strong selection to fix either allele and a polymorphism can be indefinitely maintained. The *VGLL3* locus in salmon (Barson et al. 2015) is an example of a 'resolved' yet polymorphic antagonism. On the other hand, some mechanisms of resolution will involve the fixation of each alternative allele, as might occur if an antagonistic polymorphism is duplicated and undergoes sex-specific regulation in each paralog. In this

case, antagonistic polymorphisms will be short-lived if resolution is easy but persistent if it is more difficult. With this in mind, it is noteworthy that antagonistic polymorphisms in LH_M are disproportionately found among missense variants (Chapter 3), which can make resolution through the evolution of sex-specific regulation more difficult (for reasons that are discussed at greater length in Chapter 3). This genomic architecture could stall the fixation of alternative alleles and contribute to the long-term persistence of antagonistic polymorphisms.

The results presented here also speak to the broader question of the evolutionary forces determining levels of molecular genetic variation (Gillespie 2004; Leffler et al. 2012). By showing that antagonistic selection results in elevated genome-wide polymorphism, these analyses highlight a compelling mechanism of balancing selection and challenge the notion that balanced polymorphisms are rare (Asthana et al. 2005; Bubb et al. 2006; Hedrick 2012). Furthermore, the full consequences of sexually antagonistic selection on molecular population genetic variation are likely underestimated in this study because the antagonistic loci segregating in LH_M only represent a small subset of all antagonistic loci segregating in wild populations. The bottleneck that occurred during *D. melanogaster*'s migration to North America (Li & Stephan 2006; Duchen et al. 2013) will have caused the loss of many African antagonistic polymorphisms, while drift during LH_M's laboratory maintenance will have caused the loss of many North American antagonistic polymorphisms. These 'lost' polymorphisms should also contribute to the overall balanced genetic variation in wild ancestral populations but are not screened here.

The discrepancy between the paucity of balanced loci detected in selection scans and the strong evidence for balanced loci presented here is not surprising. Selection scans focus on the strongest and clearest instances of balancing selection in order to exclude other processes that could generate similar molecular signatures; as such, they are very conservative (Andrés et al. 2009; Leffler et al. 2013; Fijarczyk & Babik 2015). By contrast, the approach taken here has more power because it leverages a priori data on the identity of candidate loci. It can therefore detect weak signatures of balancing selection (or even relaxed directional selection) at candidate loci relative to non-candidate loci, which would otherwise be missed in traditional selection scans. In other words, balancing selection may not be so much rare as simply difficult to detect. Further illustrating this point, a recent study by Bergland et al. (2014) used an F_{ST}-outlier approach to identify 'seasonal polymorphisms' among North American *D. melanogaster*. Analogous to the approach taken here, the authors linked the location of these polymorphisms with signatures of balancing selection across populations (and species) and demonstrated that fluctuating selection causes long-term elevations in genome-wide genetic variation.

In summary, this study provides the first demonstration that balancing selection generated by sexual antagonism elevates genome-wide polymorphism. This work motivates future research to uncover the factors that aid the long-term maintenance of antagonistic polymorphisms. This work also highlights the need for improved methods for detecting balanced polymorphisms in selection scans. Promising new methods have recently been developed (DeGiorgio et al. 2014; Siewert & Voight 2017; Bitarello et

al. 2018) and these should clarify the full extent to which balancing selection influences genome-wide genetic variation.

4.7. Figures

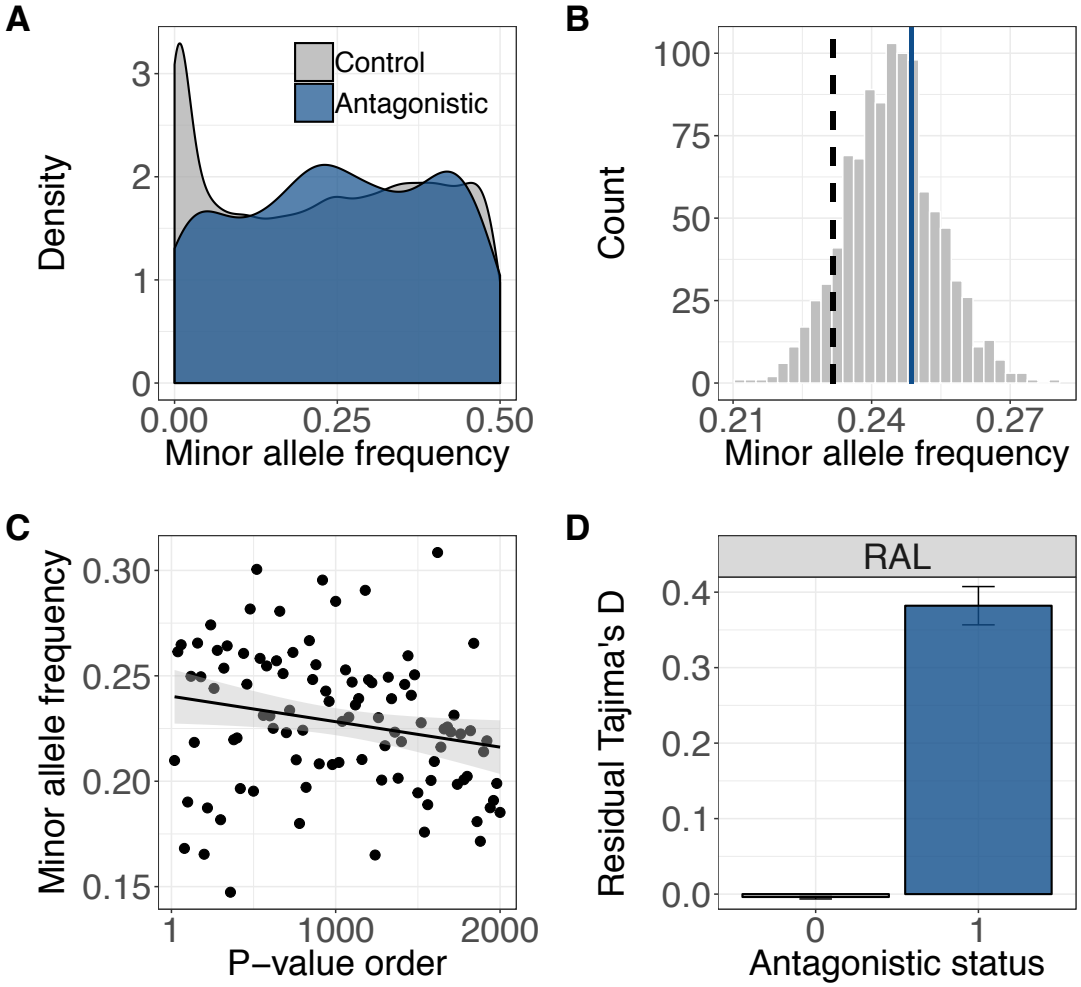

A
Control
Antagonistic
Density
Minor allele frequency
B
Count
Minor allele frequency
C
Minor allele frequency
P−value order
D
RAL
Residual Tajima's D
Antagonistic status

Figure 4.1. Signatures of balancing selection associated with antagonistic loci in RAL. A. Raw MAF spectrum for LD-pruned antagonistic (blue) and non-antagonistic ('control', grey) SNPs. **B.** Distribution of mean MAFs for 1,000 sets of LD-independent, non-antagonistic SNPs that have been frequency-matched to LH_M antagonistic SNPs (see Methods). Blue line denotes mean MAF of antagonistic SNPs; black dashed line denotes mean MAF of non-antagonistic SNPs before frequency-matching. **C.** MAF across 100 sets of LD-independent SNPs, each set matched for LH_M allele frequencies, and presented in ascending order by P-value. For visualisation purposes, a linear regression line (±95% CI) is shown. **D.** Mean (±S.E.) residual Tajima's D (corrected for linked selection, see Methods) for antagonistic windows (blue; 'antagonistic status=1') and non-antagonistic windows (grey; 'antagonistic status=0').

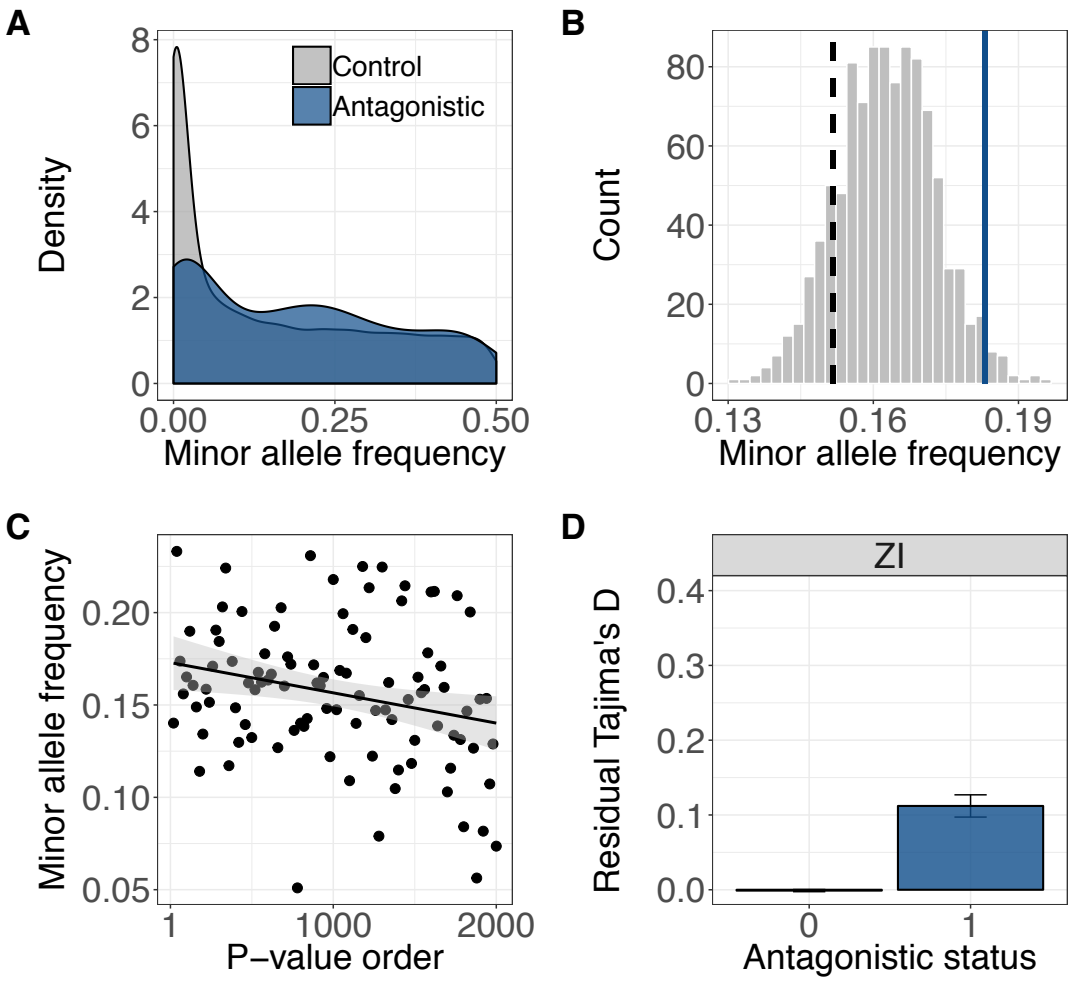


Figure 4.2. Signatures of balancing selection associated with antagonistic loci in ZI. This figure mirrors the analyses presented in Figure 4.1A-D but applies them to polymorphism data from the ZI population. See Figure 4.1A-D caption for detailed descriptions of each sub-Figure.

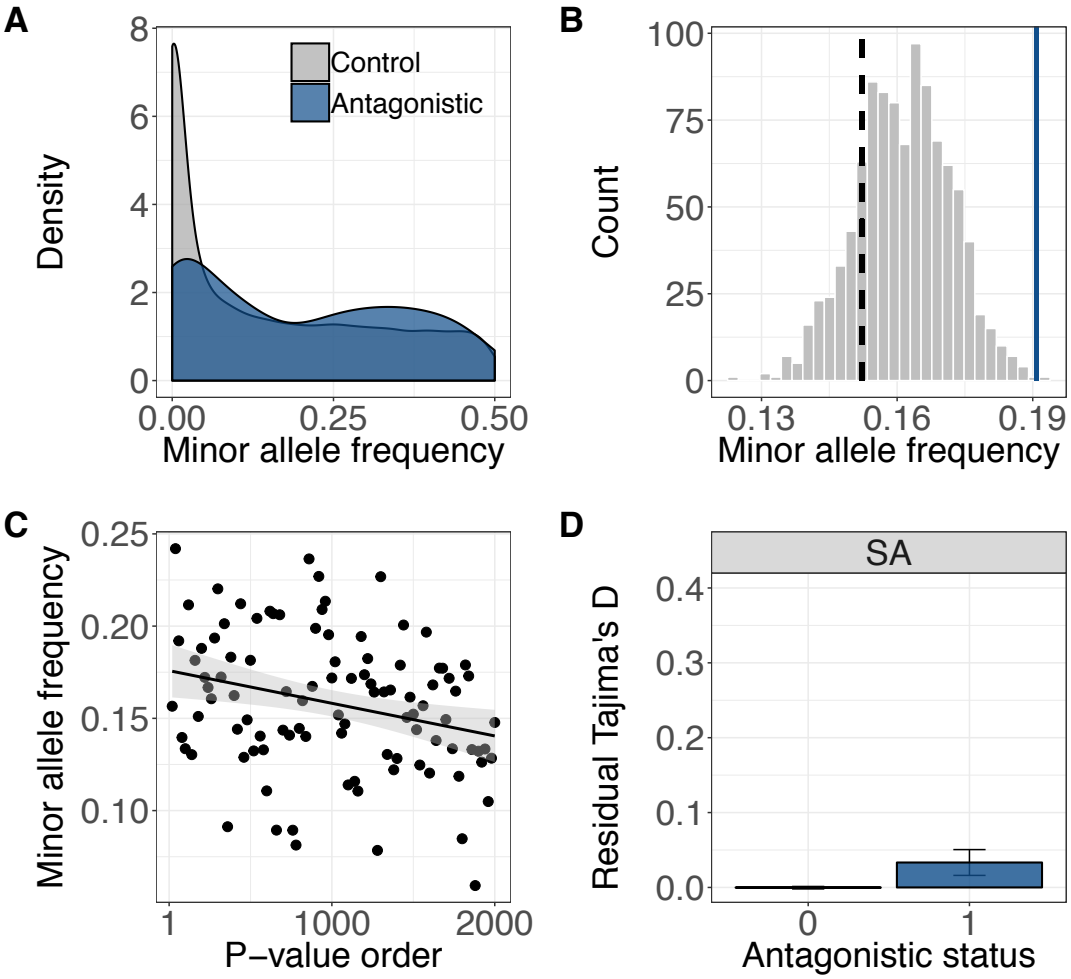


Figure 4.3. Signatures of balancing selection associated with antagonistic loci in SA. This figure mirrors the analyses presented in Figure 4.1A-D but applies them to polymorphism data from the SA population. See Figure 4.1A-D caption for detailed descriptions of each sub-Figure.

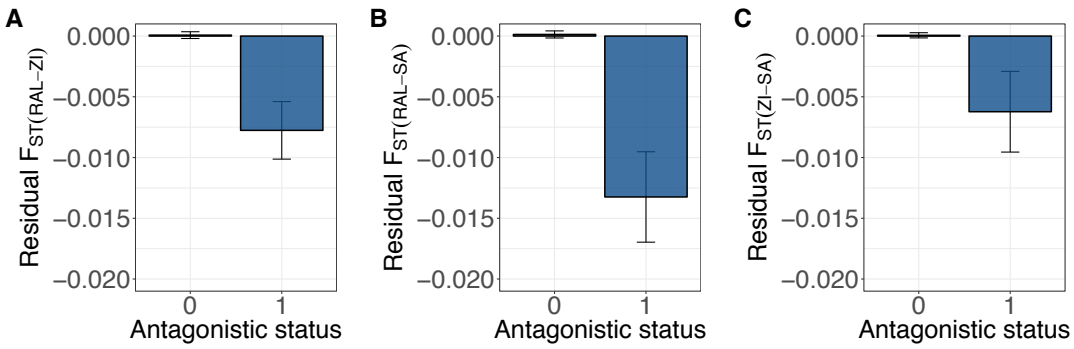


Figure 4.4. Population differentiation among antagonistic and non-antagonistic windows. Residual F_{ST} (±S.E.), corrected for linked selection (see Methods), for antagonistic and non-antagonistic windows across all three pairwise combinations of populations.

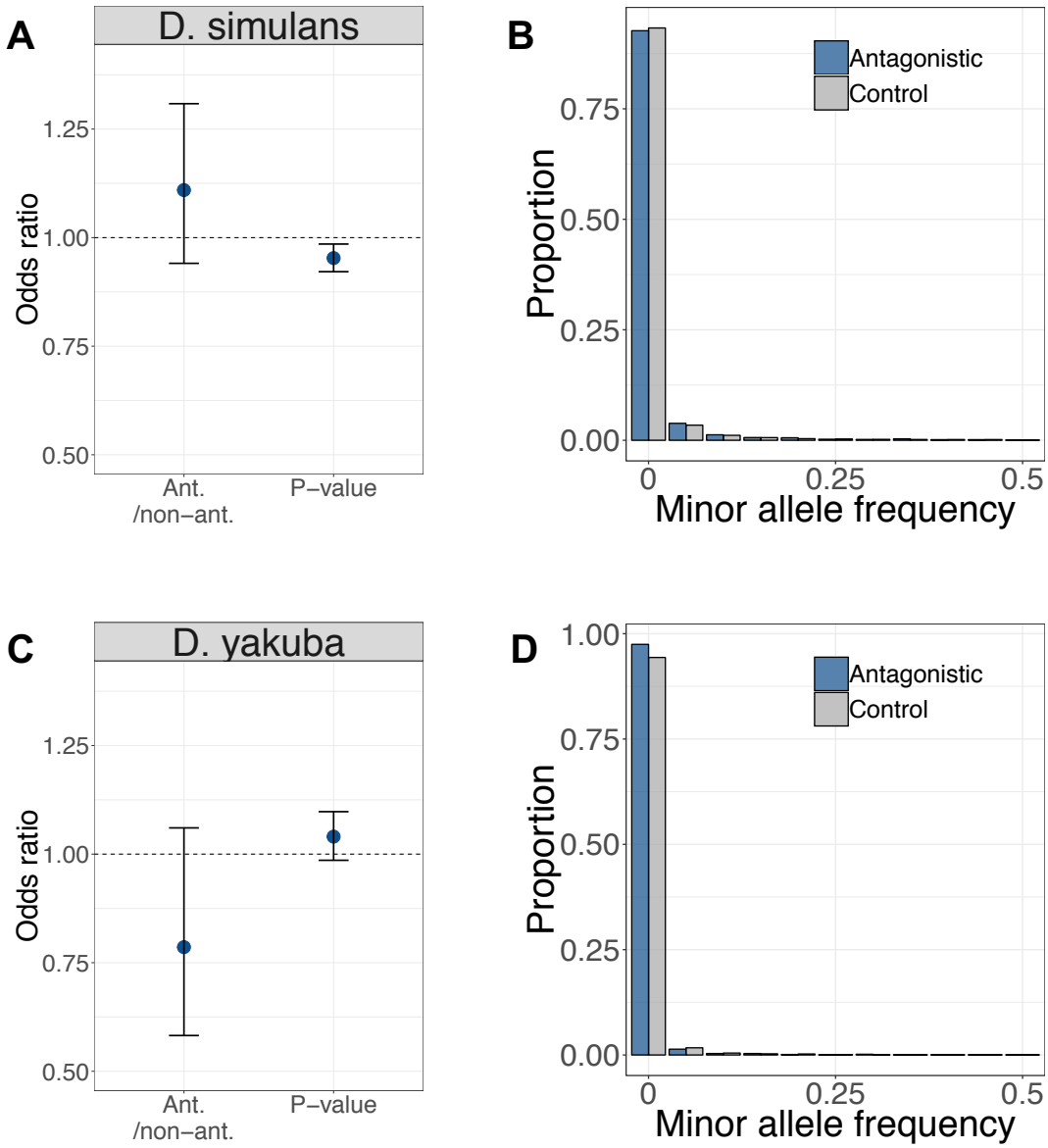

A
D. simulans
Odds ratio
1.25
1.00
0.75
0.50
Ant.
/non−ant.
P−value
B
Antagonistic
Control
Proportion
0.75
0.50
0.25
0.00
0
0.25
0.5
Minor allele frequency
C
D. yakuba
Odds ratio
1.25
1.00
0.75
0.50
Ant.
/non−ant.
P−value
D
Antagonistic
Control
Proportion
1.00
0.75
0.50
0.25
0.00
0
0.25
0.5
Minor allele frequency

Figure 4.5. Signatures of balancing selection associated with antagonistic loci in *D. simulans* and *D. yakuba*. A. (*Left*) Odds ratio (±95% CI) that antagonistic SNPs are polymorphic and share the same allelic variants in LH_M as in *D. simulans* relative to non-antagonistic SNPs. Odds ratio>1 indicates that antagonistic SNPs are more likely to be trans-specific than non-antagonistic SNPs. (*Right*) Odds ratio (±95% CI) that SNPs in LH_M are trans-specific when considering the whole range of P-values. Odds ratio<1 indicates that SNPs with lower P-values are more likely to be trans-specific. In both analyses, MAF in LH_M is included as a control in the logistic regression. **B.** Minor allele frequency spectra for antagonistic and non-antagonistic SNPs across twenty *D. simulans* genomes. **C.** Same as A. but applied to *D. yakuba* polymorphism data. **D.** Same as B. but applied to *D. yakuba* polymorphism data.

Chapter 5

Experimental evidence for balancing selection at a polymorphism in the *D. melanogaster fruitless* gene

5.1. Declaration

All work reported in this chapter is my own, except ‘5.4.1. The genetic structure of the *fru* polymorphism’ (Mark Hill, Max Houghton). Throughout this study, I have benefited from invaluable contributions from students. Harvinder Pawar and Olivia Davidson both contributed to ‘5.4.3. Frequency dynamics of *fru* alleles’. Didem Snaith contributed to exploratory experiments for ‘5.4.6. Fitness assays for isogenic allelic lines’ and Olivia Davidson performed most of the fitness work presented in ‘5.4.6. Fitness assays for isogenic allelic lines’. Max Reuter and Mark Hill produced Fig. 5.1A. I also thank Stephen Goodwin for providing the deficiency stock used in sections ‘5.4.5. Creation of isogenic allelic lines’ and ‘5.4.6. Fitness assays for isogenic allelic lines’.

5.2. Abstract

Observational data is often used to establish that balancing selection is occurring at given genetic loci. However, this type of data cannot establish that specific alleles causally affect fitness; to do so, experimental approaches must show that alleles at a locus of interest exhibit hypothesised fitness effects, while controlling for other sources of variation. Here, using experimental approaches, I consider whether a polymorphic region of a candidate gene for sexual antagonism, *fruitless* (*fru*), is evolving under sexually antagonistic selection. Exploratory bioinformatic analyses first reveal that variation in this gene is found across populations from the *D. melanogaster* distribution range, and in sister species *D. simulans*. To test whether this signal of balancing selection can be recovered experimentally, I track the frequency dynamics of alternative *fru* alleles in *D. melanogaster* cage populations initiated with skewed starting frequencies. This analysis reveals that *fru* alleles consistently evolve towards an intermediate frequency, as expected under balancing selection. I then test whether the mechanism of balancing selection is sexual antagonism by creating fly lines that carry each respective *fru* allele but are homozygous along the rest of their genome. This second analysis provides equivocal results: although I detect a significant *fru* allelic effect on female fitness, no significant effect is detectable for male fitness measurements. I discuss why this is so and suggest improvements for future experiments. Overall, this work provides a rare experimental test of balancing selection and paves the way for similar and improved experimental studies in the future.

5.3. Introduction

The selective maintenance of alleles—balancing selection—can play an important role in explaining levels of genome-wide polymorphism within populations (see Chapter 4). Balancing selection is known to occur through a variety of mechanisms, including overdominance (Hedrick 2012), frequency-dependence (Vekemans & Slatkin 1994), or opposing selection between environments, traits and sexes (Levene 1953; Kidwell et al. 1977; Rose 1982). The effects of balancing selection can be persistent, as made clear by the fact that its molecular genetic signatures are often detectable across populations and species (Leffler et al. 2013; Bergland et al. 2014).

Nevertheless, evidence for balancing selection is often observational rather than experimental. For example, the detection of ~2,000 candidate sexually antagonistic single nucleotide polymorphisms (SNPs) in LH_M (Chapter 3) is based on a correlation between sex-specific fitness effects and allelic variation. Similarly, the inference of sexually antagonistic selection across *D. melanogaster* populations (Chapter 4) is based on a correlation between the genomic location of candidate antagonistic polymorphisms in LH_M and levels of polymorphism across populations. Although these correlations are likely to reflect the signal of true antagonistic balancing selection, they cannot establish that a particular pair of alleles *causally* affects sex-specific fitness. Ultimately, it is necessary to perform experiments to demonstrate causality and validate inferences about the adaptive value of alternative alleles.

To achieve meaningful verification, it is necessary to perform experimental manipulations which isolate the effects of particular

polymorphisms on fitness and control other sources of variation. For example, to test for balancing selection, experimental approaches must show that specific alleles are maintained at a stable, intermediate equilibrium frequencies, while also ensuring that this effect either does not or cannot (owing to the nature of the experimental design) apply to 'control' alleles. Similarly, to test for sexually antagonistic selection, it is necessary to show that each allele has opposing fitness effects in each sex while ensuring that genetic variation at other loci is homogeneous. Such experimental approaches are essential, as they provide the strongest evidence for adaptation at the genetic level.

Although experimental verification is an essential step in studying the genetics of adaptation (Pardo-Diaz et al. 2015; Wilkinson et al. 2015), relatively few candidate polymorphisms for balancing selection have so far been experimentally tested for the fitness effects of their alleles (Hedrick 2012). In the specific case of sexual antagonism, only two polymorphisms have been experimentally shown to exhibit antagonistic fitness effects. One such polymorphism is the *Pax7* locus in cichlid fish, where the authors found evidence for increased *Pax7* expression among fish exhibiting the 'Orange Blotch' phenotype (Roberts et al. 2009). As the phenotype was known from previous research to benefit females but harm males, the authors were able to demonstrate that expression changes associated with the *Pax7* polymorphism are antagonistic. The second locus to be experimentally associated with antagonistic fitness effects is a transposable element polymorphism which confers DDT resistance and is situated in the *Cyp6g1* gene in *D. melanogaster* (Smith et al. 2011; Hawkes et al. 2016). Through

fitness assays of flies carrying the ‘resistant’ and ‘susceptible’ alleles, the authors of both studies were able to show that the resistant allele was beneficial to females but detrimental to males, while the susceptible allele had the opposite fitness effect. In this way, the authors demonstrated that the *Cyp6g1*-associated polymorphism is sexually antagonistic.

Given the dearth of experimental studies of individual antagonistic loci, it is important that further research is undertaken to characterise the fitness effects of specific candidate loci. One way of choosing a locus to study is the ‘candidate gene’ approach, where previous research on the properties of a specific gene is used to motivate experimental work. This approach has been successful in showing that a number of loci are subject to adaptive evolution (Martin & Orgogozo 2013).

A promising candidate gene for sexually antagonistic selection is *fruitless* (*fru*). *fru* is a transcription factor that plays a key role in sexual differentiation in *Drosophila melanogaster* and across insect groups (Gailey et al. 2006). In *D. melanogaster*, *fru* has two major functions, one of which is essential for viability in both sexes and another that helps specify a sexually dimorphic nervous system (Villella & Hall 2008). This latter, sex-specific function is achieved through alternative splicing of male- and female-specific isoforms. Experiments have shown that male-specific *fru* isoforms are expressed in the male central nervous system, helping to promote male sexual behaviour (Neville et al. 2014), whereas female-specific isoforms do not produce any functional protein (Usui-Aoki et al. 2000).

There are several reasons why *fru* is a promising candidate gene for sexual antagonism. First, the fact that *fru* governs neuronal dimorphism in *D.*

melanogaster indicates that it is under sex-specific selection. As sex-specific selection is a pre-requisite for the evolution of sexual antagonism, this property could indicate that *fru* harbours antagonistic polymorphisms. The idea that *fru* is subject to sex-specific selection is further supported by the observation that this gene is associated with interspecific differences in courtship behaviour (Lagisz et al. 2012), and that signatures of positive selection have been detected across species in regions of the gene involved in sex-specific splicing (Parker et al. 2014).

The fact that *fru* is an important regulator of sexual differentiation also means that it affects the expression of many genes (Neville et al. 2014); this high degree of pleiotropy could also favour the evolution of antagonistic polymorphisms. The rationale for this is that the evolution of sex-specific expression in pleiotropic genes is difficult to achieve without negative side-effects (Mank et al. 2008). This closes off an important avenue for the resolution of antagonistic polymorphisms, and potentially favours their accumulation in pleiotropic genes such as *fru*.

In addition to these suggestive properties, there is also evidence from exploratory analyses that balancing selection is occurring in this gene, as is expected under sexual antagonism. Specifically, Hill (2017) found that several clusters of SNPs in *fru* (henceforth the '*fru* polymorphism') were fixed between hemiclonal haplotypes with male-beneficial, female detrimental (MB) and female-beneficial, male-detrimental (FB) fitness effects. Although these MB and FB lines share unusual population histories which preclude the confident interpretation of alternative *fru* alleles in terms of antagonistic selection (see Chapter 2), the fact that SNP variation is detectable at all in

such an essential and pleiotropic gene (where variation is mostly expected to evolve under purifying selection) is highly unusual, and suggestive of balancing selection. Further exploratory work also uncovered that the fixed SNP differences between MB and FB lines identified by Hill (2017) are situated in the close vicinity of a ~50bp indel (Mark Hill, pers. comm.). The presence of a structural variant in such an essential and pleiotropic gene further indicates that this region of *fru* is likely to affect phenotypic variation, and thus perhaps fitness as well.

Motivated by the interesting genetic properties of *fru* and its suggestive associations with sexual antagonistic fitness effects, I experimentally test two major questions: (i) is the *fru* polymorphism under balancing selection, and (ii) is the *fru* polymorphism under sexually antagonistic selection?

To first establish whether the *fru* polymorphism is under balancing selection, I create replicate cage populations of *D. melanogaster* flies carrying extreme starting frequencies of each *fru* allele in an otherwise randomly variable genetic background. I then track the frequency of each allele over time and test the hypothesis that allele frequencies converge toward an intermediate, 'balanced' state. Second, I proceed to test whether each *fru* allele has opposing fitness effects in each sex, as expected under sexually antagonistic selection. To do so, I create replicate fly lines which are genetically homogeneous but carry different *fru* alleles. I then measure their sex-specific fitness and look for opposing allelic effects in each sex. Overall, this empirical study provides a rare experimental evaluation of a candidate antagonistic polymorphism in action.

5.4. Methods

5.4.1. The genetic structure of the *fru* polymorphism

Using a set of extreme hemiclonal haplotypes (Chapter 2), Hill (2017) found that a number of SNPs situated in the *fru* gene were fixed between FB and MB lines, and thus putatively associated with sexually antagonistic fitness effects (Fig. 5.1A). I henceforth refer to cluster (iv) in Fig. 5.1A as the '*fru* polymorphism' and alternative alleles across the cluster of SNPs as the '*fru* alleles', although both terms encompass more than one SNP.

Further exploratory work was then performed to investigate the genetic structure of *fru*. In this work, 96 hemiclonal lines were extracted from LH_M as hemiclones (see '1.5.3. Hemiclonal analysis'). Individual hemiclonal males were crossed with females from a deficiency strain (Df(3R)BSC509), which carries a deletion spanning the *fru* gene and a balancer complement carrying a dominant marker *Stubble* (*Sb*). DNA from the hemiclone/deletion heterozygote offspring of this cross was extracted, the *fru* polymorphism PCR amplified and Sanger sequenced. Sanger sequencing revealed that the *fru* polymorphism is situated in very close proximity (<20bp) to a 54bp-long indel. As this indel produces different fragment lengths when a PCR product is amplified and run under gel electrophoresis (Fig. 5.1B), the two alleles were termed the 'Long' (L) and 'Short' (S) alleles.

5.4.2. Molecular signatures of balancing selection at the *fru* polymorphism

I considered whether the *fru* polymorphism was also found in two wild population samples of *D. melanogaster* flies: a North American population

sample of 205 genomes (RAL) and a Zambian population sample of 197 genomes (ZI) (Mackay et al. 2012; Lack et al. 2015). To visualise the frequency of the *fru* polymorphism across these populations, I constructed a haplotype network using the SNPs found in the *fru* polymorphism. I also used data on the whole-genome sequences of 20 *D. simulans* lines (Rogers et al. 2014) to look for evidence that this cluster of SNPs is detectable across species. While these analyses do not aim to provide firm evidence of balancing selection at the *fru* polymorphism, they can provide suggestive evidence in support of the subsequent experimental studies described below.

5.4.3. Fly culture and husbandry

Unless otherwise stated, culture conditions were as follows. Flies were maintained in 25°C constant temperature rooms at 50% humidity on a 12:12hr light-dark cycle. Flies were housed in yeasted vials containing 8mL corn-agar-molasses medium, or sometimes (when stated) in yeasted bottles containing 50mL medium. When required, flies were collected as virgins, every 0-6 hours post-eclosion until sufficient numbers were collected. Virgin flies were anaesthetised using a CO_2 pad for short periods of time and manipulated using a fly aspirator.

5.4.4. Frequency dynamics of *fru* alleles

I initiated replicate cage populations with extreme ratios (9:1; 1:9) of each allele at the *fru* polymorphism and tracked their frequencies over time.

The first step in this experiment was to generate two populations fixed for each respective *fru* allele (‘L’ and ‘S’) but variable elsewhere along the

genome. Each population was established from the *Drosophila* Genetics Reference Panel (RAL (Mackay et al. 2012)), a collection of whole-genome sequenced inbred lines. Two sets of 14 lines were selected from RAL that carried either the L allele or the S allele (all obtained from the Bloomington *Drosophila* Stock Centre). To make up the L population, the 14 inbred lines carrying the L allele were crossed in a round-robin design (♀1 x ♂2, ♀2 x ♂3, ♀3 x ♂4, etc.) on corn-agar-molasses medium supplemented with tetracycline (0.25mg/mL) to cure *Wolbachia* infection. 25 virgin offspring of each sex were collected from each cross (N=700 in total), mixed and placed across 3 yeasted culture bottles. The same procedure was applied to the 14 inbred lines carrying the S allele to establish the S population. The two populations thus produced were variable throughout their genome but fixed for each alternative allele at the *fru* polymorphism.

Bottle culture lasted for 3 generations, and to maximise genetic mixing within each respective population during bottle culture, flies were collected as virgins at each generation from each of the 3 bottles, mixed, and placed in new yeasted bottles (N~700 across three bottles).

Once L and S populations were created, the next step was to track the frequency dynamics of the L and S alleles. To do so, I established five population cages for each of two experimental manipulations: (i) 'High-S' cages, which contained 450 virgin flies (225 males 225 females) from the S population and 50 virgin flies (25 males, 25 females) from the L population (S:L=9:1); (ii) 'Low-S' cages, which contained 50 virgin flies (25 males, 25 females) from the S population and 450 virgin flies (225 males 225 females) from the L population and (S:L=1:9). Each cage was initially supplied with 3

yeasted culture bottles containing 50mL of corn-agar-molasses medium. Subsequently, each cage was supplemented with 3 additional culture bottles on a weekly basis until each cage contained twelve culture bottles (in the fourth week). From this point on, each week the oldest 3 culture bottles were replaced with fresh culture bottles, such that each cage always contained 12 culture bottles.

Cage populations were set up in July 2017. In order to track *fru* allele frequencies in each cage, approximately 100 flies were sampled from each cage population, approximately 4 and 6 months (~8 and 12 generations) after cages were set up, in November 2017 and January 2018, respectively. DNA was extracted from 96 individuals from each cage at each time point using a standard DNA extraction protocol. A ~500bp DNA fragment encompassing the *fru* polymorphism with its linked 54bp indel was amplified using PCR. The genotype of each individual could be visually determined from the length of the PCR products using agarose gel electrophoresis (Figure 5.1B). Allele frequencies were then estimated from genotypic frequencies.

Note that due to time constraints, the January 2018 sample did not contain genotype information for 4 out of the 10 cages (2 High-S and 2 Low-S cages).

5.4.5. Creation of isogenic allelic lines

In order to assess the sex-specific fitness effects of the L and S alleles, I created fly lines homozygous for either one or the other allele but isogenic across the rest of their genome (‘isogenic allelic lines’). This process involved

three phases (see Figure 5.2. for a schematic representation). In Phase 1, hemiclonal lines from the LH_M population were extracted and genotyped for the allele of interest (L or S). In Phase 2, an isogenic background was repeatedly introgressed into the hemiclonal line while ensuring the transmission the original hemiclonal allele (L or S) through the use of a deficiency stock. In Phase 3, introgression lines (which were heterozygous for the allele of interest and the deficiency during the introgression process) were made homozygous for the allele of interest. Each step is described in more detail below.

Phase 1 involved the extraction of 96 hemiclonal lines from LH_M (these lines were the same as those used in '5.4.1. The genetic structure of the *fru* polymorphism' and the genotyping process is described in more detail there). From the 96 lines, three lines carrying the L allele and three lines carrying the S allele were kept for further analysis and maintained as hemiclones using standard hemiclonal amplification (see '1.5.3. Hemiclonal analysis').

In Phase 2, the LH_M genetic background of the six allelic lines was replaced with an isogenic Canton-S background. To achieve this, each of the six hemiclonal lines was repeatedly crossed with a strain that carries a deficiency spanning the *fru* polymorphism ($Df(3R)fru^{4\text{-}40}$) in an isogenic Canton-S background. The third chromosomes of this strain, which carries the deficiency, is balanced with TM6, a standard balancer chromosome marked with the dominant mutation *Tubby* (*Tb*). Since balancer and deletion homozygotes are lethal in homozygous state and balancer chromosomes are marked, the offspring of a cross between a hemiclonal female and a $Df(3R)fru^{4\text{-}40}$/TM6 male are always identifiable as hemiclone/deletion

heterozygotes. By repeatedly backcrossing hemiclone/deletion heterozygotes females to Df(3R)fru$^{4\text{-}40}$/TM6 males, the original hemiclonal genome is gradually eroded through recombination in females and replaced with the isogenic Canton-S background of the deficiency line. After 7 generations of introgression, the allelic lines should carry on average less than 1% of the original hemiclonal haplotype.

The final step in the creation of allelic isogenic lines (Phase 3) was to create lines which were homozygous for the *fru* polymorphism. To achieve this, a two-step crossing procedure was performed. In the initial cross, the virgin balancer offspring of an introgression line x Df(3R)fru$^{4\text{-}40}$/TM6 cross were set up in pairs (dyads A, B, C, see Fig. 5.2). Depending on the genotype of the introgression line parent, this cross can produce non-*Tb* F1 offspring of one of two genotypes, either introgression line homozygotes (the genotype interest to be kept) or introgression line/deficiency heterozygotes (to be discarded). Since these two offspring genotypes are indistinguishable phenotypically (neither the *fru* allele of interest nor the deficiency are associated with a phenotypic marker), an additional 'test cross' was performed where both F1 offspring were backcrossed to Df(3R)fru$^{4\text{-}40}$/TM6 males. Based on the F2 phenotype, the genotype of the F1 could be inferred, as homozygous introgression line F1 parents produce a 1:1 ratio of wild-type to *Tb* offspring, whereas introgression line/deficiency heterozygotes produce 1:2 ratio of wild-type to *Tb* offspring. Introgression line homozygote F1s which produced a ratio of wild-type to *Tb* that was significantly different from 1:2 (assessed from a χ^2 test) were kept for fitness assays.

5.4.6. Fitness assays for isogenic allelic lines

I performed experiments to assess male and female fitness of each isogenic allelic line. Throughout the fitness assays, flies were housed in constant temperature rooms at 25°C and 80% humidity (to maximise larval viability). Flies of the Df(3R)fru$^{4\text{-}40}$ deficiency strain were used as competitors in both male and female assays. Flies whose fitness was measured (referred to as 'focal flies') were reared in vials; competitor flies were reared in bottles due to the large numbers required. Environmental variation across all fitness assays was minimised by standardising density by controlling the number of eggs placed in the rearing vials and bottles. For focals, density-control was performed by manually transferring eggs laid on agar medium into a fresh vial (50 eggs per vial); for competitors, density-control was performed by manually transferring eggs laid on corn-agar-molasses medium into each bottle (~250 eggs per bottle).

Male fitness was measured as competitive fertilisation success. Virgins were collected on days 12-14 after egg laying, and aged 3-5 days prior to use in fitness assays. Vials were then set up containing a single virgin focal male from an isogenic allelic line, a virgin Df(3R)fru$^{4\text{-}40}$/TM6 competitor male and a virgin Df(3R)fru$^{4\text{-}40}$/TM6 competitor female. After 90 minutes—a period of time which maximises the chance of a mating happening while also ensuring that multiple mating is unlikely to occur—both males were removed and the competitor female left to oviposit for a further 48 hours. After 14 days, paternity was assigned by examining offspring for the presence of wild-type pupae (which would indicate that the focal male had gained a mating) and pupae with the *Tb* phenotype (which would indicate competitor

paternity). Assays were performed in a blocked design across two blocks. The sample size across all lines and blocks was N=462.

Female fitness was measured as competitive fecundity. Virgins were collected on days 12-14 after egg laying, and aged 3-5 days prior to use in fitness assays. Vials containing 5mg of live yeast were set up with 8 virgin females from each isogenic allelic line, 8 Df(3R)fru$^{4\text{-}40}$/TM6 competitor females and 16 Df(3R)fru$^{4\text{-}40}$/TM6 competitor males. After 48 hours, males were removed, focal females separated into two groups of four and housed in 'primary oviposition' vials. Here they were left to oviposit on clear agar medium for 18 hours. Females were then moved, in their respective groups, onto 'secondary oviposition' vials containing standard medium for a further 24 hours. Assays were performed in a blocked design across three blocks. The sample size across all lines and blocks was N=156 for egg counts and N=131 for offspring counts.

Eggs from the primary oviposition vials were photographed using webcamSeriesCapture (github.com/groakat/webcamSeriesCapture) software and counted manually. These egg counts provided a first measure of female fitness. Eggs from the secondary oviposition vials were left to develop and the eclosed offspring were frozen and counted after 14 days. These offspring counts provided a measure of female fitness that integrates egg production and egg-to-adult viability. Vials where a focal female died during oviposition phases were excluded from further analysis.

5.4.7. Statistical analyses

The frequency of the S allele in cage populations was modelled using a binomial GLMM, with time, manipulation (High-S or Low-S) and their interaction treated as fixed effects and cage treated as a random effect.

For male fitness, differences between lines were first assessed by fitting a binomial GLM with line and block as fixed effects. To assess differences between genotypes, male competitive fitness was fitted using a binomial GLMM with genotype and block as fixed effects, and line as a random effect.

For female fitness, egg and offspring counts of focal females were fitted using a quasipoisson GLM, with line and block as fixed effects. The effect of genotype on female egg and offspring production was fitted using a negative binomial GLMM, with genotype and block as fixed effects, and line as a random effect.

General Linear Mixed Models (GLMMs) were implemented in the *lme4* package. Overdispersion was assessed by comparing the residual deviance to residual degrees of freedom and looking for deviations from a 1:1 ratio. When such deviations were detected (>1.1:1 ratio), corrections for this effect were applied—either using quasibinomial and quasipoisson error structures (in GLMs), or by using a negative binomial rather than Poisson error structure (in GLMMs). Significance of all models was assessed using χ^2 tests. All statistical analyses were performed in *RStudio* (RStudio Team 2015).

5.5. Results

5.5.1. Signatures of balancing selection at the *fru* polymorphism

Looking at the haplotype structure of the *fru* polymorphism identified by Hill (2017) in RAL and ZI populations, I found that haplotypes do not cluster by population. Instead, alternative alleles of the *fru* polymorphism form the two major haplotypes present across both populations (Fig. 5.1C). In other words, each alternative allele is found at intermediate frequencies in both populations. Given the large evolutionary distances between the North American hemiclonal haplotypes used by Hill (2017) to identify *fru* polymorphisms, and the ZI population used in the construction of the haplotype network, this is suggestive evidence that the *fru* polymorphism is under balancing selection. Furthermore, patterns of polymorphism data in *D. simulans* reveal that one SNP in the *fru* polymorphism is trans-specific, supplementing the notion that the *fru* polymorphism is a promising candidate locus for balancing selection. As these patterns are only suggestive, I performed further experiments to test this hypothesis explicitly.

5.5.2. Frequency dynamics of the *fru* polymorphism

I tested whether *fru* polymorphism is under balancing selection by tracking the frequency of the S allele over time in populations initiated at either high initial frequencies of the S allele (‘High-S’, 90%) or low initial frequencies of S (‘Low-S’, 10%). Under the hypothesis that the *fru* polymorphism is evolving under balancing selection, the frequency of the S allele should converge towards an intermediate state. In line with this, the ‘High-S’ populations see a consistent decrease in S allele frequency over time (average of 70.5% after 4

months; 71.3% after 6 months), while the 'Low-S' populations see a consistent increase in S allele frequency (29.3% after 4 months; 23.9% after 6 months) (Fig. 5.3). In line with these patterns, a significant negative time-by-manipulation interaction is observed (χ_1^2=222.32, P<0.001). A significant positive time effect is also observed (χ_1^2=84.43, P<0.001)

5.5.3. Effect of *fru* alleles on sex-specific fitness

I measured the fitness of fly lines that were isogenic across their genome but carried different alleles at the *fru* locus.

For male competitive fertilisation success (Fig. 5.4A), I found that focal males were generally less successful at fertilising females than the competitors (mean=22.5% matings obtained across all lines). Although different lines had significantly different male fitness (χ_5^2=28.10, P<0.001), there was no significant difference effect of the identity of the *fru* allele (S or L) on male fitness (χ_1^2=1.82, P=0.177).

For female fitness, I found that egg production and offspring production differed significantly between lines (egg count: χ_5^2=317.81, P<0.001; offspring count: χ_5^2=555.93, P<0.001). Egg production was significantly higher among lines carrying the S genotype (mean=14.1 eggs) compared to lines carrying the L genotype (mean=7.1 eggs) (χ_1^2=10.47, P=0.001, 5.4B). This pattern was not repeated when considering offspring production, where S- and L-carrying lines did not exhibit significantly different offspring counts (χ_1^2=1.14, P=0.286, 5.4C).

5.6. Discussion

Few examples of experimental validation of individual candidate balanced loci exist (Pardo-Diaz et al. 2015; Wilkinson et al. 2015). Here, I assessed the frequency dynamics of alleles at the *fru* polymorphism and showed that its frequency evolves towards an intermediate state when initiated at highly skewed starting frequencies. While this is consistent with balancing selection, measurements of sex-specific fitness did not strongly support sexual antagonism as the mechanism. I discuss the strength of evidence for antagonistic balancing selection at *fru* and suggest improvements for future work.

The wealth of recent population genomic data has motivated the search for specific loci under balancing selection (Andrés et al. 2009; Leffler et al. 2013; Bergland et al. 2014; Siewert & Voight 2017; Bitarello et al. 2018). These studies have successfully identified a number of balanced loci. However, relatively few candidate loci for balancing selection have been experimentally validated (Hedrick 2012). In the specific case of sexually antagonistic selection, only two loci, *Cyp6g1* (Smith et al. 2011; Hawkes et al. 2016) and *Pax7* (Roberts et al. 2009), have been experimentally validated. Even in the case of *Pax7*, it was only established that this locus causally affected trait variation, not fitness variation.

In this work, I build on a study which showed that a polymorphism in the *fru* gene was potentially associated with antagonistic fitness effects (Hill 2017). Through population genomic analyses, I showed that the *fru* polymorphism is found in populations across the *D. melanogaster* distribution range—and even in sister species *D. simulans*. While the analyses

presented are not exhaustive, the suggestion of balancing selection at *fru* is particularly intriguing given the otherwise remarkable degree of conservation of this gene (Gailey et al. 2006) across insect groups (>250 Million years), and given previous indication that the gene is primarily under purifying selection and thus exhibits low genetic diversity (Parker et al. 2014).

I tested whether this long-term signal of balancing selection is causal, by tracking the frequencies of *fru* alleles in cage populations set up with skewed starting frequencies. Over ~8-12 generations, I found that the frequency of the S allele across both manipulations increases significantly—*i.e.*, there is a significantly positive time effect. However, I also found evidence for a significantly negative time-by-manipulation interaction, indicating that while the S allele increases in population cages initiated at low starting frequencies (1:9), it decreases in population cages initiated at high starting frequencies (9:1).

The negative time-by-manipulation effect is difficult to explain through any mechanism other than balancing selection. For instance, if *fru* was evolving under directional selection, the beneficial allele (either L or S) would be expected to increase in frequency and the detrimental allele to decrease in frequency, irrespective of the starting frequency of each allele. In other words, one would expect to see a significant time effect but a non-significant time-by-manipulation interaction. The absence of a parallel response in both manipulation is inconsistent with directional selection at *fru* itself.

The absence of a parallel response is also inconsistent with the idea that *fru* itself is neutral but variation linked to *fru* is under directional selection. This might be expected to occur if, for example, the L or S populations used

to make up cage populations had somewhat different fitness values. In this scenario, allele frequencies among the fitter initial population (L or S) would systematically increase in frequency and generate long-range LD with the (neutral) *fru* allele, thus 'dragging' it to higher frequencies.

Alternatively, if the *fru* polymorphism were evolving neutrally (and linked genomic regions were equally fit), then the frequency of *fru* alleles should vary stochastically around the initial 9:1 or 1:9 ratio. Again, the systematic change in allele frequencies towards an intermediate state is inconsistent with neutral evolution.

While the results presented here strongly suggest that balancing selection causally maintains variation at the *fru* locus, the data cannot altogether exclude the possibility that balancing selection is acting on a different balanced polymorphism for which the *fru* polymorphism acts as linked marker. This scenario is plausible, given that the L and S populations that were used to make up the initial cage populations were created from a small number of inbred lines with little time for recombination before the initiation of the cage experiment—a set-up that has the potential to generate long-range LD. Nevertheless, while a number of balanced loci will potentially be in long-range LD with *fru*, it is unlikely that these linked markers will also respond in the same way as *fru* because their starting frequencies are unlikely to be similarly skewed at the outset of the cage experiment. For example, one could imagine that *fru* in strong LD with another balanced locus, but that the starting frequencies of the alleles at this other locus are 10% in both L and S populations. In this case, the frequency dynamics of the *fru* allele which are measured in this experiment tell us little about the

frequency dynamics of this alternative balanced allele. Overall, the frequency dynamics that are detectable in this experiment are highly likely to represent the balanced frequency dynamics of *fru* itself.

Having established that balancing selection is occurring at *fru*, I measured the sex-specific fitness of replicate isogenic lines carrying different *fru* alleles to test whether sexually antagonistic selection is the mechanism of balancing selection. The data here are equivocal. On the one hand, sex-specific fitness measurements showed that *fru* alleles have a significant effect on female egg production, with elevated egg counts among the three replicate lines carrying the S allele relative to the L allele. Furthermore, the S allele also generally decreased male fitness—as might be expected under sexual antagonism. On the other hand, the difference between alleles in male fitness is quantitatively small and not statistically significant. Furthermore, the difference between alleles in female egg production is not mirrored in another measure of female fitness, offspring production.

A number of factors could explain the mixed evidence for sexual antagonism at the *fru* polymorphism. First, there is a large amount of variance between different isogenic lines, as made clear from the fact that in all assays, isogenic lines have significantly different sex-specific fitness. Although differences between isogenic lines are expected if the *fru* polymorphism causes differences in fitness, such differences are also observable between lines carrying the same *fru* allele. This indicates that some residual between-line variance present in the progenitor wild-type genome remains present. Although lines were introgressed with an isogenic stock for 7 generations, which predicts that less than 1% of the original wild-

type genome will remain present, there will be some stochastic variation in the realised proportions around this expectation, due to the random segregation of and recombination in the hemiclone/deletion heterozygotes during the introgression process. The fitness differences between lines may be therefore be due to fitness differences between the original wild-type genome that remains after introgression. For example, some isogenic allelic lines may carry more genetic variants with recessive deleterious effects in the introgressed region than other lines, which would explain the fitness difference between lines.

The expression of deleterious recessive alleles from the original wild-type genome may also help explain why the isogenic allelic lines are poor competitors in male fitness assays—despite the fact that they are competed with males from the isogenic line that the wild-type genome has been replaced with. Additionally, the expression of deleterious recessives could help explain the apparent discrepancy between the two measurements of female fitness. Since egg production measures only the fitness of focal female, while offspring production represents a combined measure of focal female fitness and the fitness of the chromosomal complement present in the offspring, this chromosomal complement could differentially mask some of the deleterious recessives present in the focal female, thus generating differences between egg and offspring counts.

A second potentially important factor in explaining the mixed evidence for sexual antagonism is epistasis. Interactions between the original wild-type genome and the isogenic line genome would tend to add variance between lines and potentially mask the fitness effects of the *fru* polymorphism.

Alternatively, epistatic interactions between the *fru* polymorphism itself and the remaining wild-type genome could contribute to between-line variance and the masking of *fru* effects. Interestingly, some exploratory analyses (Snaith et al., unpublished) have found evidence that intact hemiclonal genomes carrying different *fru* alleles show large differences in fitness when complemented with two different isogenic backgrounds, potentially suggesting a role for epistatic interactions at the *fru* polymorphism.

A final possibility is that the results of the fitness assays can be accounted for if the *fru* polymorphism is not evolving under sexual antagonism but instead under another mechanism of balancing selection, such as overdominance or frequency-dependence. This would then help to reconcile the strong evidence for balancing selection and the equivocal evidence for sexual antagonism occurring at the *fru* polymorphism.

The results presented in this study motivate future work to characterise selection patterns and fitness effects at the *fru* polymorphism. For example, it will be important to confirm that the frequency dynamics of *fru* alleles are related to the effects of *fru* and not a different, linked allele. One way to test this is to examine frequency dynamics of other polymorphisms along the genome, which can act as controls. For example, one could perform pooled sequencing of whole-genomes from individuals at each time point and then track the frequency of *fru* in the context of genome-wide genetic variants.

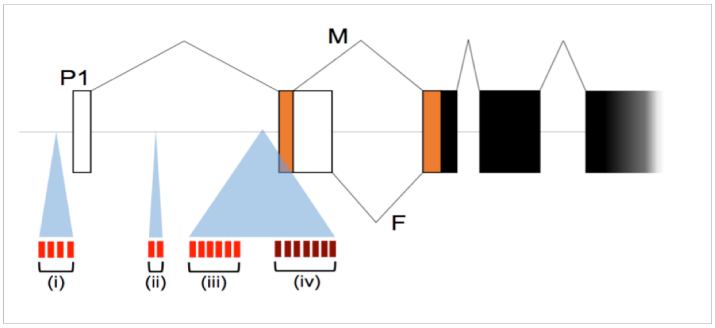
Another important avenue for future research is to improve upon the fitness assays presented here. For example, additional generations of introgression could be performed and thus circumvent the issue of between-line variance that is due to residual variation from the original wild-type

genome. Alternatively, the *fru* isogenic lines could be complemented with genetic variation from the wild-type LH_M population. Although this design would tend to increase within-line variance (due to the outbred nature of the LH_M stock population), it would help to mask some of the deleterious recessive effects that inflate between-line variance in this study.

In summary, this study provides experimental insights into the evolutionary dynamics of a candidate antagonistic polymorphism found in a highly pleiotropic and essential gene. These results constitute a rare experimental test of adaptation at the genetic level and pave the way for future work aimed at characterising the genetic basis of sexual antagonism.

5.7. Figures

A



B

C

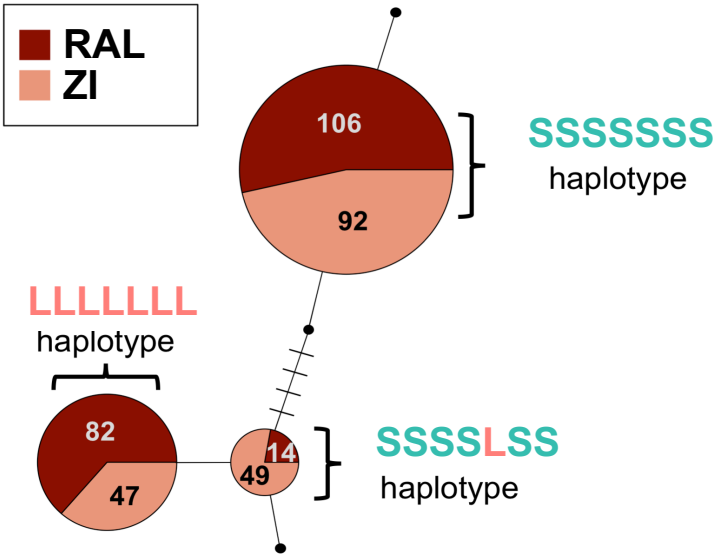


Figure 5.1. The *fru* polymorphism: gene structure and signatures of balancing selection. A. Schematic representation of the 5'-end of the *fruitless* gene model. Exons are shown as boxes (non-coding exons in white). Sex-specific isoforms of FRU are produced from the P1 promoter. Male-specific isoforms are produced via splicing labelled 'M' and include exons (highlighted in orange) that encode a sex-specific protein domain which is lacking in the female specific splice-form (F). Candidate SNPs (red bars) from Hill (2017) are arranged in four clusters (i-iv). Clusters (i), (ii), (iii), and (iv), span 61bp, 16bp, 46bp, and 57bp respectively. The latter three clusters are situated close to the male-specific exon. The SNPs of cluster (iv, dark red) were used to construct the *fruitless* haplotype network shown in C. **B.** Gel electrophoresis of the *fru* polymorphism, showing the visible differences between LL and SS homozygous genotypes that are due to a 54bp indel situated within 50bp of cluster (iv). **C.** Haplotype network for SNPs of cluster (iv) in A. Each circle represents a unique haplotype and is annotated with its frequency in each population; small black circles represent very infrequent haplotypes (0.5-2% of individuals). Notches indicate mutational steps between each haplotype.

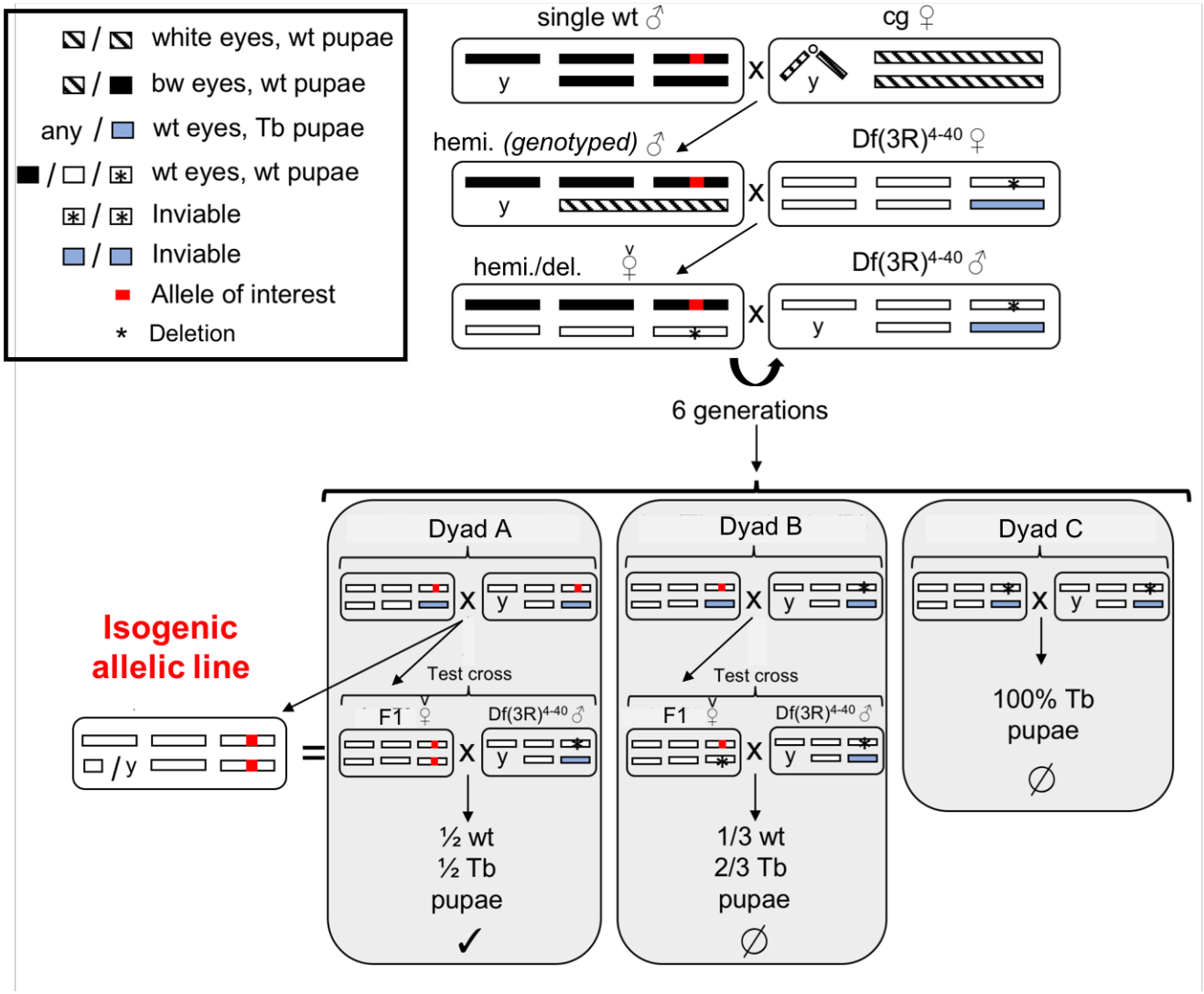

white eyes, wt pupae
bw eyes, wt pupae
wt eyes, Tb pupae
wt eyes, wt pupae
Inviable
Inviable
Allele of interest
Deletion
single wt ♂
cg ♀
hemi. (genotyped) ♂
Df(3R)4-40 ♀
hemi./del.
Df(3R)4-40 ♂
6 generations
Phase 1
Phase 2
Phase 3
Isogenic allelic line
Dyad A
Dyad B
Dyad C
Test cross
F1
Df(3R)4-40 ♂
½ wt ½ Tb pupae
1/3 wt 2/3 Tb pupae
100% Tb pupae

Figure 5.2. Crossing scheme for isogenic allelic lines. The scheme is described in more detail in section ‘5.4.5. Creation of isogenic allelic lines’. Colours represent different types of chromosomes: balancer (blue), wild-type LH_M (black), clone-generator (striped black) and non-balancer deficiency (white). Wt=wild-type; hemi.=hemiclone; del.=deletion; cg=clone-generator.

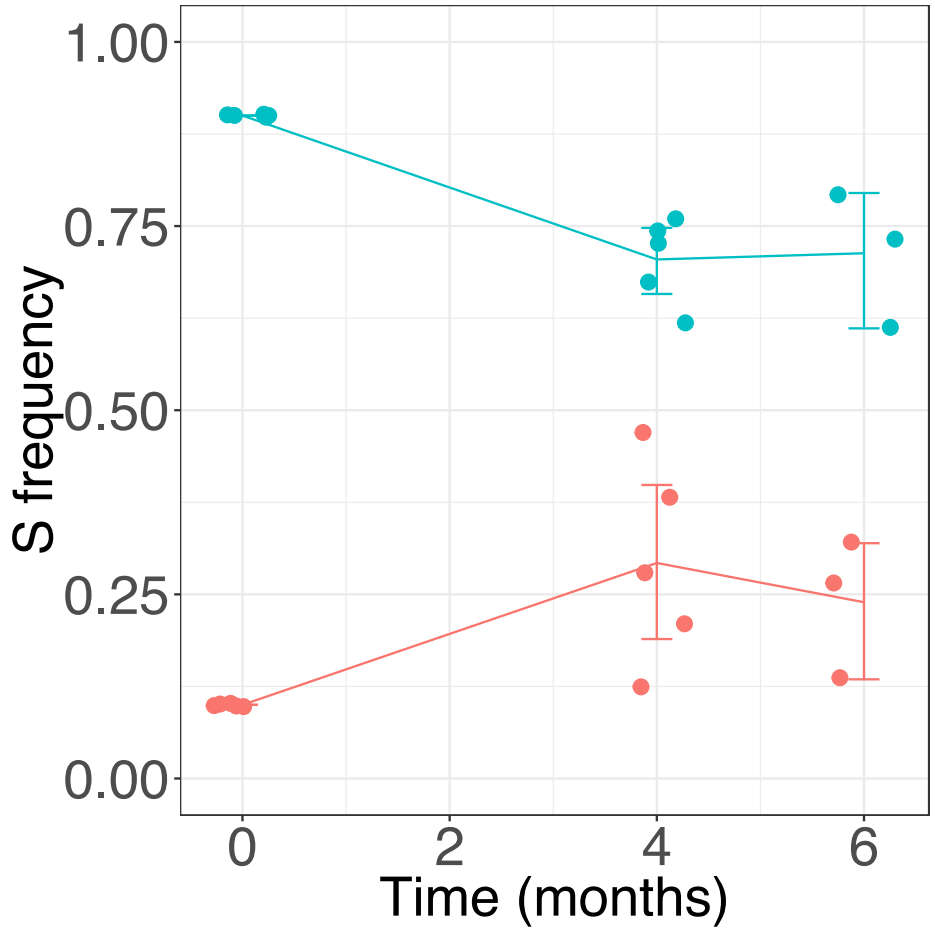


Figure 5.3. Frequency dynamics of the *fru* polymorphism. Each dot represents the frequency of the S allele estimated from genotyping of individual flies (time=4; time=6) or known exactly (time=0). Colour denotes manipulation (blue='High-S' manipulation; red='Low-S' manipulation). Bars represent means for each manipulation/time combination, with bootstrapped standard errors.

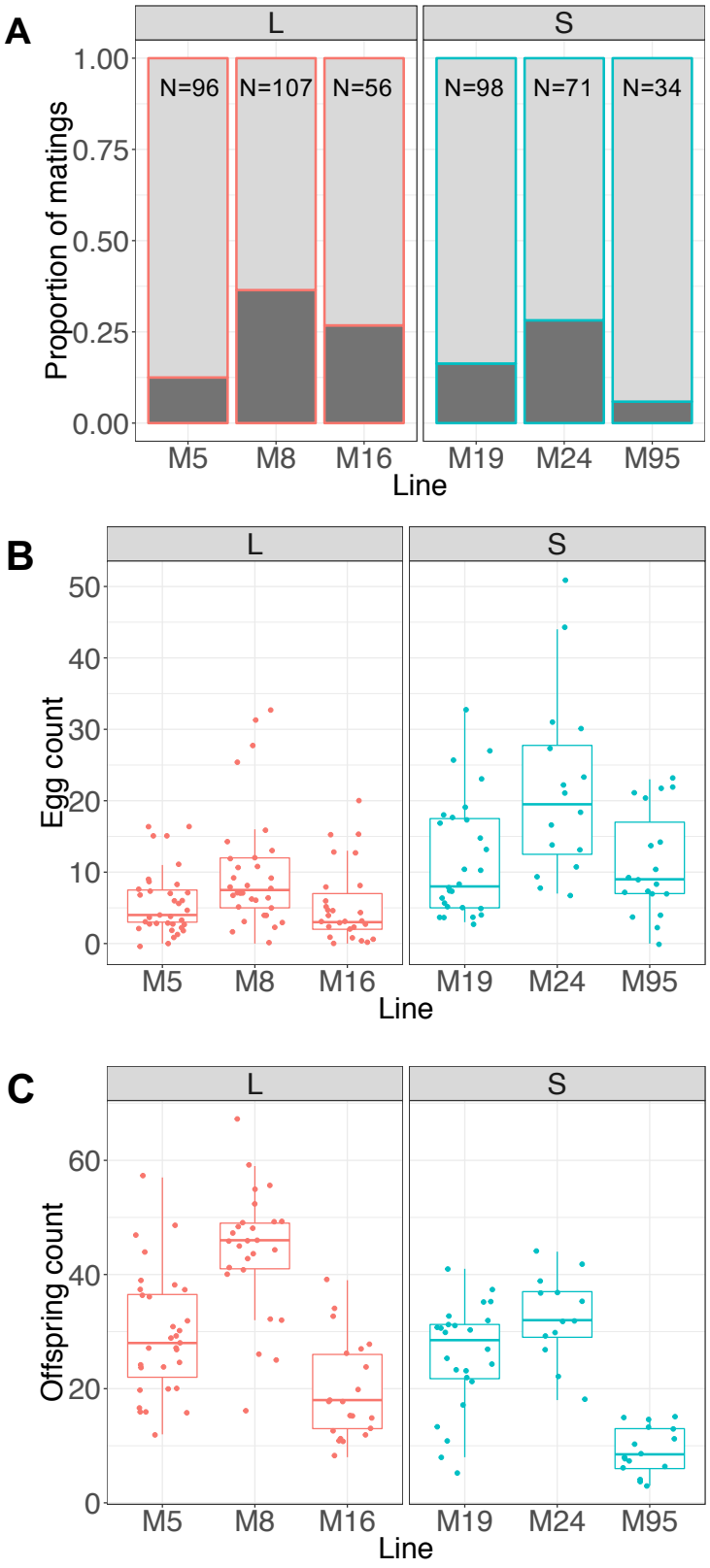

A
L
S
1.00
0.75
0.50
0.25
0.00
Proportion of matings
N=96
N=107
N=56
N=98
N=71
N=34
M5
M8
M16
M19
M24
M95
Line
B
L
S
50
40
30
20
10
0
Egg count
M5
M8
M16
M19
M24
M95
Line
C
L
S
60
40
20
0
Offspring count
M5
M8
M16
M19
M24
M95
Line

Figure 5.4. Sex-specific fitness among six isogenic allelic lines. In all plots, colours indicate the *fru* genotype of each line (red=LL homozygotes, blue=SS homozygotes). **A.** Male fitness, measured as the proportion of matings acquired by focal males (dark grey) and competitor males (light grey). **B.** Focal female fecundity, measured as the number of eggs laid by a group of four focal females on agar medium over 18hrs. **C.** Focal female fecundity, measured as the number of offspring produced by a group of four focal females laying on standard medium for 24hrs.

Chapter 6

General Discussion

6.1. Overview

Sex-specific selection favours the evolution of divergent phenotypes in males and females (Andersson 1994; Arnqvist & Rowe 2005). However, the shared genome generates positive genetic correlations between homologous traits in each sex that can constrain sex-specific adaptation (Lande 1980). At the genetic level, this constraint manifests itself as segregating alleles with opposing fitness effects in each sex—sexually antagonistic polymorphisms (Bonduriansky & Chenoweth 2009; Van Doorn 2009). Previous theory has made a number of predictions regarding the identity and functional properties of antagonistic polymorphisms (Rice 1984; Stewart et al. 2010) and their effects on quantitative and population genetic fitness variation (Connallon 2010; Connallon & Clark 2012). Until now, however, empirical investigations into the genetic basis of sexual antagonism have been limited, and analyses of its evolutionary dynamics are non-existent.

In this thesis, I have addressed a number of these knowledge gaps. Here, I begin by briefly summarising the main results in each chapter. I then proceed to discuss these findings in the broader context of adaptation across correlated environments, where I draw parallels between sexual antagonism and other forms of antagonistic pleiotropy. I end by suggesting promising avenues for future research that build upon the results presented here.

6.2. Summary of main results

In the first data chapter (Chapter 2), I illustrated the limitations of our current understanding of sexual antagonism by re-analysing phenotypic and genotypic data from two previous studies (Innocenti & Morrow 2010; Hill

2017). I first showed that male and female fitness data from 100 *D. melanogaster* hemiclonal lines sampled from the laboratory-adapted LH_M population (Innocenti & Morrow 2010) cluster phenotypically into two 'sets' ('H-lines' and 'P-lines') with significantly different male and female fitness values. The sets correspond to two independent events of notionally random sampling of genotypes from LH_M. I then examined genetic differentiation between sets by comparing sequence data from a subset of sequenced H- and P-lines with male-beneficial, female-detrimental (MB) and female-beneficial male-detrimental (FB) fitness effects respectively. The data strongly supports the idea that these two sets cluster into distinct genetic clades, thus providing an explanation for the observed differences in fitness between sets. I proceeded to show that non-random sampling among hemiclonal lines adversely affects previous phenotypic and genetic inferences. First, correcting the sex-specific fitness data for the effect of non-random sampling no longer supports a negative intersexual genetic correlation for fitness (r_{mf}^{W}) in LH_M, as had been previously reported (Innocenti & Morrow 2010). Second, the non-independence of clade and fitness effect among sequenced hemiclones substantially inflates the false discovery rate (from ~25% to ~85%) among previously identified candidate antagonistic polymorphisms (Hill 2017). Hence the asserted interpretation of candidate antagonistically expressed genes (Innocenti & Morrow 2010) is also likely compromised.

Having highlighted deficiencies among previous genetic studies in Chapter 2, I proceeded to remedy this in Chapter 3. Here, I performed a genome-wide association study of sex-specific fitness and sexual

antagonism, using fitness and whole-genome sequence data from ~200 hemiclonal lines. I found high heritabilites for sex-specific fitness and sexual antagonism, but no evidence for any large-effect loci. The numerous independent clusters of antagonistic single nucleotide polymorphisms (SNPs) suggest that antagonistic fitness variation is highly polygenic. There was no notable enrichment in terms of gene functions, but I found a significant association between antagonistic polymorphisms and missense SNP variants, implicating conflict over optimal sex-specific proteins as a key barrier to the evolution of sexual dimorphism. Contrary to classic theory, the X chromosome was not enriched for candidate antagonistic polymorphisms. I assessed the causes of high sex-specific fitness variance in general and showed that fitness variation in LH_M is explained by the joint contributions of polymorphisms with sexually antagonistic and sexually concordant effects. The relatively even contributions of both types of variation result in an r_{mf}^{W} that is not significantly different from zero.

In Chapter 4, I went on to test a fundamental prediction from theoretical studies of sexual antagonism: that antagonistic selection should elevate levels of polymorphism due to opposing selection pressures on individual alleles in each sex. I tested this by comparing signatures of balancing selection between antagonistic and non-antagonistic SNPs across three large population samples spanning *D. melanogaster*'s distribution range. Doing so, I found that antagonistic polymorphisms exhibit the hallmarks of balancing selection: elevated minor allele frequencies, elevated regional polymorphism, and reduced population differentiation. These patterns of elevated polymorphism were found in populations that share common

ancestry with LH_M dating to at least 10,000 years, or hundreds of thousands of generations. There is even some indication that antagonistic selection elevates polymorphism in sister species *D. simulans*, which is separated from *D. melanogaster* by ~1.5 million years. These analyses demonstrate that sexually antagonistic selection not only generates signatures of balancing selection, but also that the effects of antagonistic selection are highly evolutionarily persistent in time and space.

In Chapter 5, I presented experimental work aimed at testing whether a candidate polymorphism in the *fruitless* (*fru*) gene is (i) under balancing selection, and (ii) under sexually antagonistic selection. By creating *D. melanogaster* cage populations with extreme initial frequencies of each of two *fru* alleles and tracking allele frequencies over time, I was able to show that alleles converge towards an intermediate frequency, as expected under balancing selection. By creating genetic lines to isolate the effect of each *fru* allele on sex-specific fitness, I tested whether a given *fru* allele has sex-specific fitness effects consistent with the ongoing presence of sexual antagonism. I found a significant effect of the *fru* allele on female but not male fitness. Even though the results are inconclusive, this study paves the way for future work aiming to experimentally validate the fitness effects of specific candidate polymorphisms for sexual antagonism.

6.3. Sexual antagonism in the context of correlated evolution

In '1.1. Genetic constraints on the evolution of sexual dimorphism', I made an analogy between sexual antagonism, where opposing selection pressures in each sex are constrained by a shared genome, and the difficulty of adapting

to different local environments in the presence of migration between environments. This analogy has been made previously (Levene 1953; Kidwell et al. 1977) and illustrates that research into sexual antagonism sits in the broader context of research into 'antagonistic pleiotropy'—*i.e.*, opposing selection on allelic variants across environments, sexes or traits. The results presented in this thesis are therefore relevant to the general theme of correlated evolution. To place my findings in perspective, I now discuss the findings presented in this thesis in this broader context.

An emerging theme from studies of antagonistic pleiotropy (in the broad sense) is that the action of opposing selection pressures has a pervasive effect on patterns of genomic variation. For instance, in Chapter 3, I showed that there are 2,327 SNPs (226 LD-independent clusters) associated, at modest false discovery rate, with antagonistic fitness effects. Similarly, Bergland et al. (2014) identified 1750 sites associated (using the same false discovery rate threshold) with seasonal variation in *D. melanogaster*. The results of Chapter 3 are in line with previous theory indicating that there are ample opportunities for sexually antagonistic polymorphisms to arise throughout the genome (Connallon & Clark 2014a; Connallon & Clark 2014b). Similarly, Bergland et al.'s (2014) results have been corroborated by recent theory, where it has been shown that the conditions permitting the evolution of polymorphisms under fluctuating selection are similarly permissive (Wittmann et al. 2017).

An additional emerging pattern is that antagonistically pleiotropic polymorphisms are evolutionarily persistent. I showed in Chapter 4 that sexually antagonistic loci bear the hallmarks of balancing selection—

elevated heterozygosity, reduced population differentiation and elevated linkage disequilibrium—across three large population samples from *D. melanogaster*. The samples cover large geographical distances across the species' distribution range as well as significant temporal depth, as they are separated from LH_M by over 10,000 years. Similar to the signature of antagonistic balancing selection, Bergland et al. (2014) found that signatures of elevated polymorphism associated with seasonally fluctuating (and hence seasonally selected) SNPs are detectable across populations and species of fruit flies. Although there exists a set of genetic loci associated with pleiotropy between traits in humans (Pickrell et al. 2016), no study has yet linked the identity of these pleiotropic polymorphisms with signatures of antagonistic balancing selection.

Thus, in general, it appears that the genetic effects of antagonistic pleiotropy are both pervasive and evolutionarily persistent. Why is this so? Fundamentally, the answer depends on the ease with which correlations between traits, environments and sexes can be broken down, and thus alleviate the root cause of antagonistic pleiotropy. New research, including results presented in this thesis, shed some light on the factors that allow (or prevent) such resolution of adaptive conflict to occur.

First, an important way of resolving antagonistic pleiotropy is for the genome to acquire sex-specific, trait-specific or environment-specific gene regulation—*i.e.* to evolve phenotypic plasticity. The importance of gene regulatory changes in overcoming antagonistically pleiotropic constraints has been highlighted in a recent study of a bacterial system (*E. coli*). In this system, Yi et al. (2016) examined the genetic substitutions that allowed

bacterial populations to overcome a trade-off between two traits (growth and chemotaxis) and found that a nonsynonymous change in the *FLiA* transcription factor permitted this trade-off to be alleviated (see Reuter et al. (2017) and Appendix B). However, while evolving different levels of expression of a particular transcript in different contexts may be relatively straightforward, evolving different transcripts altogether—*i.e.*, overcoming conflict over coding variation—may be more difficult, as it potentially requires duplication followed by sex-specific expression of each paralog (Stewart et al. 2010; Connallon & Clark 2011b). With this in mind, it is interesting to note that both Bergland et al. (2014), the GWAS of the antagonism index presented in Chapter 3 and Pickrell et al. (2016) detect an enrichment of candidate SNPs in coding sequences, as is expected if pleiotropy over coding variation is most difficult to resolve.

A second potential factor that could explain the persistence of antagonistically pleiotropic polymorphisms is the extent to which they affect many other genes. In Chapter 5, I presented evidence that the *fruitless* gene, a master regulator of sexual differentiation in *D. melanogaster* that affects the expression of many other genes, is under balancing selection. If the mechanism for balancing selection is sexual antagonism (which requires further experiments to demonstrate), this could indicate that the extent to which genes affect the expression of other genes helps prevent resolution of sexual antagonism (Mank et al. 2008). It should be noted that, contrary to this hypothesis, additional analyses performed by Mark Hill (Appendix A) do not detect a general correlation between the presence of sexually antagonistic polymorphisms (Chapter 3) and proxies for pleiotropy (the

extent to which genes are expressed across tissues, or the number of protein-protein interactions they exhibit). The predicted relationship between tissue pleiotropy and sexual antagonism is therefore currently not directly supported by any data.

A third mechanism that could help alleviate the genetic constraints inherent in antagonistic pleiotropy is sex-, trait- or environment-specific dominance (Fry 2010; Spencer & Priest 2016; Wittmann et al. 2017), where each allele is preferentially expressed in the context in which it is beneficial. For example, under sex-specific dominance, the male-beneficial allele is preferentially expressed in males and the female-beneficial allele preferentially expressed in females when heterozygous. Some indication that context-dependent dominance could be an important mechanism for alleviating genetic conflict comes from a study of the *VGLL3* locus in salmon (*Salmo salar*), where it has been shown that this sexually antagonistic locus has different dominance coefficients in each sex (Barson et al. 2015). Furthermore, sex-specific dominance may help explain why sexually antagonistic polymorphisms are not disproportionately found on the X chromosome (Chapter 3), as theory predicts (Fry 2010). As an aside, an interesting property of context-dependent dominance is that it alleviates genetic conflict *while also* maintaining genetic polymorphisms. Thus, although the presence of antagonistic polymorphisms across long timescales is often used as a substitute for the idea that 'genetic constraints' are present across populations, this is not necessarily so if polymorphisms are maintained under context-specific dominance where the constraint has been alleviated.

While all the mechanisms for resolving antagonistic pleiotropy discussed so far may be quite general, some mechanisms are more specific to certain types of pleiotropy. For example, the constraints imposed by fluctuating selection can be overcome through low migration between environments, such as might be generated by physical barriers between populations and allowing each population to become locally adapted (Yeaman & Otto 2011). Numerous instances of local adaptation under allopatry show that this mechanism for resolving pleiotropy between environments is widespread. However, low migration to resolve antagonistic pleiotropy between environments may be more difficult in the case of temporal rather than spatial fluctuations (McDonald & Yeaman 2018). Similarly, low migration is not possible in the case of negative trade-offs between traits, since trait correlations arise from the phenotypic expression of a single genome. Here, the resolution of antagonistic pleiotropy may require relatively rare events, such as the nonsynonymous mutation in the *FLiA* gene in *E. coli* (Yi & Dean 2016). Sexually antagonistic polymorphisms are a somewhat intermediate case, as 'migration' cannot be limited on autosomes but can be limited by the accumulation of polymorphisms on chromosomes with sex-limited or sex-biased transmission, such as the sex-determining regions (Rice 1987) or the X chromosome (Rice 1984; Patten & Haig 2009). For example, the antagonistic *Pax7* locus is situated near the sex-determining region, where each allele can be preferentially expressed in the sex it benefits (Roberts et al. 2009). Interestingly, and contrary to these classic predictions, the GWAS data presented in Chapter 3 does not support

an enrichment of candidate antagonistic polymorphisms on the X chromosome (Chapter 3).

Finally, the extent to which antagonistically pleiotropic polymorphisms constrain evolution will also depend on population-specific attributes. In particular, differences in effective population sizes will be important, since antagonistic polymorphisms are highly sensitive to genetic drift relative to non-antagonistic polymorphisms (Connallon & Clark 2012; Mullon et al. 2012; Hesketh et al. 2013). This population-specific effect may help explain why, on the one hand, I reported that the intersexual correlation for fitness rapidly evolved from negative to non-negative in the small laboratory LH_M population (Chapters 2 and 3), but on the other hand, I also showed that sexually antagonistic polymorphisms are balanced for long time periods across very large wild populations of *D. melanogaster* (Chapter 4 and Langley et al. 2012).

6.4. Future directions

6.4.1. Sexual antagonism and sex-biased gene expression

To identify sexual antagonism over gene expression, Innocenti & Morrow (2010) tested for a significant sex-by-fitness interaction on gene expression levels. However, the analyses presented in Chapter 2 demonstrate that the lines included in their transcriptomic analysis were not random samples from LH_M. As a result, many of the candidate antagonistic genes identified are likely to be false positives. To gain a more accurate picture of sexually antagonistic expression, it would be beneficial to repeat the gene expression analyses performed by these researchers while correcting for population

structure, in a way analogous to the re-analyses of sex-specific fitness presented in Chapter 2.

Furthermore, given the issues detected with Innocenti & Morrow's (2010) data, it would be interesting to revisit the conclusions drawn based on this data by follow-up studies (Griffin et al. 2013; Cheng & Kirkpatrick 2016). For example, Cheng & Kirkpatrick (2016) observed that antagonistically expressed genes—as identified by Innocenti & Morrow (2010)—tend to have moderately sex-biased gene expression, while non-antagonistically expressed genes tend to have either unbiased expression or extremely sex-biased expression. From these patterns, they proposed a 'Twin Peaks' model of the relationship between sexually antagonistic selection and sex-biased gene expression, where antagonistic genes are most likely to have intermediate levels of sex-biased gene expression. In light of the re-analyses presented in Chapter 2, it is important to test whether this relationship is robust to a re-analysed set of antagonistically expressed genes that takes into account population structure. Alternatively, the 'Twin Peaks' model could be tested by looking at sex-biased expression among candidate antagonistic genes identified in the GWAS presented in Chapter 3.

Some analyses have already been performed with this goal in mind. For example, it has shown that antagonistic genes are less sex-biased than the genome-wide average (Mark Hill, Appendix A). Nevertheless, a detailed portrait of the relationship between sexually antagonistic selection and sex-biased gene expression remains to be painted.

6.4.2. Experimental verification of candidate antagonistic polymorphisms

An important task for future studies is to experimentally verify the fitness effects of candidate sexual antagonistic polymorphisms identified in the GWAS (Chapter 3). Experimental verification is needed because GWAS studies are correlational and cannot on their own establish a causal relationship between genotype and phenotype. The GWAS data presented in Chapter 3 provides a particularly promising starting point for experimental work because it is an unbiased screen for sexually antagonistic polymorphisms, in contrast to the 'candidate gene' approach employed in the case of *fru* (Chapter 5).

With this goal in mind, one possibility is simply to replicate the experiments presented in Chapter 5 using a single GWAS candidate as a starting point. One could then initiate fly populations with extreme allele frequencies at a candidate locus and track frequency changes over time, with the expectation that sexually antagonistic selection will cause allele frequencies to converge to a stable intermediate frequency. Alternatively, one could introgress alternative homozygous genotypes (*i.e.*, male-beneficial and female-beneficial genotypes respectively) into a standard genetic background in order to isolate the fitness effects of the genotype of interest while keeping genetic variation elsewhere constant. Furthermore, one could extend the analyses presented in Chapter 5 to estimate selection coefficients associated with each genotype, as has been done elsewhere (Loog et al. 2017).

There is also scope to improve upon the experiments presented in Chapter 5. For example, additional precision could be gained relative to the introgression approach by employing genome editing, such as the CRISPR/Cas9 RNA-guided nuclease system (Lamb et al. 2017). In contrast to introgression, which requires many generations to be effective and relies on random recombination events throughout the genome to homogenise the original genetic background, genome editing ensures that only the candidate site of interest is modified and leaves the rest of the genome unchanged. Any fitness effects subsequently measured can be attributed with certainty to the candidate site of interest.

Finally, experimental verification need not be confined to a single site. It could encompass genome-wide candidate polymorphisms. For instance, applying sex-limited selection (Morrow et al. 2008) and measuring genome-wide allele frequencies before and after selection (for example, through pooled sequencing (Schlötterer et al. 2014)) can provide a test of the fitness effects of genome-wide antagonistic polymorphisms. If a candidate polymorphism has antagonistic fitness effects, applying male-limited selection should increase the frequency of the male-beneficial allele, and applying female-limited selection should increase the frequency of the female-beneficial allele. Sex-limited evolution and measurement of allele frequencies before and after selection has in fact been undertaken in LH_M (Hill 2017) and this data could be analysed in the light of the candidate antagonistic SNPs uncovered in Chapter 3.

6.4.3. Mechanisms of conflict resolution

A major knowledge gap in the field is a detailed understanding of mechanisms that 'resolve' sexual antagonism—*i.e.*, the mechanisms that permit each allele to be expressed in the sex it confers fitness benefits to. The term 'resolution' encompasses a broad set of processes and can involve the evolution of sex-specific expression to resolve conflict over transcription levels or the simultaneous fixation of alternative alleles (e.g. through duplication and the evolution of sex-specific regulation in each paralog) to resolve conflict over coding sequences. Alleviation of conflict could also involve the emergence of sex-specific dominance, the evolution of sex-specific genomic imprinting, or the relocation of genes onto sex chromosomes (Bonduriansky & Chenoweth 2009). A better understanding of antagonistic resolution can therefore shed light on the processes that facilitate or hamper the long-term persistence of antagonistic polymorphisms. I highlight two readily testable mechanisms of resolution.

First, in Chapter 3, it was reported that antagonistic genetic variation is not disproportionately X-linked as predicted by classic theory (Rice 1984; Patten & Haig 2009). One mechanism that could explain this result is sex-specific dominance, which tends to shift the distribution of antagonistic polymorphisms toward autosomes (Fry 2010; Spencer & Priest 2016). Under sex-specific dominance, the female-beneficial allele is dominant in females and the male-beneficial allele is dominant in males, which helps maintain a polymorphism by ensuring that the sex-averaged fitness of heterozygotes is greater than the sex-averaged fitness of either homozygote. Testing for this mechanism is straightforward. It simply requires the fitness of a candidate polymorphism to be measured in all three possible states (homozygote AA,

homozygote aa, and heterozygote Aa), while keeping genetic variation elsewhere constant. If the candidate polymorphism is antagonistic but there is no sex-specific dominance, heterozygotes will have intermediate sex-averaged fitness; if there is sex-specific dominance, heterozygotes will have higher than intermediate sex-averaged fitness.

Second, in Chapter 3, I showed that there is an enrichment of antagonistic SNPs among missense variants. This implies that a key limiting factor for the resolution of sexual antagonism is the waiting time required for a duplication event to arise (Stewart et al. 2010; Connallon & Clark 2011b). It also implies that duplication events at antagonistic genes will be favoured by selection. Given that many antagonistic polymorphisms appear to have arisen many thousands of generations ago—some perhaps before the speciation event between *D. melanogaster* and *D. simulans*—comparing rates of duplication at antagonistic and non-antagonistic genes in different populations, or in a sister species like *D. simulans*, can provide a test of this theory. If duplication is an important mechanism for the resolution of sexual antagonism, an increased rate of paralogy among orthologs of *D. melanogaster* antagonistic genes should be observed. If not, then orthologs of antagonistic and non-antagonistic genes will exhibit similar numbers of paralogs. This analysis could further be coupled with an examination of sex-specific expression among the paralogs identified; under the hypothesis that paralogy events occur to resolve sexual conflict, one would expect to find increased sex-specific expression among paralogs of formerly antagonistic genes.

6.4.4. Antagonistic vs. non-antagonistic balancing selection

The analyses presented in Chapters 3 and 4 indicate that candidate antagonistic loci exhibit signatures of balancing selection and are associated with coding variation. This contrasts with previous scans for balancing selection, which have indicated that candidate balanced polymorphisms tend to be associated with regulatory variation and with immunity functions (Andrés et al. 2009; Leffler et al. 2013). The contrast between antagonistic and non-antagonistic polymorphisms represents a potentially fruitful avenue for future research. What proportion of balanced polymorphisms identified through a selection scan are candidate sexually antagonistic polymorphisms in LH_M? How do the functional properties of sexually antagonistic and non-sexually antagonistic balanced polymorphisms differ?

Answering these questions should be readily achievable. The public availability of whole-genome sequence data from >1000 *D. melanogaster* genomes (Lack et al. 2016) allows scans for balancing selection across populations to be performed. By examining the overlap between antagonistic polymorphisms and all balanced polymorphisms found in selection scan, one can place a minimum bound on the proportion of all balanced polymorphisms that are sexually antagonistic. This would clarify the relative role of antagonism in generating balancing selection in general. Furthermore, by comparing the functional properties of antagonistically balanced and non-antagonistically balanced polymorphisms, one can gain insights into the aspects of genomic architecture that facilitate or hamper the long-term maintenance of both types of balanced polymorphism. Analyses such as those described are currently being undertaken.

6.4.5. Sexual antagonism in other species

The antagonistic polymorphisms described in this thesis were identified in a single species, *D. melanogaster*, and a single population, LH_M. It important to establish whether the functional and population genetic properties of sexual antagonism described here apply in general.

Humans are one promising organism with which to perform such a task. Thanks to large-scale genotyping and phenotyping initiatives such as the UK Biobank (Sudlow et al. 2015), measurements of relative lifetime reproductive success (rLRS), with associated genomes, are now available for thousands of human males and females. Studies have generally used this data to estimate a genetic correlation between a trait of interest and rLRS, and thus to establish that natural selection affects the trait in question (Tropf et al. 2015; Beauchamp 2016; Kong et al. 2017; Sanjak et al. 2017). However, measurements of male and female rLRS could also be leveraged to identify sexually antagonistic loci. This could be achieved by correlating male rLRS with female rLRS, partitioning this correlation by genomic region (Shi et al. 2017) and identifying regions where there is a negative fitness correlation (sexual antagonism) and those where there is a positive fitness correlation (sexual concordance).

While this task may at first seem difficult—given that simultaneous measurements of male and female rLRS for a given genotype are not feasible—methods such as LD score regression (Bulik-Sullivan et al. 2015) permit summary statistics from separate GWAS datasets to be combined and allow genetic correlations between traits to be computed even in the

absence of overlapping individuals in both datasets. The fact that certain traits, like human height, are positively correlated with male rLRS and negatively correlated with female rLRS—in other words, that height is a sexually antagonistic trait (Sanjak et al. 2017)—suggests that candidate antagonistic genes should be readily detectable using this method.

In summary, this thesis has described the genetics and evolutionary dynamics of sexually antagonistic polymorphisms. It has shown that antagonistic polymorphisms are a pervasive feature of genomic variation and that this variation can constrain the evolution of sexual dimorphism over long time periods. The genomic data presented here provides a platform for future bioinformatic and experimental studies. These further analyses should provide a fuller understanding of the way in which genomes are moulded to accommodate divergent and often contradictory evolutionary interests.

References

Abbott, J.K., Bedhomme, S. & Chippindale, A.K., 2010. Sexual conflict in wing size and shape in *Drosophila melanogaster*. *Journal of Evolutionary Biology*, 23(9), pp.1989–1997.

Abbott, J.K. & Morrow, E.H., 2011. Obtaining snapshots of genetic variation using hemiclonal analysis. *Trends in Ecology and Evolution*, 26(7), pp.359–368.

Aguilar, A. et al., 2004. High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proceedings of the National Academy of Sciences of the United States of America*, 101(10), pp.3490–3494.

Albert, A.Y.K. & Otto, S.P., 2005. Sexual selection can resolve sex-linked sexual antagonism. *Science*, 310(5745), pp.119–121.

Alexander, D.H., Novembre, J. & Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), pp.1655–1664.

Allison, A.C., 1954. Protection afforded by sickle-cell trait against subtertian malarial infection. *British Medical Journal*, 1(4857), pp.290–294.

Andersson, M.B., 1994. *Sexual Selection*, Princeton University Press.

Andrés, A.M. et al., 2009. Targets of balancing selection in the human genome. *Molecular Biology and Evolution*, 26(12), pp.2755–2764.

Arnqvist, G., 2011. Assortative mating by fitness and sexually antagonistic genetic variation. *Evolution*, 65(7), pp.2111–2116.

Arnqvist, G. & Rowe, L., 2005. *Sexual conflict*, Princeton University Press.

Arnqvist, G., Vellnow, N. & Rowe, L., 2014. The effect of epistasis on

sexually antagonistic genetic variation. *Proceedings of the Royal Society B: Biological Sciences*, 281(1787), p.20140489.

Asthana, S., Schmidt, S. & Sunyaev, S., 2005. A limited role for balancing selection. *Trends in Genetics*, 21(1), pp.30–32.

Astle, W. & Balding, D.J., 2009. Population structure and cryptic relatedness in genetic association studies. *Statistical Science*, 24(4), pp.451–471.

Aulchenko, Y.S. et al., 2007. GenABEL: an R library for genome-wide association analysis. *Bioinformatics*, 23(10), pp.1294–1296.

Van der Auwera, G.A. et al., 2013. From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Current Protocols in Bioinformatics*, 43, p.11.10.1–11.10.33.

Ayroles, J.F. et al., 2009. Systems of complex genetics in *Drosophila melanogaster*. *Nature Genetics*, 41(3), pp.299–307.

Barson, N.J. et al., 2015. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature*, 528(7582), pp.405–408.

Bateman, 1948. Intra-sexual selection in *Drosophila*. *Heredity*, 2(3), pp.349–368.

Beauchamp, J.P., 2016. Genetic evidence for natural selection in humans in the contemporary United States. *Proceedings of the National Academy of Sciences of the United States of America*, 113(28), pp.7774–7779.

Bedhomme, S. et al., 2008. Reproductive behaviour evolves rapidly when intralocus sexual conflict is removed. *PLoS ONE*, 3(5), p.e2187.

Benjamini, Y. & Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal*

Statistical Society, 57(1), pp.289–300.

Berg, E.C. & Maklakov, A.A., 2012. Sexes suffer from suboptimal lifespan because of genetic conflict in a seed beetle. *Proceedings of the Royal Society B: Biological Sciences*, 279(1745), pp.4296–4302.

Berg, J.J. & Coop, G., 2014. A population genetic signal of polygenic adaptation. *PLoS Genetics*, 10(8), p.e1004412.

Berger, D. et al., 2014. Intralocus sexual conflict and environmental stress. *Evolution*, 68(8), pp.2184–2196.

Bergland, A.O. et al., 2014. Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in *Drosophila*. *PLoS Genetics*, 10(11), p.e1004775.

Bitarello, B.D. et al., 2018. Signatures of long-term balancing selection in human genomes. *Genome Biology and Evolution*, 10(3), pp.939–955.

Bonduriansky, R. & Chenoweth, S.F., 2009. Intralocus sexual conflict. *Trends in Ecology & Evolution*, 24(5), pp.280–288.

Borgia, G., 1979. Sexual selection and the evolution of mating systems. In B. & Blum, ed. *Sexual Selection and Reproductive Competition in Insects*. New York Academic Press, pp. 19–80.

Brommer, J.E. et al., 2007. The intersexual genetic correlation for lifetime fitness in the wild and its implications for sexual selection. *PLoS ONE*, 2(8), p.e744.

Brooks, R., 2000. Negative genetic correlation between male sexual attractiveness and survival. *Nature*, 406(6791), pp.67–70.

Bubb, K. et al., 2006. Scan of human genome reveals no new loci under ancient balancing selection. *Genetics*, 173(4), pp.2165–2177.

Bulik-Sullivan, B.K. et al., 2015. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*, 47(3), pp.291–295.

Bush, W.S. & Moore, J.H., 2012. Genome-wide association studies. *PLoS Computational Biology*, 8(12), p.e1002822.

Caballero, A., 1995. On the effective size of populations with separate sexes, with particular reference to sex-linked genes. *Genetics*, 139(2), pp.1007–1011.

Cabrera, C.P. et al., 2012. Uncovering networks from genome-wide association studies via circular genomic permutation. *G3: Genes, Genomes, Genetics*, 2(9), pp.1067–1075.

Calsbeek, R. & Sinervo, B., 2004. Within-clutch variation in offspring sex determined by differences in sire body size: Cryptic mate choice in the wild. *Journal of Evolutionary Biology*, 17(2), pp.464–470.

Charlesworth, B., 2015. Causes of natural variation in fitness: evidence from studies of Drosophila populations. *Proceedings of the National Academy of Sciences of the United States of America*, 112(6), pp.1662–1629.

Charlesworth, B., Coyne, J. & Barton, N., 1987. The relative rates of evolution of sex chromosomes and autosomes. *The American Naturalist*, 130(1), pp.113–146.

Charlesworth, B., Miyo, T. & Borthwick, H., 2007. Selection responses of means and inbreeding depression for female fecundity in *Drosophila melanogaster* suggest contributions from intermediate-frequency alleles to quantitative trait variation. *Genetical Research*, 89(2), pp.85–91.

Charlesworth, B., Morgan, M.T. & Charlesworth, D., 1993. The effect of

deleterious mutations on neutral molecular variation. *Genetics*, 134(4), pp.1289–1303.

Charlesworth, D., 2006. Balancing selection and its effects on sequences in nearby genome regions. *PLoS Genetics*, 2(4), p.e64.

Cheng, C. & Kirkpatrick, M., 2016. Sex-specific selection and sex-biased gene expression in humans and flies. *PLoS Genetics*, 12(9), p.e1006170.

Chippindale, A.K., Gibson, J.R. & Rice, W.R., 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*, 98(4), pp.1671–1675.

Collet, J.M. et al., 2016. Rapid evolution of the intersexual genetic correlation for fitness in *Drosophila melanogaster*. *Evolution*, 70(4), pp.781–795.

Comeron, J.M., Ratnappan, R. & Bailin, S., 2012. The many landscapes of recombination in *Drosophila melanogaster*. *PLoS Genetics*, 8(10), p.e1002905.

Connallon, T., 2010. Genic capture, sex linkage, and the heritability of fitness. *The American Naturalist*, 175(5), pp.564–576.

Connallon, T. & Clark, A.G., 2012. A general population genetic framework for antagonistic selection that accounts for demography and recurrent mutation. *Genetics*, 190(4), pp.1477–1489.

Connallon, T. & Clark, A.G., 2013. Antagonistic versus nonantagonistic models of balancing selection: Characterizing the relative timescales and hitchhiking effects of partial selective sweeps. *Evolution*, 67(3), pp.908–917.

Connallon, T. & Clark, A.G., 2011a. Association between sex-biased gene expression and mutations with sex-specific phenotypic consequences in *Drosophila*. *Genome Biology and Evolution*, 3(1), pp.151–155.

Connallon, T. & Clark, A.G., 2014a. Balancing selection in species with separate sexes: Insights from Fisher's geometric model. *Genetics*, 197(3), pp.991–1006.

Connallon, T. & Clark, A.G., 2014b. Evolutionary inevitability of sexual antagonism. *Proceedings of the Royal Society B: Biological Sciences*, 281(1776), pp.2013–2123.

Connallon, T. & Clark, A.G., 2011b. The resolution of sexual antagonism by gene duplication. *Genetics*, 187(3), pp.919–937.

Connallon, T. & Hall, M.D., 2016. Genetic correlations and sex-specific adaptation in changing environments. *Evolution*, 70(10), pp.2186–2198.

Cotton, S., Fowler, K. & Pomiankowski, A., 2004. Do sexual ornaments demonstrate heightened condition-dependent expression as predicted by the handicap hypothesis? *Proceedings of the Royal Society B: Biological Sciences*, 271(1541), pp.771–783.

Cowley, D.E. & Atchley, W.R., 1988. Quantitative genetics of *Drosophila melanogaster*. II. Heritabilities and genetic correlations between sexes for head and thorax traits. *Genetics*, 119(2), pp.421–433.

Cox, R.M. & Calsbeek, R., 2009. Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. *The American Naturalist*, 173(2), pp.176–187.

Crespi, B.J., 1989. Causes of assortative mating in arthropods. *Animal Behaviour*, 38(6), pp.980–1000.

Croze, M. et al., 2017. A genome-wide scan for genes under balancing selection in *Drosophila melanogaster*. *BMC Evolutionary Biology*, 17(1).

Cruickshank, T.E. & Hahn, M.W., 2014. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, 23(13), pp.3133–3157.

Curtsinger, J.W., Service, P.M. & Prout, T., 1994. Antagonistic pleiotropy, reversal of dominance, and genetic polymorphism. *The American Naturalist*, 144(2), pp.210–228.

Danecek, P. et al., 2011. The variant call format and VCFtools. *Bioinformatics*, 27(15), pp.2156–2158.

Darling, A.E., Mau, B. & Perna, N.T., 2010. Progressivemauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE*, 5(6), p.e11147.

Darwin, C., 1871. *The Descent of Man*, John Murray.

Day, T. & Bonduriansky, R., 2004. Intralocus sexual conflict can drive the evolution of genomic imprinting. *Genetics*, 167(4), pp.1537–1546.

DeGiorgio, M., Lohmueller, K.E. & Nielsen, R., 2014. A model-based approach for identifying signatures of ancient balancing selection in genetic data. *PLoS Genetics*, 10(8), p.e1004561.

Delcourt, M., Blows, M.W. & Rundle, H.D., 2009. Sexually antagonistic genetic variance for fitness in an ancestral and a novel environment. *Proceedings of the Royal Society B: Biological Sciences*, 276(1664), pp.2009–2014.

Delph, L. et al., 2004. Genetic constraints on floral evolution in a sexually dimorphic plant revealed by artificial selection. *Evolution*, 58(9),

pp.1936–1946.

Delph, L.F. et al., 2011. Environment-dependent intralocus sexual conflict in a dioecious plant. *New Phytologist*, 192(2), pp.542–552.

Dobzhansky, T., 1955. A review of some fundamental concepts and problems of population genetics. *Cold Spring Harbor Symposia on Quantitative Biology*, 20, pp.1–15.

Van Doorn, G.S., 2009. Intralocus sexual conflict. *Annals of the New York Academy of Sciences*, 1168(1), pp.52–71.

Van Doorn, G.S. & Kirkpatrick, M., 2007. Turnover of sex chromosomes induced by sexual conflict. *Nature*, 449(7164), pp.909–912.

Duchen, P. et al., 2013. Demographic inference reveals African and European admixture in the North American *Drosophila melanogaster* population. *Genetics*, 193(1), pp.291–301.

Ellegren, H. & Parsch, J., 2007. The evolution of sex-biased genes and sex-biased gene expression. *Nature Reviews Genetics*, 8(9), pp.689–98.

Ellegren, H. & Sheldon, B.C., 2008. Genetic basis of fitness differences in natural populations. *Nature*, 452(7184), pp.169–175.

Elyashiv, E. et al., 2016. A genomic map of the effects of linked selection in *Drosophila*. *PLoS Genetics*, 12(8), p.e1006130.

Fairbairn, D.J. & Roff, D.A., 2006. The quantitative genetics of sexual dimorphism: Assessing the importance of sex-linkage. *Heredity*, 97(5), pp.319–328.

Falconer, D.S. & Mackay, T.F.C., 1996. *Introduction to quantitative genetics*, Longman.

Fedorka, K.M. & Mousseau, T.A., 2004. Female mating bias results in

conflicting sex-specific offspring fitness. *Nature*, 429(6987), pp.65–67.

Fijarczyk, A. & Babik, W., 2015. Detecting balancing selection in genomes: Limits and prospects. *Molecular Ecology*, 24(14), pp.3529–3545.

Fisher, R.A., 1930. *The Genetical Theory of Natural Selection*, Clarendon Press.

Fitzpatrick, M.J., 2004. Pleiotropy and the genomic location of sexually selected genes. *The American Naturalist*, 163(6), pp.800–808.

Foerster, K. et al., 2007. Sexually antagonistic genetic variation for fitness in red deer. *Nature*, 447(7148), pp.1107–1110.

Forsman, A., 1995. Opposing fitness consequences of colour pattern in male and female snakes. *Journal of Evolutionary Biology*, 8(1), pp.53–70.

Fowler, K. et al., 1997. Genetic variation for total fitness in *Drosophila melanogaster*. *Proceedings of the Royal Society B: Biological Sciences*, 264(1379), pp.191–199.

Fraser, H.B., 2013. Gene expression drives local adaptation in humans. *Genome Research*, 23(7), pp.1089–1096.

Fry, J.D., 2010. The genomic location of sexually antagonistic variation: Some cautionary comments. *Evolution*, 64(5), pp.1510–1516.

Gailey, D.A. et al., 2006. Functional conservation of the *fruitless* male sex-determination gene across 250 Myr of insect evolution. *Molecular Biology and Evolution*, 23(3), pp.633–643.

Gay, L. et al., 2011. The genetic architecture of sexual conflict: Male harm and female resistance in *Callosobruchus maculatus*. *Journal of Evolutionary Biology*, 24(2), pp.449–456.

Gibson, J.R., Chippindale, A.K. & Rice, W.R., 2002. The X chromosome is a

hot spot for sexually antagonistic fitness variation. *Proceedings of the Royal Society B: Biological Sciences*, 269(1490), pp.499–505.

Gilks, W.P. et al., 2016. Whole genome resequencing of a laboratory-adapted *Drosophila melanogaster* population sample. *F1000Research*, 5, p.e2644.

Gillespie, J.H., 2004. *Population Genetics*, Johns Hopkins University Press.

Griffin, R.M. et al., 2013. The shared genome is a pervasive constraint on the evolution of sex-biased gene expression. *Molecular Biology and Evolution*, 30(9), pp.2168–2176.

Hadfield, J.D., 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software*, 33, pp.1–22.

Hahn, M.W., 2008. Toward a selection theory of molecular evolution. *Evolution*, 62(2), pp.255–265.

Haldane, J.B.S., 1949. Disease and evolution. *La Ricerca Scientifica*, 19(9), pp.68–76.

Haldane, J.B.S., 1937. The effect of variation on fitness. *The American naturalist*, 71(735), pp.337–349.

Harano, T. et al., 2010. Intralocus sexual conflict unresolved by sex-limited trait expression. *Current Biology*, 20(22), pp.2036–2039.

Hawkes, M.F. et al., 2016. Intralocus sexual conflict and insecticide resistance. *Proceedings of the Royal Society B: Biological Sciences*, 283(1843), p.20161429.

Hedrick and, P.W. & Parker, J.D., 1997. Evolutionary genetics and genetic variation of haplodiploids and X-linked genes. *Annual Review of Ecology*

and Systematics, 28(1), pp.55–83.

Hedrick, P.W., 2006. Genetic polymorphism in heterogeneous environments: The age of genomics. *Annual Review of Ecology, Evolution, and Systematics*, 37(1), pp.67–93.

Hedrick, P.W., 2012. What is the evidence for heterozygote advantage selection? *Trends in Ecology and Evolution*, 27(12), pp.698–704.

Hesketh, J., Fowler, K. & Reuter, M., 2013. Genetic drift in antagonistic genes leads to divergence in sex-specific fitness between experimental populations of *Drosophila melanogaster*. *Evolution*, 67(5), pp.1503–1510.

Hill, M.S., 2017. *The genetic basis of sexual antagonism in D. melanogaster*. Thesis (PhD), University College London.

Hinrichs, A. et al., 2006. The UCSC Genome Browser Database: update 2006. *Nucleic Acids Research*, 34, pp.D590–D598.

Houle, D., 1992. Comparing evolvability and variability of quantitative traits. *Genetics*, 130(1), pp.195–204.

Houle, D., Morikawa, B. & Lynch, M., 1996. Comparing mutational variabilities. *Genetics*, 143(3), pp.1467–1483.

Huang, W. et al., 2012. Epistasis dominates the genetic architecture of Drosophila quantitative traits. *Proceedings of the National Academy of Sciences of the United States of America*, 109(39), pp.15553–15559.

Huang, Y., Wright, S.I. & Agrawal, A.F., 2014. Genome-wide patterns of genetic variation within and among alternative selective regimes. *PLoS Genetics*, 10(8), p.e1004527.

Ingleby, F.C., Flis, I. & Morrow, E.H., 2015. Sex-biased gene expression and

sexual conflict throughout development. *Cold Spring Harbor Perspectives in Biology*, 7(1), p.a017632.

Innocenti, P. & Morrow, E.H., 2010. The sexually antagonistic genes of *Drosophila melanogaster* L. D. Hurst, ed. *PLoS Biology*, 8(3), p.e1000335.

James, J.W., 1973. Covariances between relatives due to sex-linked genes. *Biometrics*, 29(3), pp.584–588.

Johnston, S.E. et al., 2011. Genome-wide association mapping identifies the genetic basis of discrete and quantitative variation in sexual weaponry in a wild sheep population. *Molecular Ecology*, 20(12), pp.2555–2566.

Johnston, S.E. et al., 2013. Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature*, 502(7469), pp.93–95.

Kelly, J.K. & Willis, J.H., 2001. Deleterious mutations and genetic variation for flower size in *Mimulus guttatus*. *Evolution*, 55(5), pp.937–942.

Khaitovich, P. et al., 2005. Evolution: Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science*, 309(5742), pp.1850–1854.

Khaitovich, P. et al., 2004. Regional patterns of gene expression in human and chimpanzee brains. *Genome*, 14(8), pp.1462–1473.

Kidwell, J.F. et al., 1977. Regions of stable equilibria for models of differential selection in the two sexes under random mating. *Genetics*, 85(1), pp.171–183.

Kimura, M., 1958. On the change of population fitness by natural selection. *Heredity*, 12, pp.145–167.

Kimura, M., 1983. *The Neutral Theory of Molecular Evolution*, Cambridge

University Press.

Kirkpatrick, M. & Barton, N., 2006. Chromosome inversions, local adaptation and speciation. *Genetics*, 173(1), pp.419–434.

Kirkpatrick, M. & Hall, D.W., 2004. Sexual selection and sex linkage. *Evolution*, 58(4), pp.683–691.

Kirkpatrick, M. & Ryan, M.J., 1991. The evolution of mating preferences and the paradox of the lek. *Nature*, 350(6313), pp.33–38.

Kong, A. et al., 2017. Selection against variants in the genome associated with educational attainment. *Proceedings of the National Academy of Sciences of the United States of America*, 114(5), pp.e727-732.

Kosova, G., Abney, M. & Ober, C., 2010. Heritability of reproductive fitness traits in a human population. *Proceedings of the National Academy of Sciences of the United States of America*, 107(1), pp.1772–1778.

Kruuk, L.E.B. et al., 2000. Heritability of fitness in a wild mammal population. *Proceedings of the National Academy of Sciences*, 97(2), pp.698–703.

Kryvokhyzha, D. et al., 2016. The influence of population structure on gene expression and flowering time variation in the ubiquitous weed *Capsella bursa-pastoris* (Brassicaceae). *Molecular Ecology*, 25(5), pp.1106–1121.

Kusakabe, S. & Mukai, T., 1984. The genetic structure of natural populations of *Drosophila melanogaster*. XVIII. Clinal and uniform genetic variation over populations. *Genetics*, 108(3), pp.617–632.

Lack, J.B. et al., 2016. A thousand fly genomes: An expanded *Drosophila* Genome Nexus. *Molecular Biology and Evolution*, 33(12), pp.3308–3313.

Lack, J.B. et al., 2015. The Drosophila Genome Nexus: A population

genomic resource of 623 *Drosophila melanogaster* genomes, including 197 from a single ancestral range population. *Genetics*, 199(4), pp.1229–1241.

Lagisz, M. et al., 2012. Two distinct genomic regions, harbouring the *period* and *fruitless* genes, affect male courtship song in *Drosophila montana*. *Heredity*, 108(6), pp.602–608.

Lamb, A.M., Walker, E.A. & Wittkopp, P.J., 2017. Tools and strategies for scarless allele replacement in *Drosophila* using CRISPR/Cas9. *Fly*, 11(1), pp.53–64.

Lande, R., 1980. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution*, 34(2), pp.292–305.

Langley, C.H. et al., 2012. Genomic variation in natural populations of *Drosophila melanogaster*. *Genetics*, 192(2), pp.533–598.

Leffler, E.M. et al., 2013. Multiple instances of ancient balancing selection shared between humans and chimpanzees. *Science*, 339(6127), pp.1578–1582.

Leffler, E.M. et al., 2012. Revisiting an old riddle: What determines genetic diversity levels within species? *PLoS Biology*, 10(9), p.e1001388.

Levene, H., 1953. Genetic equilibrium when more than one ecological niche is available. *The American Naturalist*, 87(836), pp.331–333.

Lewontin, R.C. & Hubby, J.L., 1966. A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics*, 54, pp.595–609.

Li, H. et al., 2009. The Sequence Alignment/Map format and SAMtools.

Bioinformatics, 25(16), pp.2078–2079.

Li, H. & Stephan, W., 2006. Inferring the demographic history and rate of adaptive substitution in *Drosophila*. *PLoS Genetics*, 2(10), p.e1004775.

Long, T.A.F. et al., 2009. Estimating the heritability of female lifetime fecundity in a locally adapted *Drosophila melanogaster* population. *Journal of Evolutionary Biology*, 22(3), pp.637–643.

Long, T.A.F., Agrawal, A.F. & Rowe, L., 2012. The effect of sexual selection on offspring fitness depends on the nature of genetic variation. *Current Biology*, 22(3), pp.204–208.

Long, T.A.F. & Rice, W.R., 2007. Adult locomotory activity mediates intralocus sexual conflict in a laboratory-adapted population of Drosophila melanogaster. *Proceedings of the Royal Society B: Biological Sciences*, 274(1629), pp.3105–3112.

Loog, L. et al., 2017. Inferring allele frequency trajectories from ancient DNA indicates that selection on a chicken gene coincided with changes in medieval husbandry practices. *Molecular Biology and Evolution*, 34(8), pp.1981–1990.

Lucotte, E.A. et al., 2016. Detection of allelic frequency differences between the sexes in humans: a signature of sexually antagonistic selection. *Genome Biology and Evolution*, 8(5), pp.1489–1500.

Lunter, G. & Goodson, M., 2011. Stampy: A statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Research*, 21(6), pp.936–939.

Mackay, T.F.C., 2010. Mutations and quantitative genetic variation: Lessons from *Drosophila*. *Philosophical Transactions of the Royal Society B:*

Biological Sciences, 365(1544), pp.1229–1239.

Mackay, T.F.C. et al., 2012. The *Drosophila melanogaster* Genetic Reference Panel. *Nature*, 482(7384), pp.173–178.

Mainguy, J. et al., 2009. Father-offspring phenotypic correlations suggest intralocus sexual conflict for a fitness-linked trait in a wild sexually dimorphic mammal. *Proceedings of the Royal Society B: Biological Sciences*, 276(1675), pp.4067–4075.

Mank, J.E. et al., 2008. Pleiotropic constraint hampers the resolution of sexual antagonism in vertebrate gene expression. *The American Naturalist*, 171(1), pp.35–43.

Martin, A. & Orgogozo, V., 2013. The loci of repeated evolution: A catalog of genetic hotspots of phenotypic variation. *Evolution*, 67(5), pp.1235–1250.

Maynard Smith, J., 1998. *Evolutionary Genetics*, Oxford University Press.

Maynard Smith, J. & Haigh, J., 1974. The hitch-hiking effect of a favorable gene. *Genetical Research*, 23, pp.23–35.

McCleery, R.H. et al., 2004. Components of variance underlying fitness in a natural population of the great tit *Parus major*. *The American Naturalist*, 164(3), pp.62–72.

McDonald, J.H. & Kreitman, M., 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature*, 351(6328), pp.652–654.

McDonald, T.K. & Yeaman, S., 2018. Effect of migration and environmental heterogeneity on the maintenance of quantitative genetic variation: A simulation study. *Journal of Evolutionary Biology*, 31(9), pp.1386–1399.

McLaren, W. et al., 2010. Deriving the consequences of genomic variants

with the Ensembl API and SNP Effect Predictor. *Bioinformatics*, 26(16), pp.2069–2070.

Merilä, J. & Sheldon, B.C., 1999. Genetic architecture of fitness and nonfitness traits: Empirical patterns and development of ideas. *Heredity*, 83(2), pp.103–109.

Merilä, J. & Sheldon, B.C., 2000. Lifetime reproductive success and heritability in nature. *The American Naturalist*, 155(3), pp.301–310.

Merilä, J., Sheldon, B.C. & Ellegren, H., 1997. Antagonistic natural selection revealed by molecular sex identification of nestling collared flycatchers. *Molecular Ecology*, 6(12), pp.1167–1175.

Messina, F.J., 1993. Heritability and “evolvability” of fitness components in *Callosobruchus maculatus*. *Heredity*, 71, pp.623–629.

Mi, H. et al., 2017. PANTHER version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Research*, 45(D1), pp.D183–D189.

Mokkonen, M. et al., 2011. Negative frequency-dependent selection of sexually antagonistic alleles in *Myodes glareolus*. *Science*, 334(6058), pp.972–974.

Morrow, E.H., Stewart, A.D. & Rice, W.R., 2008. Assessing the extent of genome-wide intralocus sexual conflict via experimentally enforced gender-limited selection. *Journal of Evolutionary Biology*, 21(4), pp.1046–1054.

Mousseau, T.A. & Roff, D.A., 1987. Natural selection and the heritability of fitness components. *Heredity*, 59(2), pp.181–197.

Mullon, C., Pomiankowski, A. & Reuter, M., 2012. The effects of selection

and genetic drift on the genomic distribution of sexually antagonistic alleles. *Evolution*, 66(12), pp.3743–3753.

Neville, M.C. et al., 2014. Male-specific *fruitless* isoforms target neurodevelopmental genes to specify a sexually dimorphic nervous system. *Current Biology*, 24(3), pp.229–241.

Obbard, D.J. et al., 2012. Estimating divergence dates and substitution rates in the *Drosophila* phylogeny. *Molecular Biology and Evolution*, 29(11), pp.3459–3473.

Page, A.J. et al., 2016. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microbial Genomics*, 2(4), p.e000056.

Paradis, E., Claude, J. & Strimmer, K., 2004. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20(2), pp.289–290.

Pardo-Diaz, C., Salazar, C. & Jiggins, C.D., 2015. Towards the identification of the loci of adaptive evolution. *Methods in Ecology and Evolution*, 6(4), pp.445–464.

Parker, D.J. et al., 2014. The evolution of novelty in conserved genes: Evidence of positive selection in the *Drosophila fruitless* gene is localised to alternatively spliced exons. *Heredity*, 112(3), pp.300–306.

Parker, G.A., 1979. *Sexual Selection and Sexual Conflict*, Murray S. Blum.

Parsch, J. et al., 2010. On the utility of short intron sequences as a reference for the detection of positive and negative selection in *Drosophila*. *Molecular Biology and Evolution*, 27(6), pp.1226–1234.

Parsch, J. & Ellegren, H., 2013. The evolutionary causes and consequences of sex-biased gene expression. *Nature Reviews Genetics*, 14(2), pp.83–

7.

Patten, M.M. & Haig, D., 2009. Maintenance or loss of genetic variation under sexual and parental antagonism at a sex-linked locus. *Evolution*, 63(11), pp.2888–2895.

Patten, M.M., Haig, D. & Úbeda, F., 2010. Fitness variation due to sexual antagonism and linkage disequilibrium. *Evolution*, 64(12), pp.3638–3642.

Pavlidis, P. et al., 2012. A critical assessment of storytelling: Gene ontology categories and the importance of validating genomic scans. *Molecular Biology and Evolution*, 29(10), pp.3237–3248.

Pennell, T.M. & Morrow, E.H., 2013. Two sexes, one genome: The evolutionary dynamics of intralocus sexual conflict. *Ecology and Evolution*, 3(6), pp.1819–1834.

Pettay, J.E. et al., 2005. Heritability and genetic constraints of life-history trait evolution in preindustrial humans. *Proceedings of the National Academy of Sciences of the United States of America*, 102(8), pp.2838–2843.

Pfeifer, B. et al., 2014. PopGenome: An efficient swiss army knife for population genomic analyses in R. *Molecular Biology and Evolution*, 31(7), pp.1929–1936.

Pickrell, J.K. et al., 2016. Detection and interpretation of shared genetic influences on 42 human traits. *Nature Genetics*, 48(7), pp.709–717.

Pischedda, A. & Chippindale, A.K., 2006. Intralocus sexual conflict diminishes the benefits of sexual selection. *PLoS Biology*, 4(11), pp.2099–2103.

Poissant, J., Wilson, A.J. & Coltman, D.W., 2010. Sex-specific genetic

variance and the evolution of sexual dimorphism: A systematic review of cross-sex genetic correlations. *Evolution*, 64(1), pp.97–107.

Pomiankowski, A. & Moller, A.P., 1995. A resolution of the lek paradox. *Proceedings of the Royal Society B: Biological Sciences*, 260(1357), pp.21–29.

Pool, J.E. et al., 2012. Population genomics of sub-Saharan *Drosophila melanogaster*: African diversity and non-African admixture. *PLoS Genetics*, 8(12), p.e1003080.

Pool, J.E. & Aquadro, C.F., 2007. The genetic basis of adaptive pigmentation variation in *Drosophila melanogaster*. *Molecular Ecology*, 16(14), pp.2844–2851.

Pool, J.E. & Nielsen, R., 2007. Population size changes reshape genomic patterns of diversity. *Evolution*, 61(12), pp.3001–3006.

Prasad, N.G. et al., 2007. An evolutionary cost of separate genders revealed by male-limited evolution. *The American Naturalist*, 169(1), pp.29–37.

Price, A.L. et al., 2010. New approaches to population stratification in genome-wide association studies. *Nat. Rev. Genet.*, 11(7), pp.459–463.

Price, D.K. & Burley, N.T., 1993. Constraints on the evolution of attractive traits: genetic (co)variance of zebra finch bill colour. *Heredity*, 71(4), pp.405–412.

Price, T. & Schluter, D., 1991. On the low heritability of life-history traits. *Evolution*, 45(4), pp.853–861.

Punzalan, D., Delcourt, M. & Rundle, H., 2014. Comparing the intersex genetic correlation for fitness across novel environments in the fruit fly, *Drosophila serrata*. *Evolution in Health and Disease*, 112(2), pp.143–

148.

Purcell, S. et al., 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81(3), pp.559–575.

Qvarnström, A., Brommer, J.E. & Gustafsson, L., 2006. Testing the genetics underlying the co-evolution of mate choice and ornament in the wild. *Nature*, 441(1), pp.84–86.

Radwan, J., 2008. Maintenance of genetic variation in sexual ornaments: A review of the mechanisms. *Genetica*, 134(1), pp.113–127.

Randall, J.C. et al., 2013. Sex-stratified Genome-wide Association Studies Including 270,000 Individuals Show Sexual Dimorphism in Genetic Loci for Anthropometric Traits. *PLoS Genetics*, 9(6).

Reinhold, K., 1998. Sex linkage among genes controlling sexually selected traits. *Behavioral Ecology and Sociobiology*, 44(1), pp.1–7.

Reinhold, K. & Engqvist, L., 2013. The variability is in the sex chromosomes. *Evolution*, 67(12), pp.3662–3668.

Reuter, M. et al., 2017. Evolving plastic responses to external and genetic environments. *Trends in Genetics*, 33(3), pp.169–170.

Rice, W.R. et al., 2005. Inter-locus antagonistic coevolution as an engine of speciation: assessment with hemiclonal analysis. *Proceedings of the National Academy of Sciences of the United States of America*, 102(1), pp.6527–6534.

Rice, W.R., 1998. Male fitness increases when females are eliminated from gene pool: implications for the Y chromosome. *Proceedings of the National Academy of Sciences of the United States of America*, 95(11),

pp.6217–21.

Rice, W.R., 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evolution*, 38(4), pp.735–742.

Rice, W.R., 1996. Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. *Nature*, 381(6579), pp.232–234.

Rice, W.R., 1987. The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. *Evolution*, 41(4), p.911.

Rice, W.R. & Chippindale, A.K., 2001. Intersexual ontogenetic conflict. *Journal of Evolutionary Biology*, 14(5), pp.685–693.

Roberts, R.B., Ser, J.R. & Kocher, T.D., 2009. Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes. *Science*, 326(5955), pp.998–1001.

Roff, D.A. & Mousseau, T.A., 1987. Quantitative genetics and fitness: Lessons from *Drosophila*. *Heredity*, 58(1), pp.103–118.

Rogers, R.L. et al., 2014. Landscape of standing variation for tandem duplications in *Drosophila yakuba* and *Drosophila simulans*. *Molecular Biology and Evolution*, 31(7), pp.1750–1766.

Rose, M.R., 1982. Antagonistic pleiotropy, dominance, and genetic variation. *Heredity*, 48(1), pp.63–78.

Rowe, L., Chenoweth, S.F. & Agrawal, A.F., 2018. The genomics of sexual conflict. *The American Naturalist*, 192(2), pp.274–286.

Rowe, L. & Houle, D., 1996. The lek paradox and the capture of genetic variance by condition dependent traits. *Proceedings of the Royal Society B: Biological Sciences*, 263(1375), pp.1415–1421.

RStudio Team, 2015. RStudio: Integrated Development for R. RStudio. *RStudio, Inc., Boston, MA*. Available at: https://www.rstudio.com/.

Saitou, N. & Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4), pp.406–425.

Sanjak, J.S. et al., 2017. Evidence of directional and stabilizing selection in contemporary humans. *Proceedings of the National Academy of Sciences of the United States of America*, 115(1), pp.151–156.

Dos Santos, G. et al., 2015. FlyBase: Introduction of the *Drosophila melanogaster* Release 6 reference genome assembly and large-scale migration of genome annotations. *Nucleic Acids Research*, 43(D1), pp.D690–D697.

Schierup, M.H., Vekemans, X. & Charlesworth, D., 2000. The effect of subdivision on variation at multi-allelic loci under balancing selection. *Genetical Research*, 76(1), pp.51–62.

Schlötterer, C. et al., 2014. Sequencing pools of individuals — mining genome-wide polymorphism data without big funding. *Nature Reviews Genetics*, 15(11), pp.749–763.

Shi, H. et al., 2017. Local genetic correlation gives insights into the shared genetic architecture of complex traits. *The American Journal of Human Genetics*, 101(5), pp.737–751.

Siewert, K.M. & Voight, B.F., 2017. Detecting long-term balancing selection using allele frequency correlation. *Molecular Biology and Evolution*, 34(11), pp.2996–3005.

Slatkin, M., 1984. Ecological causes of sexual dimorphism. *Evolution*, 38(3),

pp.622–630.

Slatkin, M., 2008. Linkage disequilibrium - Understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, 9(6), pp.477–485.

Smith, D.T. et al., 2011. DDT resistance, epistasis and male fitness in flies. *Journal of Evolutionary Biology*, 24(6), pp.1351–1362.

Speed, D. et al., 2012. Improved heritability estimation from genome-wide SNPs. *American Journal of Human Genetics*, 91(6), pp.1011–1021.

Speed, D. et al., 2017. Reevaluation of SNP heritability in complex human traits. *Nature Genetics*, 49(7), pp.986–992.

Spencer, H.G. & Priest, N.K., 2016. The evolution of sex-specific dominance in response to sexually antagonistic selection. *The American Naturalist*, 187(5), pp.658–666.

Stewart, A.D., Pischedda, A. & Rice, W.R., 2010. Resolving intralocus sexual conflict: genetic mechanisms and time frame. *Journal of Heredity*, 101(1), pp.94–99.

Stewart, A.D. & Rice, W.R., 2018. Arrest of sex-specific adaptation during the evolution of sexual dimorphism in *Drosophila*. *Nature Ecology & Evolution*, 2(9), pp.1507–1513.

Storey, J.D. et al., 2007. Gene-expression variation within and among human populations. *The American Journal of Human Genetics*, 80(3), pp.502–509.

Stulp, G. et al., 2012. Intralocus sexual conflict over human height. *Biology Letters*, 8(6), pp.976–8.

Sudlow, C. et al., 2015. UK Biobank: An open access resource for identifying

the causes of a wide range of complex diseases of middle and old age. *PLoS Medicine*, 12(3), p.e1001779.

Svensson, E.I., McAdam, A.G. & Sinervo, B., 2009. Intralocus sexual conflict over immune defense, gender load, and sex-specific signaling in a natural lizard population. *Evolution*, 63(12), pp.3124–3135.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123(3), pp.585–595.

Tarka, M. et al., 2014. Intralocus sexual conflict over wing length in a wild migratory bird. *The American Naturalist*, 183(1), pp.62–73.

Teplitsky, C. et al., 2009. Heritability of fitness components in a wild bird population. *Evolution*, 63(3), pp.716–726.

The 1000 Genomes Project Consortium, 2010. A map of human genome variation from population-scale sequencing. *Nature*, 467(7319), pp.1061–1073.

Trivers, R.L., 1972. *Parental investment and sexual selection*, Aldine.

Tropf, F.C. et al., 2015. Human fertility, molecular genetics, and natural selection in modern societies. *PLoS ONE*, 10(6), p.e0126821.

Turchin, M.C. et al., 2012. Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nature Genetics*, 44(9), pp.1015–1019.

Turelli, M. & Barton, N.H., 2004. Polygenic variation maintained by balancing selection: pleiotropy, sex-dependent allelic effects and GxE interactions. *Genetics*, 166(2), pp.1053–1079.

Úbeda, F. et al., 2011. Stable linkage disequilibrium owing to sexual antagonism. *Proceedings of the Royal Society B: Biological Sciences*,

278(1707), pp.855–862.

Usui-Aoki, K. et al., 2000. Formation of the male-specific muscle in female *Drosophila* by ectopic *fruitless* expression. *Nature Cell Biology*, 2(8), pp.500–506.

Vekemans, X. & Slatkin, M., 1994. Gene and allelic genealogies at a gametophytic self-incompatibility locus. *Genetics*, 137(4), pp.1157–1165.

Villella, A. & Hall, J.C., 2008. Neurogenetics of courtship and mating in *Drosophila*. *Advances in Genetics*, 62, pp.67–184.

Vincent, C.M. & Sharp, N.P., 2014. Sexual antagonism for resistance and tolerance to infection in *Drosophila melanogaster*. *Proceedings of the Royal Society B: Biological Sciences*, 281(1788), p.20140987.

Wayne, M.L. et al., 2007. Simpler mode of inheritance of transcriptional variation in male *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*, 104(47), pp.18577–18582.

Whitehead, A. & Crawford, D.L., 2006. Variation within and among species in gene expression: Raw material for evolution. *Molecular Ecology*, 15(5), pp.1197–1211.

Whitlock, M.C. & Agrawal, A.F., 2009. Purging the genome with sexual selection: Reducing mutation load through selection on males. *Evolution*, 63(3), pp.569–582.

Wilkinson, G.S. et al., 2015. The locus of sexual selection: Moving sexual selection studies into the post-genomics era. *Journal of Evolutionary Biology*, 28(4), pp.739–755.

Williams, T. & Carroll, S., 2009. Genetic and molecular insights into the development and evolution of sexual dimorphism. *Nature Reviews Genetics*, 10(11), pp.797–804.

Wittkopp, P.J. & Kalay, G., 2011. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nature Reviews Genetics*, 13(1), pp.59–69.

Wittmann, M.J. et al., 2017. Seasonally fluctuating selection can maintain polymorphism at many loci via segregation lift. *Proceedings of the National Academy of Sciences of the United States of America*, 114(46), pp.E9932–E9941.

Yang, J. et al., 2011. Genome partitioning of genetic variation for complex traits using common SNPs. *Nature Genetics*, 43(6), pp.519–525.

Yeaman, S. & Otto, S.P., 2011. Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. *Evolution*, 65(7), pp.2123–2129.

Yi, X. & Dean, A.M., 2016. Phenotypic plasticity as an adaptation to a functional trade-off. *eLife*, 5, p.e19307.

Zhao, L. et al., 2015. Parallel gene expression differences between low and high latitude populations of *D. melanogaster* and *D. simulans*. *PLoS Genetics*, 11(5), p.e1005184.

Zhou, X. & Stephens, M., 2014. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature Methods*, 11(4), pp.407–409.

Appendix A

Genome-wide sexually antagonistic variants reveal longstanding constraints on sexual dimorphism in the fruitfly

Genome-wide sexually antagonistic variants reveal longstanding constraints on sexual dimorphism in the fruitfly

Filip Ruzicka[1#], Mark S. Hill[1,2#], Tanya M. Pennell[3], Ilona Flis[4], William P. Gilks[5], Fiona C. Ingleby[5], Kevin Fowler[1], Edward H. Morrow[5§,*], Max Reuter[1§,*]

[#, §] Equal contributions

* Corresponding authors

[1] Research Department of Genetics, Evolution and Environment, University College London, UK

[2] Department of Ecology and Evolutionary Biology, The University of Michigan, Ann Arbor, USA

[3] Daphne du Maurier Building, University of Exeter, Penryn, UK

[4] The Pirbright Institute, Pirbright, Surrey, UK

[5] School of Life Sciences, University of Sussex, Brighton, UK

The evolution of sexual dimorphism is constrained by a shared genome, leading to 'sexual antagonism' where different alleles at given loci are favoured by selection in males and females. Despite its wide taxonomic incidence, we know little about the identity, genomic location and evolutionary dynamics of antagonistic genetic variants. To address these deficits, we use sex-specific fitness data from 202 fully sequenced hemiclonal *D. melanogaster* fly lines to perform a genome-wide association study of sexual antagonism. We identify ~230 chromosomal clusters of candidate antagonistic SNPs. In contradiction to classic theory, we find no clear evidence that the X chromosome is a hotspot for sexually antagonistic variation. Characterising antagonistic SNPs functionally, we find a large excess of missense variants but little enrichment in terms of gene function. We also assess the evolutionary persistence of antagonistic variants by examining extant polymorphism in wild *D. melanogaster* populations. Remarkably, antagonistic variants are associated with multiple signatures of balancing selection across the *D. melanogaster* distribution range, indicating widespread and evolutionarily persistent (>10,000 years) genomic constraints. Based on our results, we propose that antagonistic variation accumulates due to constraints on the resolution of sexual conflict over protein coding sequences, thus contributing to the long-term maintenance of heritable fitness variation.

The divergent reproductive roles of males and females favour different phenotypes[1,2]. However, responses to these selective pressures are constrained by a shared genome, leading to 'sexual antagonism' where different alleles at given loci are favoured in the two sexes[1,3–5]. A wealth of quantitative genetic studies has established sexual antagonism as near ubiquitous across a wide range of taxa, including mammals[6] (and humans[7]), birds[8], reptiles[9], insects[10,11], fish[12,13] and plants[14]. Accordingly, sexual antagonism can be considered a major constraint on adaption and an important mechanism for the maintenance of fitness variation within populations[15].

However, despite its evolutionary importance, we have little understanding of the biological mechanisms underlying this conflict and virtually no empirical data on the identity and evolutionary dynamics of antagonistic alleles[13]. While a small number of individual antagonistic loci have been identified[12,13], these are of limited use for elucidating general properties of loci experiencing sexual antagonism. On a genome-wide scale, previous transcriptomic work in *D. melanogaster* has associated antagonistic fitness effects with patterns of gene expression[16]. But despite potentially revealing some of the molecular correlates of fitness variation, this approach cannot distinguish between causal antagonistic loci and their downstream regulatory targets. In humans, genome-wide frequency differences between males and females have been used to infer sexually antagonistic selection on viability[17], but this approach neglects important reproductive components of fitness. It is essential that we characterise causal antagonistic loci underlying lifetime reproductive success in order to understand the adaptive limits to sexual dimorphism and mechanisms of conflict resolution.

To address this shortcoming, we identified sexually antagonistic loci across the *D. melanogaster* genome and characterised their functional and

evolutionary properties. Specifically, we obtained male and female fitness data for over 200 hemiclonal lines that had been extracted from LH_M, the outbred, laboratory-adapted population in which sexually antagonistic fitness effects were first characterised[10,18]. Our fitness measurements estimate lifetime reproductive success in both sexes by replicating the regime under which LH_M has been maintained for over 20 years[19]. We combined these fitness data with high-coverage genome sequences[20] and performed a genome-wide-association-study (GWAS) to map the genetic basis of sexual antagonism. We then examined the properties of candidate antagonistic polymorphisms, including their genomic distribution across the X chromosome and autosomes, the functional characteristics of candidate polymorphisms and the genes in which they occur, and their population genomic dynamics across a number of wild *D. melanogaster* populations.

Results

Quantitative genetic analyses confirmed the presence of significant amounts of genetic variation for male and female fitness among the lines assayed (N=223). Estimating the genetic variances and covariances between the lines, we found appreciable heritabilities for fitness in both sexes (female h^2=0.42, 95% CI 0.30–0.54; male h^2=0.16, 95% CI 0.04–0.27). Comparable estimates were also obtained by treating single nucleotide polymorphisms (SNPs) as random effects in a linear mixed model and calculating SNP heritability (h^2_{SNP}) using restricted maximum likelihood (female h^2_{SNP}=0.59, SD 0.13, P<0.001; male h^2_{SNP}=0.29, SD 0.16, P=0.007). Overall, the intersexual genetic correlation for fitness did not differ significantly from zero in this sample of genotypes (r_{MF}=0.15, 95% CI −0.21–0.46). The presence of ample heritable fitness variation, combined with the lack of a strong positive intersexual genetic correlation for fitness, suggests the presence of sexually antagonistic fitness variants.

We quantified antagonistic fitness variation by calculating an 'antagonism index'. Specifically, we rotated the coordinate system of the male and female fitness plane by 45 degrees and extracted the position of individual fly lines on the axis ranging from extremely male-beneficial, female-detrimental (MB) fitness effects to extremely female-beneficial, male-detrimental (FB) fitness effects (Fig. 1A). This approach for defining an antagonism index is analogous to other linear transformations, such as the widely applied transformation of human height and weight into a Body Mass Index[21]. The antagonism index itself had high SNP heritability (h^2_{SNP}=0.51, SD 0.15, P=0.001), as expected from the heritability of its sex-specific fitness components.

To identify putative antagonistic SNPs, we performed a genome-wide association study (GWAS) based on the antagonism index and sequence polymorphism data for 765,654 common (MAF>0.05) and stringently quality-filtered SNPs across 202 of the 223 lines (see Methods; Sup. Fig. 1). We employed a linear mixed model that corrects for between-line relatedness and population structure by incorporating a genetic similarity matrix as a random effect[22] (Sup. Fig. 2). Figure 1B presents a Manhattan plot of raw P-values from SNP-wise association tests along the *D. melanogaster* genome.

Although no individual SNP reached genome-wide significance based on stringent Bonferroni correction ($P<6.53 \times 10^{-8}$), our focus was to characterise broad patterns associated with genome-wide antagonistic variation rather than identifying individual antagonistic sites with high confidence. Accordingly, we applied three main approaches to investigate the general properties of antagonistic SNPs and regions. First, we defined 2,372 candidate antagonistic SNPs (henceforth 'antagonistic SNPs') as SNP positions with false discovery rate (FDR) Q-values<0.3 (but note that the Q-

values we estimated are likely to be conservative, see Sup. Fig. 2). This threshold achieves a balance between false positives and false negatives that is suitable for genome-wide analyses[23] and allowed us to contrast the properties of antagonistic and non-antagonistic (Q-value≥0.3) SNPs. Second, we quantified the importance of different classes of SNPs (defined by chromosomal location or function) by partitioning total SNP heritability of the antagonism index ('antagonistic h^2_{SNP}') into the contribution of each class[24,25]. These contributions can then be tested for deviations from random expectations and interpreted without need for defining significance cut-offs for individual SNPs. Finally, we employed set-based association testing where the joint effect of a set of SNPs (such as those in a chromosomal window) on the phenotype is assessed. This joint analysis alleviates the multiple testing burden and can be used to define antagonistic windows with more stringent support (Q-value<0.1). Together these approaches allowed us to characterise the functional properties and evolutionary dynamics of antagonistic genetic variation.

We first examined the genomic distribution of antagonistic variants. The 2,372 antagonistic SNPs were significantly clustered along chromosome arms (median distance: 147bp on autosomes, 298bp on the X chromosome, permutation test: P<0.001 for autosomes and X, Sup. Fig. 3). Using LD clumping[26], we estimated that the antagonistic SNPs form approximately 226 independent clusters. Some previous theory[3] and empirical quantitative genetic results[17] suggest that the X chromosome should harbour a disproportionate amount of antagonistic genetic variation. This was not borne out by our data. We found that, relative to autosomes, the X chromosome neither contained a disproportionate number of antagonistic SNPs (Z-test, P>0.05, Sup. Fig. 4), nor contributed more antagonistic h^2_{SNP} than expected (Fig. 2A).

Our data also allowed us to provide some of the first insights into the biological functions that underlie sexual antagonism. At the most basic level, our results suggest that antagonism arises primarily due to adaptive conflict over coding sequences. Thus, genomic partitioning revealed that variants which result in missense changes were significantly over-represented among antagonistic SNPs (Sup. Fig. 4) and contributed significantly more antagonistic h^2_{SNP} than expected from their proportional genomic representation (Fig. 2A; Sup. Tab. 1). As expected, intergenic regions were under-represented among antagonistic SNPs and contributed qualitatively less antagonistic h^2_{SNP} than expected (Fig. 2A; Sup. Fig. 4). However, we found no evidence that SNP functions involved in expression regulation, such as 3'UTR, intronic, upstream or splice region variants, were over-represented among antagonistic SNPs or h^2_{SNP} (Fig. 2A; Sup. Fig. 4).

We next performed a series of analyses to characterise the properties of genes harbouring antagonistic SNPs (one or more antagonistic SNPs within $\pm$ 5Kb of the gene coordinates). The list of antagonistic genes included some genes known to be involved in sexual differentiation, including *male-specific-lethal-1*, *traffic jam*, and *roundabout 2*, the circadian clock gene *period*, and the Golgi-associated transport protein gene *Tango6* that has been previously found to harbour coding sequence polymorphism shared between *D. melanogaster* and *D. simulans*[27] (see Sup. Tab. 2 for a complete list of antagonistic genes). Reflecting the heterogeneous list of genes, Gene Ontology (GO) analysis revealed little evidence for preferential association of antagonistic variation with specific biological processes. Only one term, 'sodium-channel-regulator-activity', was significant after correction for multiple-testing (Q-value=0.013). However, this annotation is shared by only a few genes (N=5), a cluster of four of which carry antagonistic SNPs. It thus

appears that antagonism is not enriched in genes involved in specific functions.

While antagonistic genes were not enriched in specific functions, they did show lower than average sex-bias in gene expression. This pattern is expected because unbiased genes should be most prone to experiencing balanced, opposing selection pressures in the two sexes that would stabilise antagonistic variation. We found evidence for this enrichment both in qualitative terms, where fewer antagonistic genes than expected by chance showed significant sex-biased gene expression (observed=188, expected=212, 11.3% deficit, χ_1^2=7.78, P=0.005), and in quantitative terms, where antagonistic genes had a lower degree of sex-bias than did non-antagonistic genes (W=1309700, P<0.001, Fig. 2B). We did not, however, detect significant overlap between the antagonistic genes identified here and genes that had previously been shown to have sexually antagonistic expression patterns (opposing relationships between expression level and fitness in males and females[16], observed number of overlapping genes=41, expected number=36, χ_1^2=0.59, P=0.44). This discrepancy is not unexpected, as the causal genetic changes underlying antagonism need not primarily be associated with expression differences (as suggested by enrichment among missense variants), nor need genes showing expression divergence between the phenotypic fitness extremes necessarily carry causal genetic variants themselves. Finally, we tested whether antagonistic variation is enriched in genes that are likely to be subject to pleiotropic constraints, as this has been proposed to make sexual antagonism harder to resolve[28]. We did not find an association between antagonism and higher levels of pleiotropy, measured either as tissue-specificity[29] (τ: W=2319200, P=0.80) or as the number of protein-protein interactions[30] (PPIs: $F_{1,5276}$=2.43, P=0.12). This implies that

pleiotropy—at least as captured by τ and PPIs—does not contribute significantly to maintaining sexually antagonistic genetic variation.

In addition to assessing the functional properties of antagonistic loci, we also investigated the population genetic effects of sexual antagonism. Models predict that the opposing sex-specific fitness effects of antagonistic alleles generate balancing selection, resulting in elevated levels of genetic polymorphism at antagonistic loci[31–33]. Having identified putatively antagonistic variants, we can test this prediction by comparing levels of polymorphism at antagonistic and non-antagonistic loci. Doing so directly within the LH_M population is problematic, because the power to detect antagonistic effects is higher at more polymorphic sites, and candidates therefore tend to show above-average polymorphism. However, we can use data from independent populations and ask whether, for a given level of polymorphism in LH_M, polymorphism there is greater at antagonistic than at non-antagonistic sites. We performed this type of analysis (see Methods for details) using publicly available polymorphism data from the *Drosophila* Genetics Reference Panel[34,35] (DGRP), a collection of 205 wild-derived inbred lines. Like LH_M, the DGRP was established from a North American *D. melanogaster* population. Given the relatively recent colonisation of the continent by *D. melanogaster* (~150 years[36]), the two populations are closely related. We found that antagonistic SNPs had elevated minor allele frequencies (MAFs) in the DGRP, although owing to the close relationship between LH_M and the DGRP (and the resulting similarity in allele frequencies) this difference was not statistically significant (P=0.322, Fig. 3A,B). However, when using the P-value for the antagonistic effects of individual SNPs rather than a binary antagonistic/non-antagonistic categorisation of sites, we found a significant negative correlation between P-value and MAF in the DGRP, consistent with elevated polymorphism in the

DGRP at SNPs that are more closely associated with antagonism in LH_M (ρ=−0.055, P=0.044, Fig. 3C). This evidence for antagonism-driven balancing selection at individual sites was corroborated by patterns of regional polymorphism—measured as Tajima's D within 1000bp windows along the chromosome arms. Tajima's D was significantly higher in antagonistic windows (those with Q-value<0.1 in a window-based GWAS) than in non-antagonistic windows (Q-value≥0.1; $F_{1,115477}$=224.6, P<0.001, Fig. 3G). Overall, these analyses show that the heritable phenotypic variation in sex-specific fitness that can be generated and maintained by sexual antagonism is mirrored by a signal of increased polymorphism at the underlying genetic loci.

A key, yet so far unresolved question is whether antagonistic polymorphisms are mainly short-lived and population-specific or persist over prolonged periods of time. The analyses of polymorphism in the DGRP shed some light on this question, demonstrating that antagonistic polymorphisms are maintained at least over periods of tens to hundreds of years, or hundreds to a few thousand generations. In order to assess signals of balancing selection over longer time spans, we repeated these analyses with data from a population in *D. melanogaster*'s ancestral Sub-Saharan distribution range, in Zambia (ZI, 197 genomes from phase 3 of the *Drosophila* Population Genomics Project[37]; see also Sup. Fig. 5, which repeats the ZI analyses with identical results using 118 genomes from South Africa[38]). Just as in the DGRP, we found that antagonism generated a clear signature of balancing selection in this ancestral population sample. Analyses based on binary categories showed that antagonistic SNPs had significantly higher MAFs in ZI compared to non-antagonistic SNPs (P=0.024, Fig. 3D,E), while analyses based on P-values showed again that sites with stronger evidence for antagonistic effects had more elevated MAFs (ρ=−0.070, P=0.002, Fig. 3F).

At a larger chromosomal scale, antagonistic windows had significantly higher polymorphism (Tajima's D) than non-antagonistic windows ($F_{1,116099}$=60.63, P<0.001, Fig. 3G). Furthermore, they also exhibited lower population differentiation between DGRP and ZI (measured as F_{ST}; Wilcoxon Rank-Sum test, W=63416000, P=0.012, Fig. 3H; Sup. Fig. 5), in line with balancing selection maintaining similar frequencies across distant populations.

In addition to elevated polymorphism in antagonistic regions of the genome, we also found evidence for increased linkage disequilibrium (LD) – another hallmark of balancing selection[39]. We compared local LD (<1,000bp, measured as r^2) between pairs of antagonistic sites, pairs of non-antagonistic sites, and 'mixed' site pairs (consisting of an antagonistic and a non-antagonistic SNP) in the ZI population, which is most phylogenetically distant from LH_M and where a signal of LD should be weakest in the absence of long-term balancing selection. Consistent with selection, we found that pairs of antagonistic sites had higher LD in this population than pairs of non-antagonistic sites (Wilcoxon Rank-Sum tests, W=8346500000, P<0.001, Fig. 3I). They also had higher LD relative to mixed pairs (Wilcoxon Rank-Sum test, W=33823000, P<0.001, Fig. 3I). Thus, high LD between antagonistic sites is not an artefact of unusually low levels of recombination near antagonistic regions, but instead reflects the action of long-term balancing selection.

Taken together, these comparative population genomic analyses demonstrate that the antagonistic allelic variation identified in LH_M is neither recent nor population-specific. To a significant degree, balancing selection maintains antagonistic variation over timescales that extend beyond the extension of the species range out of Africa, more than 10,000 years ago[36].

Discussion

Our study provides the first genome-wide analysis of the identity, function and evolution of sexually antagonistic sequence polymorphisms in fruitflies. Remarkably, we find that genetic variation at antagonistic loci is stably maintained across *D. melanogaster* populations throughout the species' distribution range, indicating that the targets of antagonistic selection have been largely conserved for many millennia[36,40–42]—and several tens of thousands of generations. The geographical stability and low turnover in antagonistic sequence variation implies that adaptive conflict between males and females is rooted in a fundamental aspect of the biology of the sexes and persists even in the face of environmental variation[43]. It is therefore unaffected by the adaptation of populations to the environmental conditions that they encountered during their colonisation of the globe[36,41,44] or the continuous adaptive evolution that occurs within temperate populations over the course of the seasons[23]. More generally, our results supplement a growing body of evidence[23] suggesting that balancing selection can influence patterns of genetic variation on a genome-wide scale, rather than simply a small number of isolated loci[45], as is often assumed[46]. Sexually antagonistic selection should contribute particularly strongly to the build-up of balanced polymorphisms, given that there is abundant evidence for sex-specific selection in nature[4,47] and that sex-specific selection can generate permissive conditions for the evolution of such polymorphisms relative to alternative modes of balancing selection[33,48].

The long-term stability of sexually antagonistic polymorphisms further suggests that the evolutionary constraints on sexual dimorphism inherent in antagonism are difficult to resolve. While we do not find any evidence that there is elevated pleiotropy among genes experiencing ongoing conflict, the persistence of antagonism fits with our finding that antagonistic polymorphisms are highly enriched for missense variants. While antagonistic selection on expression levels can be accommodated by gradual evolution of

sex-specific gene expression[49], adaptive conflicts over coding sequences can only be resolved through a complex multi-step process[50] of gene duplication, sex-specific sub-functionalisation of coding sequences and the evolution of differential expression of the two paralogues. The requirement for gene duplication, in particular, would be expected to constitute a severely limiting barrier for this route towards resolution, as suitable mutation events will be exceedingly rare. This large barrier to resolution, and the resulting stochasticity in which antagonisms will undergo resolution, may help to explain the lack of GO enrichment observed among antagonistic genes.

We find no convincing evidence that the X chromosome is enriched for antagonistic variation, in contradiction to classic theory[3]. This discrepancy could be due to the presence of dominance reversal, where the beneficial allele is dominant in each sex. Such sex specific dominance has recently been documented empirically[13] and is predicted to shift enrichment of antagonism from the X to the autosomes[51]—particularly so if antagonistic loci interact epistatically[52]. Furthermore, the hemiclonal approach might miss low-frequency X-linked antagonistic polymorphisms with recessive fitness effects, as these will rarely be expressed in phenotypic assays. However, while these general effects might explain the lack of X-enrichment, our result also contradicts previous empirical findings obtained in the LH_M population[17] itself, which found that the X chromosome contributed disproportionally to antagonistic fitness variation. The previous study was based on a much smaller sample of genomes, with large uncertainties about the estimated chromosomal contributions. It was also performed more than ten years ago and much closer to the establishment of LH_M as a laboratory population. Accordingly, the discrepancy to our results might in part be explained by stronger genetic drift on the X chromosome relative to autosomes, which could in turn lead to a disproportionate loss of X-linked antagonistic polymorphisms[32].

Taken together, this study addresses a longstanding gap in our understanding of sexual antagonism, and provides a valuable resource from which to further elucidate the origin and resolution of this fundamental evolutionary phenomenon.

Methods

LH_M hemiclones

LH_M is a laboratory-adapted population of *Drosophila melanogaster* that has been maintained under a highly controlled rearing regime since 1996[53]. A random sample of 223 genetic lines was created from the population[20] using a hemiclonal approach[54]. Individuals of each line carry an identical haploid genome comprising the major chromosomes X, 2 and 3. Crosses with flies from custom stocks allows the generation of many replicate individuals—males and females—that carry a line's X-2-3 haplotype alongside a random chromosomal complement from the LH_M population that can be assayed for fitness.

Fitness measurements

Lifetime adult reproductive fitness of males and females of each line was measured using assays designed to mimic the LH_M rearing regime. For male fitness, we measured competitive fertilisation success by setting up competition vials containing 5 hemiclonal males from a given line, 10 competitor *bw* males and 15 virgin *bw* females. After two days, *bw* females were isolated into individual vials containing no additional yeast and left to oviposit for 18 hours. On day 12 post egg-laying, progeny were scored for eye colour. Male fitness was calculated as the proportion of offspring sired by the 5 hemiclonal males (those with wildtype eye-colour), combining progeny data from the 15 oviposition vials. This assay was repeated 5 times in a blocked design; estimates for each line were therefore based on fitness measurements from 25 hemiclonal males.

Female fitness was measured as competitive fecundity. Competition vials containing 5 virgin hemiclonal females from a given line, 10 competitor *bw* females and 15 *bw* males were set up. Two days later, the 5 hemiclonal females were isolated into individual vials and left to oviposit for 18 hours.

These vials were immediately chilled at 4°C and fecundity was measured by counting the number of eggs laid per female. This assay was replicated 5 times in a blocked design; each line estimate therefore measured the fitness of 25 hemiclonal females.

Fitness data were subjected to quality control and pre-processing in preparation for quantitative genetic and association analysis. Male fitness data from competition vials where not all 5 focal males were present at the end of the assay were removed from further analysis. Similarly, we omitted female oviposition vials where fewer than 2 eggs were present (indicating partial sterility or failure to mate) or where the female had died over the course of the assay. For each sex, fitness measurements were then first box-cox transformed to be normally distributed within each block, then scaled and centred. To calculate SNP heritabilities and for association analysis, data from each block were averaged to obtain one fitness estimate for each line and sex.

Quality control of whole-genome sequences

We used previously published whole genome sequences generated from the hemiclonal lines analysed here[20]. Details about DNA extraction, library preparation, sequencing, read processing and SNP calling are provided in the original publication. Prior to the association analysis performed here, further site-level quality filtering steps were performed in vcftools[55] and *PLINK*[26]. First, individual variant calls based on depth<10 and genotype quality<30 were removed. Second, individuals with>15% missing positions were removed. Third, positions with poor genotype information (<95% call rate) across all retained individuals were discarded. Finally, given the relatively small sample size of the dataset as a whole and the low power of an association test for rare variants, we retained only common variants (MAF>0.05) for further analysis. From an initial dataset of 220 hemiclones

containing 1,312,336 SNPs, this yielded a quality-filtered dataset of 765,980 SNPs from 203 hemiclones.

To detect outliers, we examined LH_M's population structure using principal components analysis (PCA). Overlapping SNP positions from the 203 LH_M genomes and from an outgroup population (*Drosophila* Genetic Reference Panel, or DGRP[34]) consisting of 205 whole-genome sequenced individuals were used as input to construct a genetic similarity matrix. This set of SNPs was pruned for linkage disequilibrium (LD) such that no two SNPs with r^2>0.2 within 10Kb remained. The leading PC axes were extracted in LDAK ("Linkage-Disequilibrium Adjusted Kinships"[56]). After removal of one outlier (see Sup. Fig. 1A), the final dataset used for association analysis contained 202 individuals and 765,764 SNPs.

Heritability analyses

We estimated the variance-covariance matrix for fitness and sex-specific residual variances by fitting a model using MCMCglmm[57] implemented in R. Specifically, we fitted the model $Y_{ijk} = X_{ij} + \varepsilon_{ijk}$, where Y_{ijk} is the scaled and centred fitness of individual k of genotype j and sex i, X_{ij} is the sex-specific random effect of genotype j in sex i, and ε_{ijk} describes the sex-, genotype- and individual-specific residual. The genotypic fitness effects in males and females follow a bivariate normal distribution $X_{ij} \sim N(0, \mathrm{G})$, where

$$G = \begin{pmatrix} \sigma^2_{G,m} & Cov_{G,mf} \\ Cov_{G,mf} & \sigma^2_{G,f} \end{pmatrix}$$

is the genetic variance-covariance matrix across sexes (composed of male and female additive genetic variances $\sigma^2_{G,m}$ and $\sigma^2_{G,f}$ and the intersexual genetic covariance $Cov_{G,mf}$). Residuals follow a normal distribution $\varepsilon_{ijk} \sim N(0, \sigma^2_{R,i})$, where $\sigma^2_{R,i}$ is the sex-specific residual variance, and are assumed to be uncorrelated across sexes.

From these variance estimates, we calculated male and female heritabilities of fitness as $h_i^2 = 2\sigma_{G,i}^2/(\sigma_{G,i}^2 + \sigma_{R,i}^2)$, where the subscript i indicates either male or female. The factor 2 in the heritability calculation reflects the fact that with the hemiclonal approach, individuals assayed share half their genetic material (the hemizygous hemiclonal genome). The intersexual genetic correlation was calculated as $r_{mf} = Cov_{G,mf}/\left(\sqrt{\sigma_{G,m}^2}\sqrt{\sigma_{G,f}^2}\right)$. The quantitative genetic parameters h_m^2, h_f^2 and r_{mf} were calculated for each sample from the Monte Carlo Markov chain. From these series of values we obtained point estimates (averages) and 95% credible interval (using the function HPDintervals).

As a complementary approach, we estimated the SNP heritability (h_{SNP}^2) of male and female fitness in LDAK[56]. This approach uses Restricted Maximum Likelihood (REML)[58] to fit a linear mixed model that expresses the vector of phenotypes Y as a function genome-wide SNP genotypes, treated as random effects:

$$Y \sim N(0, \sigma_{SNP}^2 K + \sigma_e^2 I)$$

where K the kinship matrix, σ_{SNP}^2 a vector of additive genetic variances for each SNP, σ_e^2 the vector of residual variances and I an individual identity matrix. SNP heritability is then estimated as $h_{SNP}^2 = \sigma_{SNP}^2/(\sigma_{SNP}^2 + \sigma_e^2)$.

LDAK corrects for local linkage when calculating SNP heritabilities to avoid inflation of h_{SNP}^2 in clusters of linked sites that otherwise arises because several SNPs tag the same causal polymorphism. SNPs are weighted inversely proportional to their local linkage, such that SNPs in high LD contribute less to h_{SNP}^2 than SNPs in low LD. This model has been shown to substantially improve heritability estimates across a wide range of traits[24]. LDAK also allows to set the parameter α that determines how SNPs are weighted by their MAF (as MAF$^{\alpha}$) when calculating the kinship matrix K. We used the default of α=−0.25 which provides a steeper relationship between

MAF and h^2_{SNP} than the value of -1 that is frequently used in studies on humans. Significance of h^2_{SNP} estimates was assessed by permuting phenotype labels 1,000 times, re-calculating h^2_{SNP} on each permutation as above, and calculating the number of permuted estimates which exceeded the observed.

Quantification and association analysis of sexual antagonism

To identify loci underlying sexual antagonism, we defined an antagonism index (see main text, Fig. 1A). We calculated its SNP heritability ('antagonistic h^2_{SNP}') in LDAK, following the same procedure and settings as those for estimating sex-specific SNP heritabilities.

We performed a GWAS by applying a linear mixed model to test the effect of allelic variants at each SNP on the antagonism index, while including the kinship matrix as a random effect to account for the heritable portion of genetic variation attributable to kinship between individuals. This approach has been shown to effectively control the false positive rate and increase power to detect true associations in samples with moderate degrees of population structure and close relatedness, such as LH_M[22,59]. The GWAS was implemented in LDAK (settings as above) and a Wald χ^2 test was used to generate P-values for each position.

The genomic inflation factor[60] of λ_{median}=0.967 (calculated as median χ^2_{obs} / median χ^2_{exp} in GenABEL[61] in R) suggests that genetic confounding has been well controlled in our GWAS. This is corroborated by the fact that the distribution of P-values when permuting the phenotype labels (100 times and applying the same linear mixed model) was not enriched for low P-values (such a pattern would be expected if residual genetic confounding remained in our sample).

Defining candidate antagonistic SNPs and regions

We corrected for multiple testing using a False Discovery Rate (FDR) approach and converted P-values into Q-values. We defined antagonistic SNPs as sites with FDR Q-values<0.3 and non-antagonistic SNPs as sites with Q-value≥0.3.

For analyses which consider larger genomic regions (windows), we ran a set-based association test implemented in LDAK (options using '–calc-genes-reml', 'ignore-weights YES' and α=−0.25). The test calculates set-wide h^2_{SNP} via REML, corrects for local relatedness using the predictors in each window, and computes a P-value using a likelihood ratio test (LRT). The sets we used were 1000bp windows (500bp step) defined according to *Drosophila* Reference 5 genome coordinates, and subsequently converted (using UCSC's liftOver tool[62]) to Release 6 coordinates. This was a necessary step, as publicly available polymorphism data was mapped to Release 5 of the *D. melanogaster* genome, whereas the GWAS data was mapped to Release 6. We then calculated window-based Q-values from the LRT P-values and defined antagonistic windows as those with Q-value<0.1.

Genomic distribution of antagonistic SNPs

To estimate the number of independent antagonistic regions, we performed LD-clumping in *PLINK*. We used a significance threshold of 0.00093 for the index SNP (the maximum, least significant, P-value across all antagonistic SNPs), and clustered ('clumped') neighbouring antagonistic SNPs by specifying an r^2 threshold of 0.4 and a distance threshold of 10Kb.

We also quantified the clustering by calculating the median distance between all pairs of adjacent antagonistic SNPs across chromosome arms. We did this separately for the autosomes and X chromosome, to accommodate for the lower SNP density on the X chromosome. We tested for significant clustering by using a permutation test, where antagonistic/non-antagonistic labels was permuted among all SNPs, distances between

adjacent SNPs labelled as 'antagonistic' after permutation were recalculated as before and the median distance recorded. This process was repeated 1,000 times in order to generate a null distribution of median distances. Significance of clustering among true antagonistic SNPs was calculated as the proportion of median distances in the null distribution that were lower than or equal to the true median distance.

To examine the proportional contribution of autosomal and X-linked antagonistic variants to total h^2_{SNP}, we used two complementary methods. First, we partitioned the genome into X chromosome and autosome subsets, and calculated h^2_{SNP} via REML in LDAK each subset in turn (default parameters; α=−0.25). The observed proportion of h^2_{SNP} contributed by each compartment was then compared to the expected proportion (*i.e.*, the fraction of LD-weighted predictors belonging to each compartment). We tested whether the two compartments contributed significantly more h^2_{SNP} than expected using a two-sample Z-test. Second, we compared the proportion of antagonistic SNPs (Q-value<0.3) to the proportion of all SNPs mapping to each chromosomal compartment, using Z-tests. The under- or over-representation of antagonistic SNPs (deficit or excess of antagonistic compared to all SNPs) in each compartment is therefore unaffected by differences in SNP density between chromosome arms, such as the lower density on the X chromosome.

Functional analyses of antagonistic loci

We used the variant effect predictor (Ensembl VEP[63]) to map SNPs to functional categories. We partitioned total antagonistic h^2_{SNP} into functional subsets, and estimated the observed proportion of h^2_{SNP} contributed by each subset using REML in LDAK (settings as above). We then used a permutation test to compare observed and expected h^2_{SNP} for each functional category, where we shifted genome-wide annotations to a random starting

point along a 'circular genome'. This procedure breaks the relationship between each SNP and its annotation while preserving the order of annotations and their associated LD structure[64]. h^2_{SNP} was re-calculated via REML for each of 1,000 permuted datasets and two-tailed P-values determined as the sum of permuted estimates with more extreme absolute values than the observed. As a complementary approach, we compared the proportion of antagonistic SNPs to the proportion of all SNPs mapping to each functional category. We then assessed enrichment for each functional category in turn using Z-tests.

We also used the VEP to map SNPs to genes. We included extended gene regions (+/- 5kb of gene coordinates, VEP default) in our gene definition. To gain preliminary insights into the functions of antagonistic genes we used the Gorilla[65] Gene Ontology tool, with FDR correction for multiple testing across GO terms. All genes covered in the final SNP dataset were used as the background set.

To examine the relationship between antagonistic genes and sex-biased gene expression we used the Sebida online database[66] to annotate genes as having either sex-biased or unbiased expression profiles (meta-class identifier). We then used a χ^2 test to compare the sex-biased expression status of antagonistic and non-antagonistic genes. We additionally examined the quantitative degree of sex-bias using this same dataset. We took the absolute value of the log2 transformed 'M_F' bias variable, such that large values indicate more extreme sex bias in expression irrespectively of whether this bias is towards males or females. We compared the distributions of this variable between antagonistic and non-antagonistic genes using a Wilcoxon Rank-Sum test.

To assess the degree of overlap between antagonistic genes identified here and those associated with sexually antagonistic expression patterns in a previous study[16], we included only genes covered in both datasets, and

only those genes in both datasets that were adult-expressed. To determine whether genes were adult-expressed we used the *Drosophila* gene expression atlas (FlyAtlas[67]). Conservatively, we considered a gene 'adult-expressed' if its transcript was detected as present in at least one library of one adult-derived sample. We then used a χ^2 test to assess the degree of overlap between the datasets.

We used the tissue-specificity index (τ) to compare pleiotropy between antagonistic and non-antagonistic genes. We used gene expression data from FlyAtlas[67] to get average expression values for each gene and in each tissue and then calculated τ as:

where is the proportional expression level of the gene in tissue and is the number of tissues. We compared values of τ for antagonistic and non-antagonist genes using a Wilcoxon rank-sum test.

As an additional proxy for pleiotropy we examined the number of protein-protein interactions (PPIs) between antagonistic and non-antagonistic genes. We used the physical interactions table from FlyBase[68] to summarise the total number of PPIs for all genes and then compared candidate and non-candidate genes using a general linear model (GLM) with quasipoisson error structure to account for overdispersion.

Comparative population genomic data

To analyse SNP polymorphism outside the LH_M population, we used publicly available population genomic data from three wild *D. melanogaster* populations. The first is an introduced population from North America (*Drosophila* Genetic Reference Panel[34,35], denoted DGRP: 205 whole-genome sequences derived from inbred lines). The two others come from *D. melanogaster*'s ancestral distribution range in sub-Saharan Africa (Zambia–ZI: 197 whole-genome sequences derived from haploid embryos; South

African–SA: 118 whole-genome sequences derived from inbred lines, combines data from sub-populations 'SD' and 'SP', which have very low population differentiation[38]).

All genome sequences were downloaded as FASTA files from the Drosophila Genome Nexus website (www.johnpool.net/genomes.html). These files had been generated following standardised alignment and quality filtering steps[37] and were further quality-filtered for admixture and identity-by-descent using scripts provided on the Genome Nexus website. We used snp-sites[69] to call SNPs and convert the multiple sequence alignments to vcf format. Allele frequencies in the three populations were calculated using vcftools. We further excluded tri-allelic and poorly covered sites (call rate<20).

SNP-based analyses of balancing selection

To test whether antagonistic sites are associated with signatures of balancing selection, we closely followed the approach of Turchin et al.[70] and looked for an increased minor allele frequency (MAF) at antagonistic relative to non-antagonistic sites (as identified in LH_M) in the three comparison populations (DGRP, ZI, SA). By focussing on the contrast between classes of SNPs, we ensured that demographic differences between populations did not confound our analyses.

We first LD-pruned the LH_M dataset by clumping (in *PLINK*) to avoid pseudo-replication due to correlations between SNPs. For antagonistic sites, we used the 226 index SNPs identified in the previous clumping. For non-antagonistic sites, a non-antagonistic SNP was randomly chosen as an index SNP and clumped by clustering all SNPs within 10kb with $r^2>0.4$. Pruning in this manner reduced the original dataset of 765,764 SNPs to 36,319 “LD-independent” SNPs. For each of these SNPs, we then estimated MAFs in each comparison population. We assigned MAF=0 to sites which were monomorphic in a comparison population and those where a comparison

population was polymorphic for variants other than those segregating at that site in the LH_M.

We then used this LD-independent dataset to compare MAF between antagonistic and non-antagonistic SNPs. We did this using a Monte Carlo approach where, 1,000 times, we paired the 226 antagonistic SNPs with 226 randomly drawn non-antagonistic "control" SNPs. The latter were carefully frequency-matched to the 226 antagonistic SNPs. The matching procedure first corrected LH_M MAF for 'linked selection'[71] by taking the residuals of a linear regression of LH_M MAF on estimates of linked selection. These estimates quantify local recombination rates and proximity to functional sequences in *D. melanogaster*. They thereby account for factors that affect polymorphism along the genome, such as background selection and selective sweeps. We then drew sets of 226 non-antagonistic SNPs to match the residual LH_M MAF distribution of the 226 antagonistic SNPs and for each set calculated the mean MAF in the comparison population. The 1,000 sets generated in this way provided a null distribution of MAFs for non-antagonistic sites in each comparison population. P-values for deviations in polymorphism between antagonistic and non-antagonistic sites were then calculated by comparing, in each population, the mean MAF of the 226 antagonistic SNPs to the null MAF distribution.

A second analysis used the same LD-independent dataset but considered the whole spectrum of P-values, rather than a binary split of SNPs into antagonistic/non-antagonistic categories. To this end, we performed binning in two dimensions, by residual LH_M MAF (20 quantiles) and P-values (100 quantiles). We then drew one SNP from each of these MAF/P-value bins (2,000 SNPs in total), recorded the MAF for each in the comparison population of interest, and finally correlated these MAF values with P-values of the associated SNPs in the LH_M using a Spearman's rank correlation. Under the hypothesis of antagonism-mediated balancing

selection, SNPs with low P-values should tend to have higher MAFs in the population under consideration than SNPs with high P-values.

Window-based analyses of balancing selection

We performed genome-wide sliding window analyses (1,000bp windows, 500bp step size) to investigate regional signatures of balancing selection. Tajima's D, which compares SNP polymorphism (nucleotide diversity, π) to SNP abundance (Watterson's estimator, θ_W), was compared for windows defined as antagonistic (Q-value<0.1) or non-antagonistic (Q-value≥0.1) from the set-based analysis (see section 'Defining candidate antagonistic SNPs and regions'). Under the hypothesis that antagonism generates balancing selection, Tajima's D is expected to be elevated in antagonistic windows. We calculated Tajima's D for each comparison population using PopGenome in R[72]. As in SNP-based analyses, we incorporated estimates of linked selection[71] (estimated in 1,000bp windows) by calculating the residuals of a regression of Tajima's D on estimates of linked selection. Since estimates of linked selection were not available for windows on the X chromosome, we instead used estimates of recombination rate on this chromosome[73]. We then used a generalised linear model (GLM), assuming Gaussian error structure, to compare residual Tajima's D between antagonistic and non-antagonistic windows.

We also tested for another signature of balancing selection, reduced population differentiation. Measures such as F_{ST} are often considered problematic because they do not correct for the dependency of F_{ST} on local levels of polymorphism[74]. However, the availability of genome-wide estimates of linked selection in *D. melanogaster* allowed us to incorporate this confounding variable explicitly. We therefore estimated F_{ST} over windows, using PopGenome, correcting F_{ST} for linked selection in a way analogous to that used for Tajima's D. Since the distribution of F_{ST} values is

not normally distributed, we contrasted residual F_{ST} between antagonistic and non-antagonistic windows using Wilcoxon Rank-Sum tests.

Linkage-based analyses of balancing selection

We examined the extent to which antagonistic haplotypes are selectively maintained by investigating whether antagonistic SNPs have unusually high linkage disequilibrium (LD) in the ZI population, the population that is most distant from LH_M and where levels of LD between antagonistic SNPs should be weakest in the absence of long-term balancing selection. Thus, for all SNPs situated within 1000bp of one another in ZI and which were also covered in LH_M (*i.e.*, SNPs which could be inferred to be either antagonistic or non-antagonistic), we calculated pairwise LD (r^2) in PLINK. We then compared r^2 values between pairs of antagonistic SNP and two control pairs: non-antagonistic pairs, and 'mixed' pairs (antagonistic/non-antagonistic). Comparing pairs of antagonistic SNPs to the mixed pairs allowed us to consider only SNPs located close to an antagonistic SNP, thus effectively controlling for possible non-random distributions of antagonistic pairs and non-antagonistic pairs with respect to genome-wide recombination rates.

To test for significant differences in LD between antagonistic pairs and the two control pairs, we modelled variation in r^2 as a declining exponential function of chromosomal distance, and assessed differences in residual r^2 (once distance was regressed out) using Wilcoxon Rank-Sum tests.

Statistical software

All statistical analyses were carried out in RStudio (version 1.0.136[75]).

Data availability

Phenotypic data will be deposited on Dryad prior to publication.

Population genomic data from LH_M is available at https://zenodo.org/record/159472.
Population genomic data from the DGRP, ZI and SA is available at http://www.johnpool.net/genomes.html.

Code availability

Analysis code is available on request.

References

1. Bonduriansky, R. & Chenoweth, S. F. Intralocus sexual conflict. Trends Ecol. Evol. 24, 280–288 (2009).
2. Van Doorn, G. S. Intralocus sexual conflict. Ann. N. Y. Acad. Sci. 1168, 52–71 (2009).
3. Rice, W. R. Sex chromosomes and the evolution of sexual dimorphism. Evolution 38, 735–742 (1984).
4. Cox, R. M. & Calsbeek, R. Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. Am. Nat. 173, 176–187 (2009).
5. Pennell, T. M. & Morrow, E. H. Two sexes, one genome: The evolutionary dynamics of intralocus sexual conflict. Ecol. Evol. 3, 1819–1834 (2013).
6. Mokkonen, M. et al. Negative frequency-dependent selection of sexually antagonistic alleles in Myodes glareolus. Science 334, 972–974 (2011).
7. Stulp, G., Kuijper, B., Buunk, A. P., Pollet, T. V & Verhulst, S. Intralocus sexual conflict over human height. Biol. Lett. 8, 976–8 (2012).
8. Tarka, M., Åkesson, M., Hasselquist, D. & Hansson, B. Intralocus sexual conflict over wing length in a wild migratory bird. Am. Nat. 183, 62–73 (2014).

9. Svensson, E. I., McAdam, A. G. & Sinervo, B. Intralocus sexual conflict over immune defense, gender load, and sex-specific signaling in a natural lizard population. Evolution 63, 3124–3135 (2009).
10. Rice, W. R. Sexually antagonistic genes: experimental evidence. Science 256, 1436–1439 (1992).
11. Berger, D., Berg, E. C., Widegren, W., Arnqvist, G. & Maklakov, A. A. Multivariate intralocus sexual conflict in seed beetles. Evolution 68, 3457–3469 (2014).
12. Roberts, R. B., Ser, J. R. & Kocher, T. D. Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes. Science 326, 998–1001 (2009).
13. Barson, N. J. et al. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. Nature 528, 405–408 (2015).
14. Delph, L. F. et al. Environment-dependent intralocus sexual conflict in a dioecious plant. New Phytol. 192, 542–552 (2011).
15. Kidwell, J. F., Clegg, M. T., Stewart, F. M. & Prout, T. Regions of stable equilibria for models of differential selection in the two sexes under random mating. Genetics 85, 171–183 (1977).
16. Innocenti, P. & Morrow, E. H. The sexually antagonistic genes of Drosophila melanogaster. PLoS Biol. 8, e1000335 (2010).
17. Lucotte, E. A., Laurent, R., Heyer, E., Ségurel, L. & Toupance, B. Detection of allelic frequency differences between the sexes in humans: a signature of sexually antagonistic selection. Genome Biol. Evol. 8, 1489–1500 (2016).
18. Chippindale, A. K., Gibson, J. R. & Rice, W. R. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in Drosophila. Proc. Natl. Acad. Sci. U. S. A. 98, 1671–1675 (2001).

19. Rice, W. R. et al. Inter-locus antagonistic coevolution as an engine of speciation: assessment with hemiclonal analysis. Proc. Natl. Acad. Sci. U. S. A. 102, 6527–6534 (2005).
20. Gilks, W. P., Pennell, T. M., Flis, I., Webster, M. T. & Morrow, E. H. Whole genome resequencing of a laboratory-adapted Drosophila melanogaster population sample. F1000Research 5, 2644 (2016).
21. Locke, A. E. et al. Genetic studies of body mass index yield new insights for obesity biology. Nature 518, 197–206 (2015).
22. Price, A. L., Zaitlen, N. A., Reich, D. & Patterson, N. New approaches to population stratification in genome-wide association studies. Nature Reviews Genetics 11, 459–463 (2010).
23. Bergland, A. O., Behrman, E. L., O'Brien, K. R., Schmidt, P. S. & Petrov, D. A. Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in Drosophila. PLoS Genet. 10, e1004775 (2014).
24. Speed, D., Cai, N., Johnson, M. R., Nejentsev, S. & Balding, D. J. Reevaluation of SNP heritability in complex human traits. Nat. Genet. 49, 986–992 (2017).
25. Yang, J. et al. Genome partitioning of genetic variation for complex traits using common SNPs. Nat. Genet. 43, 519–525 (2011).
26. Purcell, S. et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575 (2007).
27. Langley, C. H. et al. Genomic variation in natural populations of Drosophila melanogaster. Genetics 192, 533–598 (2012).
28. Mank, J. E., Hultin-Rosenberg, L., Zwahlen, M. & Ellegren, H. Pleiotropic constraint hampers the resolution of sexual antagonism in vertebrate gene expression. Am. Nat. 171, 35–43 (2008).
29. Yanai, I. et al. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. Bioinformatics 21, 650–659 (2005).

30. He, X. & Zhang, J. Toward a molecular understanding of pleiotropy. Genetics 173, 1885–1891 (2006).
31. Gavrilets, S. & Rice, W. R. Genetic models of homosexuality: generating testable predictions. Proceedings of the Royal Society B: Biological Sciences 273, 3031–3038 (2006).
32. Mullon, C., Pomiankowski, A. & Reuter, M. The effects of selection and genetic drift on the genomic distribution of sexually antagonistic alleles. Evolution 66, 3743–3753 (2012).
33. Connallon, T. & Clark, A. G. Balancing selection in species with separate sexes: Insights from Fisher's geometric model. Genetics 197, 991–1006 (2014).
34. Mackay, T. F. C. et al. The Drosophila melanogaster Genetic Reference Panel. Nature 482, 173–178 (2012).
35. Huang, W. et al. Natural variation in genome architecture among 205 Drosophila melanogaster Genetic Reference Panel lines. Genome Res. 24, 1193–1208 (2014).
36. Duchen, P., Živković, D., Hutter, S., Stephan, W. & Laurent, S. Demographic inference reveals African and European admixture in the North American Drosophila melanogaster population. Genetics 193, 291–301 (2013).
37. Lack, J. B. et al. The Drosophila Genome Nexus: A population genomic resource of 623 Drosophila melanogaster genomes, including 197 from a single ancestral range population. Genetics 199, 1229–1241 (2015).
38. Lack, J. B., Lange, J. D., Tang, A. D., Corbett-Detig, R. B. & Pool, J. E. A thousand fly genomes: An expanded Drosophila Genome Nexus. Mol. Biol. Evol. 33, 3308–3313 (2016).
39. Kelly, J. K. A test of neutrality based on interlocus associations. Genetics 146, 1197–1206 (1997).
40. Lachaise, D. et al. Historical biogeography of the Drosophila melanogaster species subgroup. Evol. Biol. 22, 159–225 (1988).

41. Li, H. & Stephan, W. Inferring the demographic history and rate of adaptive substitution in Drosophila. PLoS Genet. 2, e1004775 (2006).
42. Thornton, K. & Andolfatto, P. Approximate Bayesian inference reveals evidence for a recent, severe bottleneck in a Netherlands population of Drosophila melanogaster. Genetics 172, 1607–1619 (2006).
43. Connallon, T. & Hall, M. D. Genetic correlations and sex-specific adaptation in changing environments. Evolution 70, 2186–2198 (2016).
44. Pool, J. E. et al. Population Genomics of Sub-Saharan Drosophila melanogaster: African Diversity and Non-African Admixture. PLoS Genet. 8, e1003080 (2012).
45. Charlesworth, D. Balancing selection and its effects on sequences in nearby genome regions. PLoS Genet. 2, e64 (2006).
46. Asthana, S., Schmidt, S. & Sunyaev, S. A limited role for balancing selection. Trends in Genetics 21, 30–32 (2005).
47. Poissant, J., Wilson, A. J. & Coltman, D. W. Sex-specific genetic variance and the evolution of sexual dimorphism: A systematic review of cross-sex genetic correlations. Evolution 64, 97–107 (2010).
48. Connallon, T. & Clark, A. G. Evolutionary inevitability of sexual antagonism. Proc. R. Soc. B Biol. Sci. 281, 2013–2123 (2014).
49. Connallon, T. & Clark, A. G. Sex linkage, sex-specific selection, and the role of recombination in the evolution of sexually dimorphic gene expression. Evolution 64, 3417–3442 (2010).
50. Stewart, A. D., Pischedda, A. & Rice, W. R. Resolving intralocus sexual conflict: genetic mechanisms and time frame. J. Hered. 101, S94-9 (2010).
51. Fry, J. D. The genomic location of sexually antagonistic variation: Some cautionary comments. Evolution 64, 1510–1516 (2010).
52. Arnqvist, G., Vellnow, N. & Rowe, L. The effect of epistasis on sexually antagonistic genetic variation. Proc. R. Soc. B Biol. Sci. 281, 20140489 (2014).

53. Rice, W. R. et al. Inter-locus antagonistic coevolution as an engine of speciation: assessment with hemiclonal analysis. Proc. Natl. Acad. Sci. U. S. A. 102, 6527–6534 (2005).
54. Abbott, J. K. & Morrow, E. H. Obtaining snapshots of genetic variation using hemiclonal analysis. Trends Ecol. Evol. 26, 359–368 (2011).
55. Danecek, P. et al. The variant call format and VCFtools. Bioinformatics 27, 2156–2158 (2011).
56. Speed, D., Hemani, G., Johnson, M. R. & Balding, D. J. Improved heritability estimation from genome-wide SNPs. Am. J. Hum. Genet. 91, 1011–1021 (2012).
57. Hadfield, J. D. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. J. Stat. Softw. 33, 1–22 (2010).
58. Yang, J. et al. Common SNPs explain a large proportion of heritability for human height. Nat. Genet. 42, 565–569 (2010).
59. Astle, W. & Balding, D. J. Population structure and cryptic relatedness in genetic association studies. Stat. Sci. 24, 451–471 (2009).
60. Yang, J. et al. Genomic inflation factors under polygenic inheritance. Eur. J. Hum. Genet. 19, 807–812 (2011).
61. Aulchenko, Y. S., Stephan, R., Isaacs, A. & van Duijn, C. M. GenABEL: an R library for genome-wide association analysis. Bioinformatics 23, 1294–1296 (2007).
62. Casper, J. et al. The UCSC Genome Browser database: 2018 update. Nucleic Acids Res. 46, D762–D769 (2018).
63. McLaren, W. et al. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics 26, 2069–2070 (2010).
64. Cabrera, C. P. et al. Uncovering networks from genome-wide association studies via circular genomic permutation. G3 Genes, Genomes, Genet. 2, 1067–1075 (2012).

65. Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10, 48 (2009).
66. Gnad, F. & Parsch, J. Sebida: a database for the functional and evolutionary analysis of genes with sex-biased expression. Bioinformatics 22, 2577–2579 (2006).
67. Chintapalli, V. R., Wang, J. & Dow, J. A. T. Using FlyAtlas to identify better Drosophila melanogaster models of human disease. Nat. Genet. 39, 715–720 (2007).
68. Gramates, L. S. et al. FlyBase at 25: Looking to the future. Nucleic Acids Res. 45, D663–D671 (2017).
69. Keane, J. A. et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. Microb. Genomics 2, e000056 (2016).
70. Turchin, M. C. et al. Evidence of widespread selection on standing variation in Europe at height-associated SNPs. Nat. Genet. 44, 1015–1019 (2012).
71. Elyashiv, E. et al. A genomic map of the effects of linked selection in Drosophila. PLoS Genet. 12, e1006130 (2016).
72. Pfeifer, B., Wittelsburger, U., Ramos-Onsins, S. E. & Lercher, M. J. PopGenome: An efficient swiss army knife for population genomic analyses in R. Mol. Biol. Evol. 31, 1929–1936 (2014).
73. Comeron, J. M., Ratnappan, R. & Bailin, S. The many landscapes of recombination in Drosophila melanogaster. PLoS Genet. 8, e1002905 (2012).
74. Cruickshank, T. E. & Hahn, M. W. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. Mol. Ecol. 23, 3133–3157 (2014).
75. RStudio Team. RStudio: Integrated Development for R. RStudio. RStudio, Inc., Boston, MA (2015). Available at: https://www.rstudio.com/.

Acknowledgements

We are grateful to Aida Andrés for her suggestions on the analysis of balancing selection and to Laurent Keller, Andrew Pomiankowski and François Balloux for helpful comments on earlier versions of the manuscript. FR was funded by a London NERC DTP studentship (NERC grant NE/L002485/1), MSH was funded by a UCL IMPACT PhD studentship, EHM, TMP, IF and FCI by a European Research Council Grant (280632) and a Royal Society University Research Fellowship and MR and KF by a Natural Environment Research Council research grant (NE/G0189452/1).

Author contributions

EHM, MR, KF, MSH and FR conceived the project; TMP, IF and EHM conducted phenotypic experiments; WPG generated sequencing data; FR, MSH, MR and FCI analysed the data; FR, MR, MSH and KF wrote the manuscript.

Competing financial interests

The authors declare no competing interests.

Figures

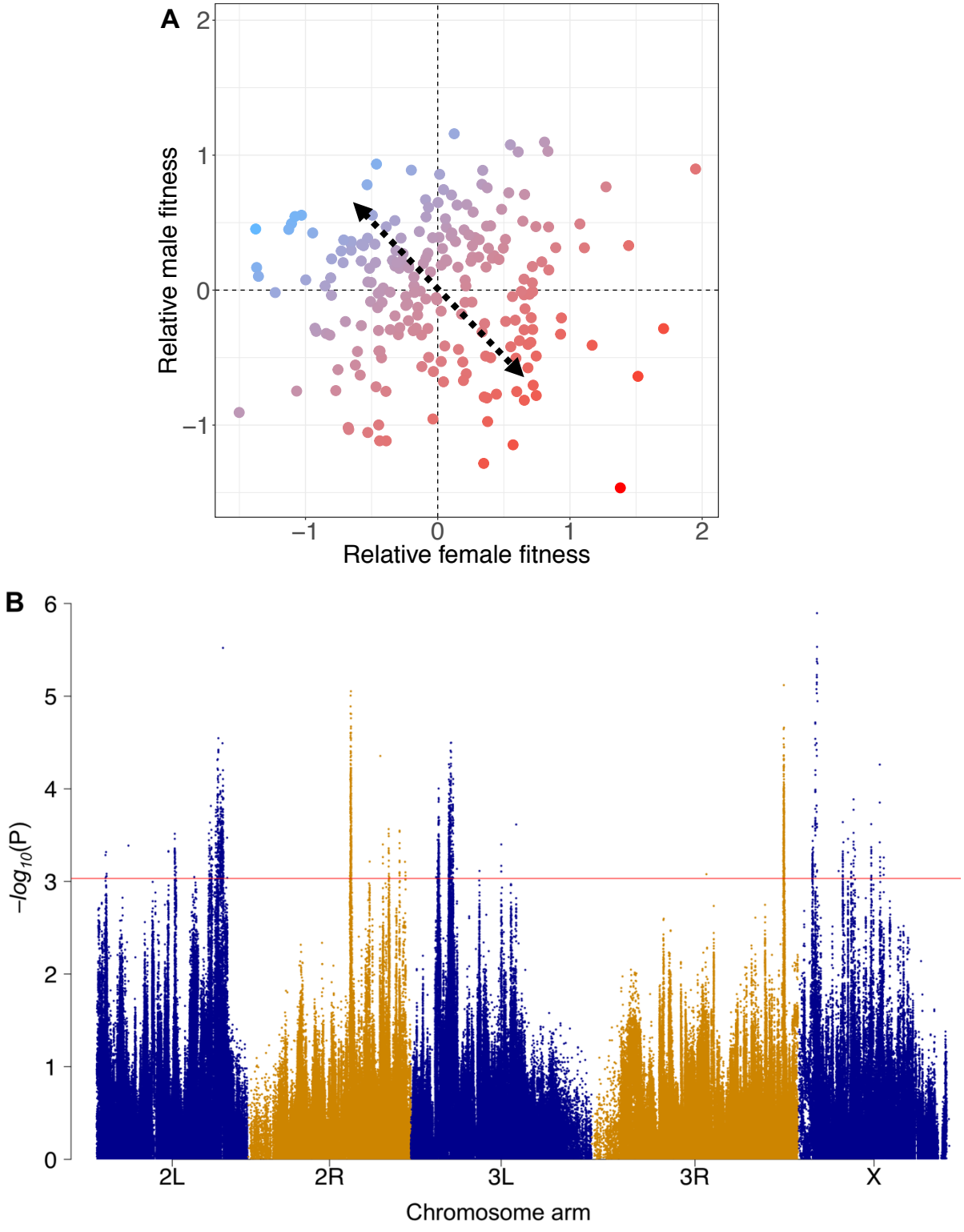
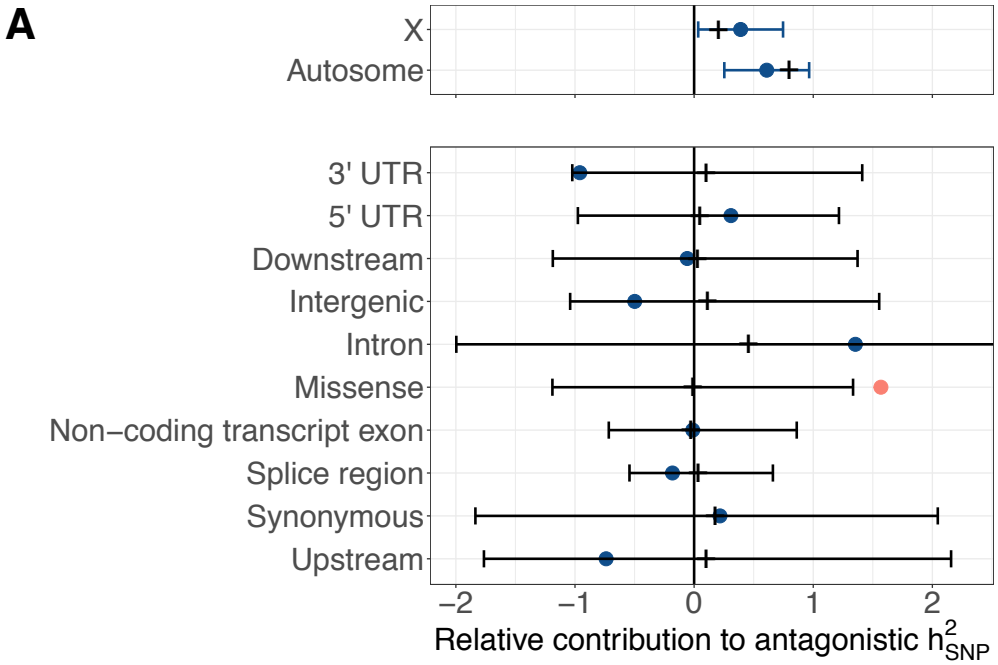

A
Relative male fitness
Relative female fitness
2
1
0
−1
−1
0
1
2
B
−log10(P)
6
5
4
3
2
1
0
2L
2R
3L
3R
X
Chromosome arm

Figure 1. Genome-wide association mapping of sexual antagonism. A. Relative male and female lifetime reproductive fitness estimates for 223 *D. melanogaster* hemiclonal lines. Fitness measures have been scaled to be normally distributed. Colours denote each line's antagonism index, *i.e.* their position along a spectrum (arrows) ranging from male-beneficial, female-detrimental fitness effects (blue), to female-beneficial, male-detrimental effects (red). **B.** Association of each SNP with the antagonism index along the five major *D. melanogaster* chromosome arms, presented as a Manhattan plot where each point represents the $-log_{10}$(P) value from a Wald χ^2 association test. Colours denote chromosome arms, the horizontal line represents the Q-value cut-off (0.3) used to define candidate antagonistic SNPs.


A
X
Autosome
3' UTR
5' UTR
Downstream
Intergenic
Intron
Missense
Non−coding transcript exon
Splice region
Synonymous
Upstream
−2
−1
0
1
2
Relative contribution to antagonistic h2SNP

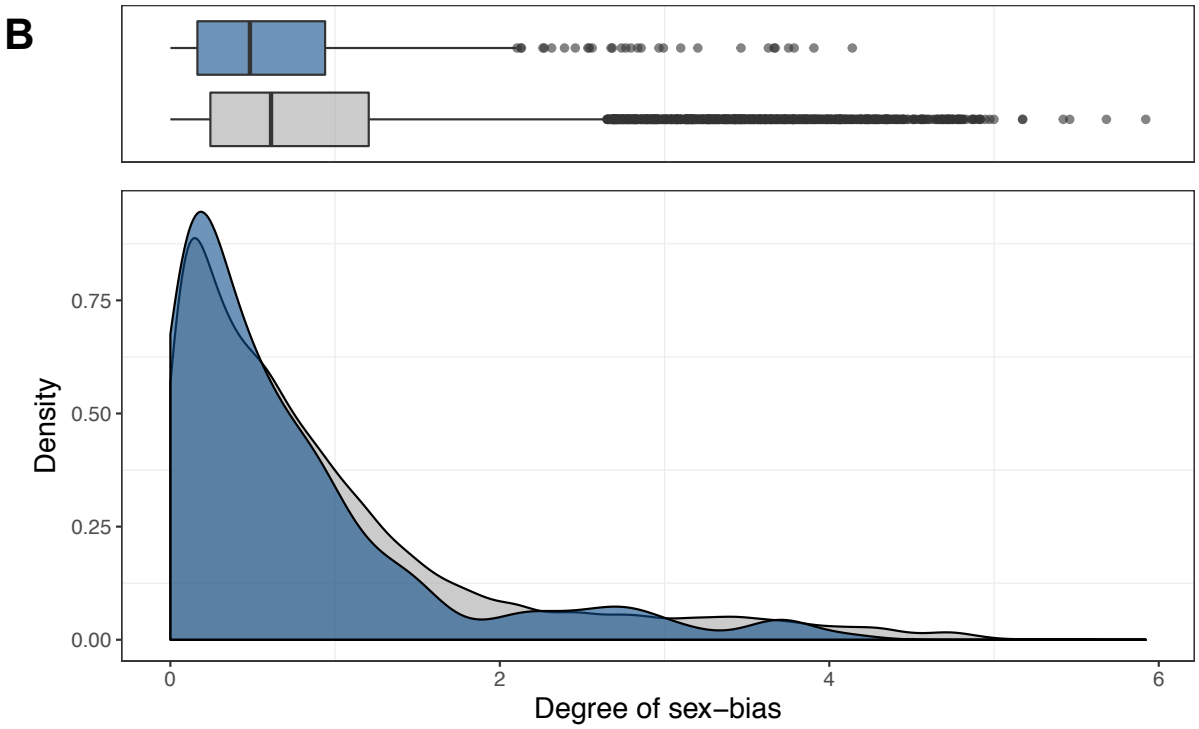

B
0.75
0.50
0.25
0.00
Density
0
2
4
6
Degree of sex−bias

Figure 2. Genomic distribution and functional characteristics of antagonistic variants. A. Relative contribution (proportional share) of different chromosomal compartments (top) and functional categories (bottom) to total antagonistic SNP heritability (h^2_{SNP}). Dots represent estimated h^2_{SNP} contributions (±95% CI, for chromosomal compartments), with colours indicating significant under or over-representation (red: P-value<0.05; blue: P-value>0.05). Expected h^2_{SNP} contributions are presented either as black notches (fixed values for chromosomal compartments) or mean±95% CI of the empirical null distribution computed through permutation (for functional categories). See Methods for additional details. **B.** Distributions of the absolute degree of sex-biased expression for antagonistic (blue) and non-antagonistic (grey) genes.

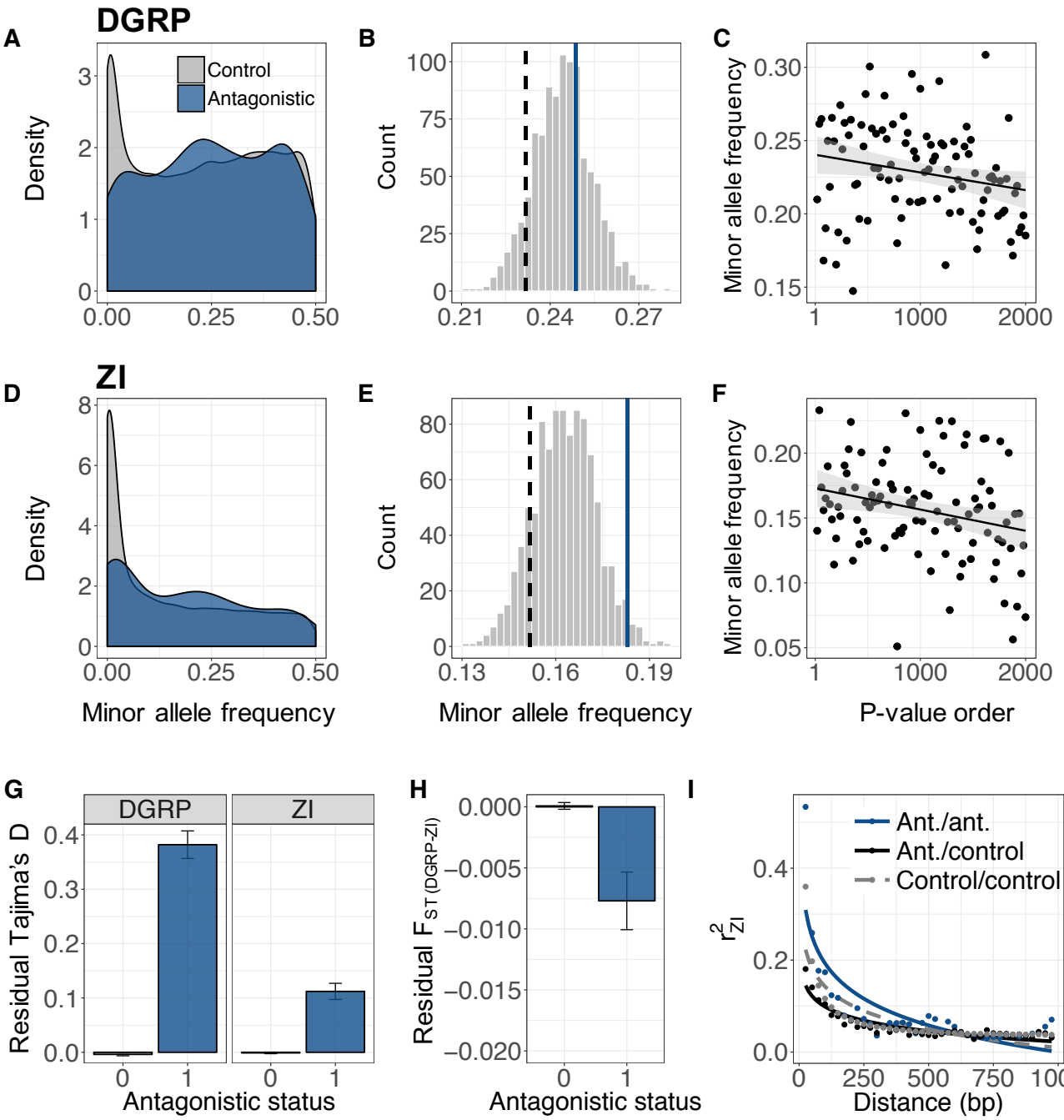

A
DGRP
Control
Antagonistic
Density
B
Count
C
Minor allele frequency
D
ZI
Density
Minor allele frequency
E
Count
Minor allele frequency
F
Minor allele frequency
P-value order
G
DGRP
ZI
Residual Tajima's D
Antagonistic status
H
Residual $F_{ST\ (DGRP-ZI)}$
Antagonistic status
I
Ant./ant.
Ant./control
Control/control
r^2_{ZI}
Distance (bp)

Figure 3. Signatures of balancing selection associated with antagonistic variants in two independent populations (DGRP and ZI; see Sup. Fig. 5 for SA population). A,D. Spectra of raw minor allele frequencies (MAF) for LD-pruned antagonistic (blue) and non-antagonistic ('control', grey) SNPs in the DGRP and ZI populations. **B,E.** Distribution of mean MAFs for 1,000 sets of LD-independent, non-antagonistic SNPs that have been frequency-matched to LH_M antagonistic SNPs (see Methods). Blue line denotes mean MAF of antagonistic SNPs; black dashed line denotes mean MAF of non-antagonistic SNPs before frequency-matching. **C,F.** MAF in the DGRP and ZI populations across 100 sets LD-independent SNPs, each set matched for LH_M allele frequencies, and presented in ascending order by P-value. For visualisation purposes, a linear regression line (±95% CI) is shown. **G.** Mean (±S.E.) residual Tajima's D (corrected for linked selection, see Methods) for antagonistic windows (blue; 'antagonistic status=1') and non-antagonistic windows (grey; 'antagonistic status=0') in the DGRP and ZI populations. **H.** Residual F_{ST} (±S.E.), corrected for linked selection (see Methods), for antagonistic and non-antagonistic windows in the DGRP and ZI populations. **I.** Linkage disequilibrium (r^2) in the ZI population between pairs of antagonistic SNPs (blue, 'Ant./ant.'), pairs of non-antagonistic SNPs (grey, 'Control/control') and mixed pairs (black, 'Ant./control'). Points represent mean r^2 across 25bp bins; r^2 is modelled as a declining exponential function of distance (fitted lines).