Sexual antagonism exerts evolutionarily persistent genomic constraints on sexual differentiation in *Drosophila melanogaster*

Mark S. Hill[1#], Filip Ruzicka[1#], Sara Fuentes[1], Julie M. Collet[2], Edward H. Morrow[3], Kevin Fowler[1], & Max Reuter[1*]

[#] Equal contribution
* Corresponding author

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

[2] Unité Mixte de Recherche, Centre de Biologie pour la Gestion des Populations, Institut National de la Recherche Agronomique, Avenue du Campus Agropolis, 34988 Montferrier-sur-Lez, France

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

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]. Sexual antagonism is taxonomically widespread and imposes an important evolutionary constraint on adaptation[1,4]. It also maintains genetic variation with sex-specific deleterious effects[6,7], contributing to disease susceptibility in humans[8,9]. Yet, despite its prevalence and importance, we know virtually nothing about the evolutionary dynamics of sexually antagonistic variation[10–12] or the general properties of the biological processes underlying antagonistic loci[13]. Here we report the first genome-wide identification and characterisation of sexually antagonistic SNPs in a population of *D. melanogaster*. We identify 4,899 antagonistic SNPs, distributed in 1,462 local clusters across the genome. Contrary to longstanding predictions[3], we show that these loci are significantly underrepresented on the X chromosome. Comparative analyses reveal that sexual antagonism exerts widespread and evolutionarily persistent constraints on the species' genome. Thus, antagonistic loci generate a detectable signature of balancing selection in populations across the distribution range of *D. melanogaster*, and often segregate as extended haplotypes with homogeneous sex-specific fitness effects. In functional terms, we find that sexual antagonism is rooted in the regulation of development and associated primarily with cis-regulatory elements. We further detect multiple associations with the sexual differentiation cascade, including the gene fruitless, a major driver of neuronal and behavioural differentiation between males and females. This reveals that sexual antagonism resides at the core of sex-specific development. Collectively, these results provide unprecedented insights into the biology and evolution of sexual antagonism, and open opportunities for more detailed studies of constraints on sex-specific adaptation and mechanisms that resolve them.

A wealth of quantitative genetic studies has established sexual antagonism as near ubiquitous across a wide range of taxa, including mammals[14] birds[15], reptiles[16], insects[17,18], fish[13,19] and plants[20]. Accordingly, sexual antagonism can be considered a major constraint on adaption and an important mechanism for the maintenance of fitness variation within populations[6]. 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[13,19], these are of limited use for elucidating general properties of loci experiencing sexual antagonism. On a genome-wide scale, previous transcriptomic work has associated antagonistic fitness effects with patterns of gene expression[21]. But despite revealing some of the molecular correlates of antagonism, this approach cannot distinguish between causal antagonistic loci and their downstream regulatory targets. We need to identify and characterise causal antagonistic loci in order to understand the adaptive limits to sexual dimorphism and how mechanisms of conflict resolution might arise.

To address this shortcoming, we combined experimentation, population genomics and bioinformatics to identify causal antagonistic loci across the *D. melanogaster* genome. Building on previous studies, we assayed the LH_M population in which sexually antagonistic fitness effects were first characterised[17,22]. To identify causal sexually antagonistic SNPs, we compared the genomic sequences of hemiclonal fly lines[21] that exhibited extreme male-beneficial/female-detrimental (N_{MB}=5) and female-beneficial/male-detrimental (N_{FB}=4) fitness effects. Out of >1x10^6 informative SNPs, 6,275 variants showed a pattern of perfect

segregation where the two fitness classes of genomes were fixed for opposing alleles. These SNPs constitute primary candidates for the underlying causal sexually antagonistic loci.

Initial analyses support the credibility of our candidate sites. Compared to all informative SNPs, candidate SNPs were enriched in genic regions (including upstream, downstream, and 3'UTR regions) and depauperate in intergenic regions (Fig. 1a), the pattern expected if variants are functional and under selection. Second, the number of candidate SNPs detected in our screen far exceeded the random expectation. Simulated sampling of alleles based on genome-wide estimates of allele frequencies in the LH_M population[23] returned significantly fewer incidences of perfectly segregating sites than the observed data (Fig. 1b). Additionally, the number of perfectly segregating sites was much larger in the original dataset than when permuting genomes between fitness classes (Fig. 1c), indicating that our approach detected a clear genetic signal of the phenotypic fitness effects. The relatively consistent random expectation of ~1000 perfectly segregating sites obtained in the two Monte Carlo approaches also provided an estimate of the number of false positives among our candidates of ~20%.

The 6,275 primary candidate SNPs were significantly clustered (Fig. 1d, Supp. fig. 1), with a median distance of 197bp and 6980bp between adjacent candidates on the autosomes and the X, respectively. We exploited the clustering of antagonistic SNPs to reduce the contribution of false positives by retaining candidates that were part of local clusters and removing those that were isolated and not in linkage disequilibrium with other candidate SNPs (Supp. fig. 2). As expected, isolated SNPs show properties that more closely resemble non-candidates than clustered candidates (Supp. fig. 3). After removal, 4,899 SNPs across 1,462 local clusters remained for downstream analysis. For simplicity, we hereafter refer to these as antagonistic SNPs.

We examined the genomic distribution of antagonistic SNPs and found that they were significantly underrepresented on the X chromosome (Fig. 1e). This result contradicts classical theory which predicts that the X chromosome will be a hotspot for the accumulation of antagonistic variation[3]. The low prevalence of antagonistic polymorphisms on the X is more in line with recent models[24,25] which suggest that the emergence of autosomal variation is facilitated by sex-specific dominance of antagonistic variants. Dominance of the beneficial allele in each sex increases the average fitness of heterozygotes and generates strong balancing selection, as documented recently for an individual sexually antagonistic polymorphism in salmon[13].

We next 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[7,26,27]. Our power to discern such a signal in the source population LH_M is low, because considering SNPs that were polymorphic across nine antagonistic hemiclonal genomes imposes a strong ascertainment bias that elevates heterozygosity across all SNPs used and thereby masks differences between antagonistic and non-antagonistic SNPs (Supp. fig. 4). Despite this, we found a signal of balancing selection in LH_M. Regional polymorphism (measured as Tajima's D) was significantly higher in windows containing antagonistic SNPs than in windows containing only non-antagonistic SNPs (Supp. fig. 4). This shows that the increased phenotypic variation in sex-specific fitness generated 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 signature of balancing selection at antagonistic loci in the LH_M population allow us to address this question by looking for a matching signal in other populations. To do so, we analysed publicly available population genomic data from the *Drosophila* Genetics Reference Panel[28,29] (DGRP) and phase 3 of the *Drosophila* Population Genomics project[30] (DPGP3). Each dataset provides information on whole-genome polymorphism from ~200 isogenic lines, with the DGRP constituting a sample from a recently introduced North American population (similar to LH_M), while the DPGP3 was sampled in Zambia in *D. melanogaster*'s ancestral distribution range. Just as in LH_M, we found that antagonism generated a clear signature of balancing selection in these two independent population samples. Antagonistic SNPs had significantly elevated heterozygosity in both the DGRP and DPGP3 populations compared to non-antagonistic SNPs (Fig. 2a). At a larger scale, antagonistic windows had significantly higher polymorphism (Tajima's D) than non-antagonistic windows (Fig. 2b), while also exhibiting lower population differentiation between the DGRP and DPGP3 (measured as Φ_{ST}; Fig. 2c). Importantly, these results held after controlling for genome-wide variation in recombination rate and proximity to functional regions ('linked selection'[31]). Taken together, these comparative analyses demonstrate that the antagonistic allelic variation identified in LH_M is neither recent nor specific to this population. To a significant degree, it has been conserved since before the extension of the species range beyond Africa ca. 10,000 years ago[32–35].

The population genomic whole-genome sequence data further revealed non-random associations between antagonistic SNP variants. Looking at clusters of antagonistic SNPs in the DPGP3, the population most phylogenetically distant from LH_M, we found that alleles frequently segregated as haplotypic blocks of variants with concordant fitness effects (see example in Fig. 2d). Furthermore, clusters where fitness effects were most homogeneous, and male-beneficial and female-beneficial haplotypes most equal in frequency, showed the most stable frequencies across populations (lowest Φ_{ST}; Fig. 2e, Supp. fig. 5), indicating that balancing selection was strongest on these clusters. These observations are consistent with theory predicting that linkage disequilibrium between alleles with similar fitness effects helps to maintain antagonistic polymorphisms[36].

Another fundamental knowledge gap concerns the biological processes that underlie sexual antagonism. Our data allow us to characterize these processes for the first time. At the most basic level, our results suggest that antagonism arises mainly due to adaptive conflict over gene expression, rather than over coding sequences. Antagonistic SNP variants did not cause missense changes more often than expected but were enriched in genic regions that contributed to expression regulation (up- and downstream, UTRs, Fig. 1a). In line with antagonistic selection on gene expression, we also found that the 1,204 antagonistic genes defined by our SNP data (those containing one or more antagonistic SNPs within ±5kb of the gene coordinates) significantly overlapped with genes that were previously shown to have sexually antagonistic expression patterns (opposing relationships between expression level and fitness in males and females[21]) (Fig. 3a). In terms of specific regulatory functions, we did not detect a global enrichment of transcription factors among antagonistic genes (*trans* regulation). We did find, however, evidence for associations between antagonism and specific *cis*-regulatory functions. Thus, antagonistic genes were enriched for targets of specific regulators, with binding site motifs of 38 transcription factors being over-represented upstream of antagonistic genes (Fig. 3b, Supp. tab. 1). Furthermore, short sequence stretches immediately flanking antagonistic SNPs were centrally enriched for binding motifs for 6

transcription factors (Fig. 3c). These results suggest that antagonism tends to be clustered in specific branches of gene regulatory cascades.

The general biological functions performed by antagonistic genes tended to relate to development and morphogenesis (Gene Ontology analysis; Supp. tabs. 2-4). Most notably, we found a significant association with the sex-determination and differentiation pathway. Antagonistic genes were overrepresented among regulators of sexual differentiation (observed=14, expected=7, χ^2=6.2, P=0.01) and included developmental genes that interact with the sex-determining cascade, such as *Abdominal A* and *Abdominal B*. Of particular note is *fruitless*, a direct target of the *Drosophila* sex-determination cascade and an important component of sex-specific neuronal development, male courtship behaviour, and physiology[37–39]. Despite its fundamental and highly evolutionarily conserved role in sexual differentiation[40], *fruitless* contained several clusters of antagonistic SNPs (Supp. fig. 6). Significant associations between antagonism and sexual differentiation also appeared at the *cis*-regulatory level. Thus, antagonistic genes were enriched for targets of several sexual differentiation genes (Fig. 3b, Supp. tab. 1), including daughterless, *Sex combs reduced*, *longitudinals lacking*, and again, *Abdominal B*. In addition, the motif most significantly associated with the flanking regions of antagonistic SNPs was a putative binding site for *longitudinals lacking* (Fig. 3c).

We experimentally verified the fitness effects of antagonistic loci by performing a replicated, short-term, evolve-and-resequence experiment under sex-limited selection[41]. Male-limited selection allows male-beneficial (and normally female-detrimental) alleles to rapidly increase in frequency, while female-limited selection in contrast allows female-beneficial alleles to increase. To track frequency changes, we performed pooled sequencing on populations at the start and the end of the experiment after 3 generations. This allowed us to estimate selection coefficients at autosomal antagonistic SNPs identified above (see Supplementary Information for details). Despite the limited power of the experiment, the distribution of selection coefficients strongly differed under the two regimes (Fig. 4). Selection coefficients of female-beneficial variants under female-limited selection were positive and significantly different from zero, although we failed to recover the mirrored effect under male-limited selection, where more stochastic variance in reproductive success may mask phenotypic effects.

Our characterisation of causal loci has provided unprecedented insights into the evolutionary dynamics and functional basis of sexual antagonism. Our data show that variation at antagonistic loci is stably maintained across *D. melanogaster* populations throughout the species' distribution range, aided by the long-term persistence of extended antagonistic haplotypes. The low turnover in antagonistic sequence variation implies that the phenotypic and genetic targets of sexually antagonistic selection have remained remarkably stable over time and space. It also suggests that the evolutionary constraints on sexual dimorphism inherent in antagonism are difficult to resolve, possibly requiring complex genetic changes that are unlikely to arise in single mutational steps. The resolution of antagonism, and the evolution of dimorphism, is then likely to proceed in a punctuated, stepwise manner[23]. Recently, we have described a possible example of such a resolution event in a LH_M stock population[23], where we observed a significant reduction in antagonism at the phenotypic level and rapid, chromosomally localised frequency change at the sequence level. Interestingly, the loci affected there show significant overlap with the antagonistic loci identified here (genes: observed=122, expected=91, χ^2=12.1, P=0.0005; individual SNPs: observed=24, expected=16, χ^2=4.0, P=0.046), corroborating the proposed model of long-term stasis and punctuated resolution.

The prevalence of antagonism among regulators of sexual differentiation suggests that the evolution of dimorphism is impeded by the rate with which genes can acquire sex-specific regulation. This is supported by the fact that antagonistic genes show a lower incidence of sex-biased expression than the genome at large (obs=520, exp=594, χ^2=29.1, P<0.001). With this constraint in place, antagonism permeates across developmental regulatory cascades and up to their very top, as in the case of *fruitless*. Superficially, the association of antagonism with such a fundamental and tightly selected biological process might be surprising. However, the sexual differentiation cascade provides an effective regulatory lever whereby mutations can simultaneously affect many aspects of male and female phenotypes. This will allow them to generate the large and balanced fitness effects that characterise the antagonistic polymorphisms most likely to be maintained over prolonged periods of time[11,26]. Having filled a longstanding and major gap in our understanding of sexual antagonism now provides a platform from which to further elucidate the origin and resolution of this important evolutionary phenomenon.

References

1. Bonduriansky, R. & Chenoweth, S. F. Intralocus sexual conflict. *Trends Ecol. Evol.* **24,** 280–8 (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 (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. Kidwell, J. F., Cleggl, 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,** (1977).
7. Gavrilets, S. & Rice, W. R. Genetic models of homosexuality: generating testable predictions. *Proc. R. Soc. B Biol. Sci.* **273,** 3031–8 (2006).
8. Morrow, E. H. & Connallon, T. Implications of sex-specific selection for the genetic basis of disease. *Evol. Appl.* **6,** 1208–17 (2013).
9. Casarini, L. & Brigante, G. The polycystic ovary syndrome evolutionary paradox: a genome-wide association studies–based, in silico, evolutionary explanation. *J. Clin. Endocrinol. Metab.* **99,** E2412–E2420 (2014).
10. Stewart, A. D., Pischedda, A. & Rice, W. R. Resolving intralocus sexual conflict: genetic mechanisms and time frame. *J. Hered.* **101,** S94–S99 (2010).
11. Connallon, T. & Clark, A. G. A general population genetic framework for antagonistic selection that accounts for demography and recurrent mutation. *Genetics* **190,** 1477–89 (2012).
12. Dean, R., Perry, J. C., Pizzari, T., Mank, J. E. & Wigby, S. Experimental evolution of a novel sexually antagonistic allele. *PLoS Genet.* **8,** e1002917 (2012).
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. Mokkonen, M. *et al.* Negative Frequency-Dependent Selection of Sexually Antagonistic Alleles in *Myodes glareolus*. *Science* **334,** 972–974 (2011).
15. 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).

16. 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).
17. Rice, W. R. Sexually antagonistic genes: experimental evidence. *Science* **256,** 1436–1439 (1992).
18. Berger, D., Berg, E. C., Widegren, W., Arnqvist, G. & Maklakov, A. A. Multivariate intralocus sexual conflict in seed beetles. *Evolution* **68,** 3457–3469 (2014).
19. 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).
20. Delph, L. F. *et al*. Environment-dependent intralocus sexual conflict in a dioecious plant. *New Phytol*. **192,** 542–552 (2011).
21. Innocenti, P. & Morrow, E. H. The sexually antagonistic genes of *Drosophila melanogaster*. *PLoS Biol*. **8,** e1000335 (2010).
22. 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).
23. Collet, J. M. *et al*. Rapid evolution of the intersexual genetic correlation for fitness in *Drosophila melanogaster*. *Evolution* **70,** 781–795 (2016).
24. Fry, J. D. The genomic location of sexually antagonistic variation: some cautionary comments. *Evolution* **64,** 1510–6 (2009).
25. Jordan, C. Y. & Charlesworth, D. The potential for sexually antagonistic polymorphism in different genome regions. *Evolution* **66,** 505–516 (2012).
26. 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).
27. Connallon, T. & Clark, A. G. Balancing selection in species with separate sexes: insights from Fisher's geometric model. *Genetics* **197,** 991–1006 (2014).
28. Mackay, T. F. C. *et al*. The *Drosophila melanogaster* Genetic Reference Panel. *Nature* **482,** 173–8 (2012).
29. Huang, W. *et al*. Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. *Genome Res*. **24,** 1193–1208 (2014).
30. 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).
31. Elyashiv, E. *et al*. A genomic map of the effects of linked selection in *Drosophila*. *PLoS Genet*. **12,** e1006130 (2016).
32. Lachaise, D. *et al*. Historical biogeography of the *Drosophila melanogaster* species subgroup. *Evol. Biol*. **22,** 159–225 (1988).
33. Li, H. & Stephan, W. Inferring the demographic history and rate of adaptive substitution in *Drosophila*. *PLoS Genet*. **2,** e166 (2006).
34. 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).
35. Duchen, P., Zivkovic, 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).
36. Patten, M. M., Haig, D. & Úbeda, F. Fitness variation due to sexual antagonism and linkage disequilibrium. *Evolution* **64,** 3638–3642 (2010).
37. Kimura, K. I., Ote, M., Tazawa, T. & Yamamoto, D. Fruitless specifies sexually dimorphic neural circuitry in the *Drosophila* brain. *Nature* **438,** 229–233 (2005).

38. Villella, A. & Hall, J. C. Neurogenetics of Courtship and Mating in *Drosophila*. *Advances in Genetics* **62,** 67–184 (2008).
39. Neville, M. C. *et al*. Male-Specific Fruitless Isoforms Target Neurodevelopmental Genes to Specify a Sexually Dimorphic Nervous System. *Curr. Biol*. **24,** 229–241 (2014).
40. Gailey, D. A. *et al*. Functional conservation of the fruitless male sex-determination gene across 250 Myr of insect evolution. *Mol. Biol. Evol*. **23,** 633–643 (2006).
41. Morrow, E. H., Stewart, A. D. & Rice, W. R. Assessing the extent of genome-wide intralocus sexual conflict via experimentally enforced gender-limited selection. *J. Evol. Biol*. **21,** 1046–1054 (2008).

Acknowledgements

We thank the Centre for Genomic Research, University of Liverpool, and the Institute of Child Health, UCL, for assistance with high-throughput sequencing, Claire Fourcade for exploratory analyses of antagonistic gene regulatory effects and Daniel Wegmann for generous help with running ApproxWF. We are grateful to Laurent Keller, Andrew Pomiankowski and François Balloux for useful comments on the manuscript. MSH was funded by a UCL IMPACT PhD studentship, FR by a London NERC DTP PhD studentship, SF and JC by a NERC research grant (NE/G0189452/1) to MR and KF, and EHM by a European Research Council Grant (280632) and a Royal Society University Research Fellowship.

Figures

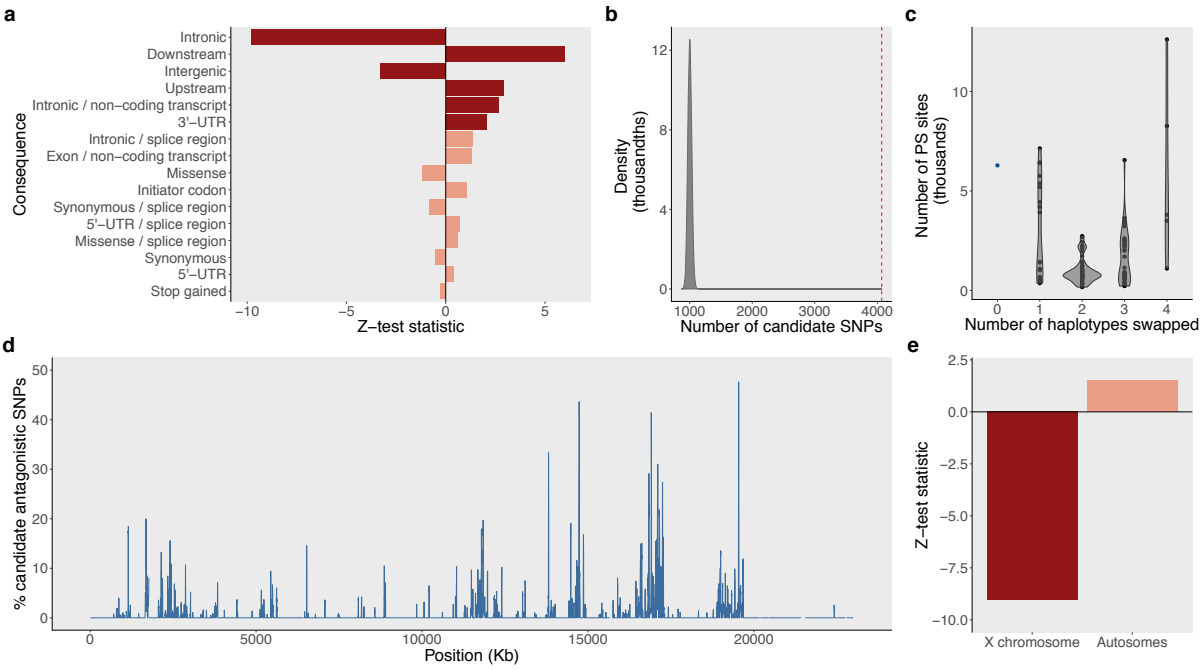


Figure 1 | Characteristics of candidate sexually antagonistic SNPs. a, Predicted variant effects of sexually antagonistic candidate SNPs relative to all SNPs covered in the haplotype sequencing. Dark red bars indicate statistically significant over-/under-representation, lighter red bars P>0.05. **b,** Null distribution of the expected number of candidate SNPs from simulated sampling of alleles according to their frequency in the LH_M population (grey surface). The median number of pseudo-candidates across permuted datasets was equal to 1,000. The dashed red line indicates the observed number of candidate antagonistic SNPs that are covered in the population genomic data used for simulated sampling (N=4,071). The overrepresentation of candidate SNPs in the observed data was statistically significant (P<0.001). **c,** Distributions of the number of perfectly segregating (PS) sites as a function of the number of haplotypes swapped between fitness classes (0 = original fitness classes). **d,** Sliding window plot (window size=10,000bp, step size=2,500bp) of chromosome arm 2L showing the percentage of informative SNPs in each window that are antagonistic candidates. **e,** Representation of antagonistic SNPs, relative to all informative SNPs, on autosomes and the X chromosome. Dark red bars indicate statistically significant over-/under- representation, lighter red bars P>0.05.

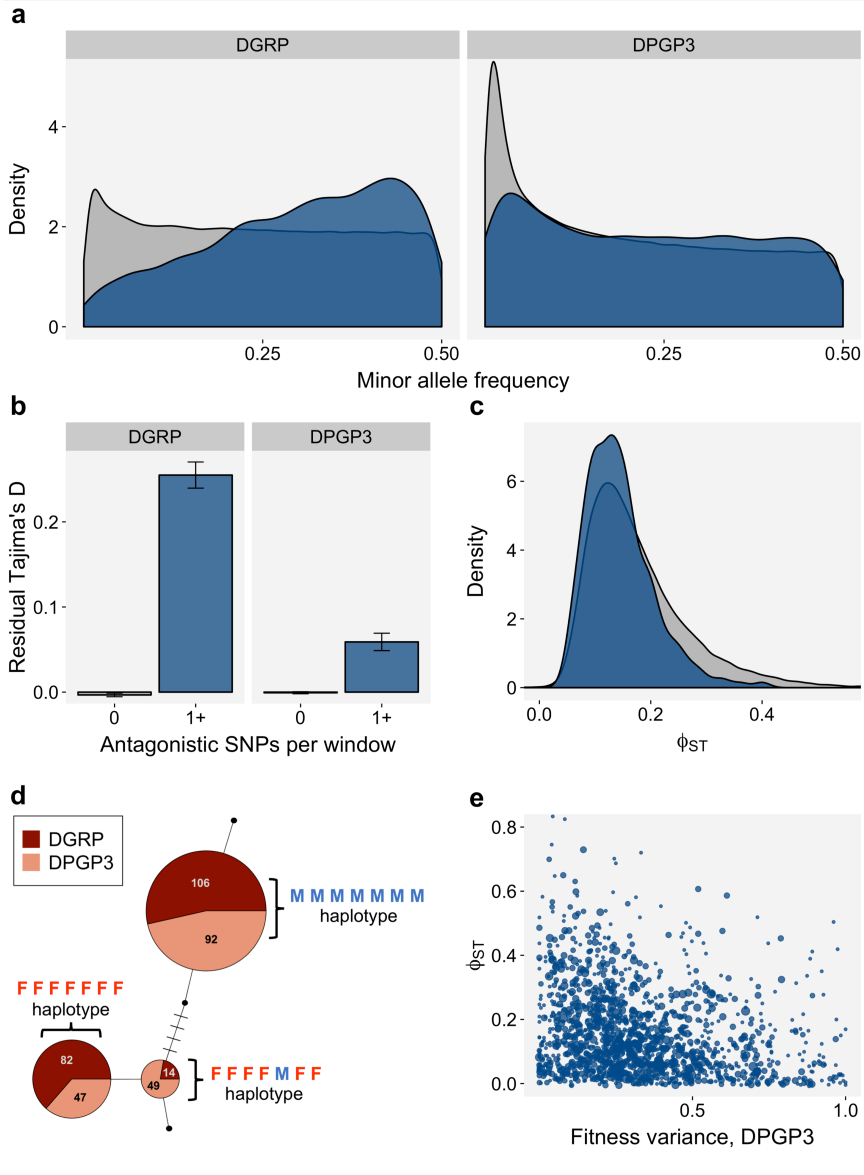
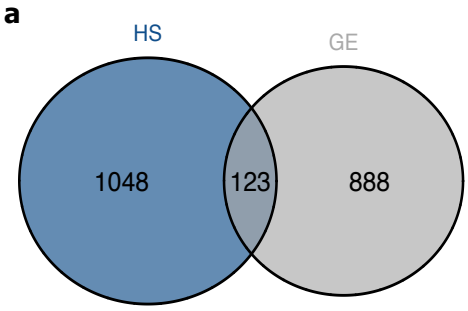
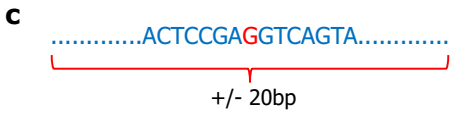


Figure 2 | Signatures of balancing selection generated by sexually antagonistic SNPs across *D. melanogaster* populations. a, Site frequency spectrum for highly covered (≥100) SNPs in the DGRP and DPGP3 that our haplotype sequencing data identify as antagonistic (blue, N_{DGRP}=4,605; N_{DPGP3}=4,158) or non-antagonistic (grey, N_{DGRP}=911,395; N_{DPGP3}=811,901). Antagonistic SNPs have elevated minor allele frequencies, resulting in elevated heterozygosity (Wilcoxon test, P<0.001) in both populations. **b,** Mean±s.e.m. of Tajima's D, corrected for the effects of linked selection, in the DGRP and DPGP3. Blue bars indicate sliding windows (window size=1000bp, step size=500bp) containing one or more antagonistic SNPs (N_{DGRP}=1,459; N_{DPGP3}=1,460), grey bars indicate windows containing only non-antagonistic SNPs (N_{DGRP}=111,733; N_{DPGP3}=113,091). Antagonistic windows display elevated Tajima's D in both populations (DGRP: $F_{1,113190}$=205.6, P<0.001; DPGP3: $F_{1,114548}$=65.83, P<0.001). **c,** Population differentiation (Φ_{ST}) between the DGRP and DPGP3, calculated in sliding windows as above. Blue and grey density surfaces correspond to windows with (N=1,459) or without (N=112,267) antagonistic SNPs. Differentiation is

significantly lower in windows containing antagonistic SNPs (Wilcoxon test, $P<0.001$). This remains true when accounting for linked selection (Wilcoxon test, $P<0.001$). **d,** Haplotype network for a cluster of seven antagonistic SNPs (M: male-beneficial allele; F: female-beneficial allele) situated in the gene *fruitless*. 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. The two major haplotypes are identical to those identified as perfectly segregating in the LH_M population; recombinant haplotypes are comparatively infrequent (~17%). **e,** Negative correlation between Φ_{ST} and fitness variance, both calculated across clusters of antagonistic SNPs (Spearman rank correlation, $\rho=-0.260$, $P<0.001$, $N=1,457$). Each point depicts a cluster of antagonistic SNPs, the relative sizes denote the number of SNPs per cluster. Fitness variance quantifies the extent to which clusters of SNPs have homogeneous sex-specific fitness effects and are maintained at even frequencies between male- and female-beneficial haplotypes (see Supplementary Information). Clusters where fitness variance is high tend to be at intermediate frequencies in both populations, thus lowering the proportion of variance explained by population (Φ_{ST}).



FBgn0000015	*Abdominal B*
FBgn0033749	*achintya*
FBgn0013753	*Big brother*
FBgn0000611	*extradenticle*
FBgn0001235	*homothorax*
FBgn0003300	*runt*
FBgn0033748	*vismay*
FBgn0029822	*CG12236*
FBgn0004050	*zeste*
FBgn0015381	*dissatisfaction*



FBgn0283521	*longitudinals lacking*
FBgn0015936	*elbow A*
FBgn0003053	*pebbled*
FBgn0037207	*mesoderm-expressed 2*
FBgn0029822	*CG12236*
FBgn0005638	*slow border cells*

Figure 3 | Functional properties of antagonistic loci. a, Overlap of antagonistic genes identified in our study by haplotype sequencing (HS, blue) and antagonistically expressed genes identified in previous study of antagonistic gene expression[21] (GE, grey). The number of genes shared between the two studies was significantly greater than expected by chance (observed overlap=123, expected overlap=94, χ^2=10.01, P=0.002). **b,** Top 10 genes associated with motifs enriched in upstream regions (5kb upstream of transcription start sites) of antagonistic genes. Upstream regions are denoted by the red bracket, antagonistic genes are shown in blue. **c**, Genes associated with motifs which were centrally enriched in short sequences flanking antagonistic SNPs (red nucleotide in the example sequence).

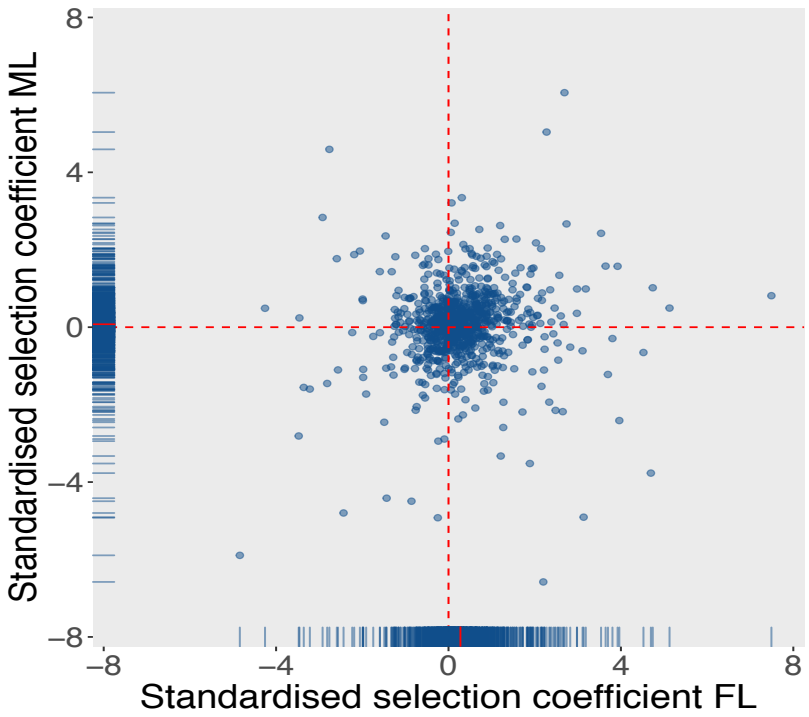


Figure 4 | Selection on autosomal antagonistic SNP clusters under sex-limited selection. Standardised selection coefficients of female-beneficial variants under female- (FL) and male-limited (ML) selection, summed across all antagonistic SNPs of a cluster that were covered in the evolve-and-resequence data. The red line in each axis ribbon denotes mean standardised selection coefficient under FL and ML selection. The distribution of FL and ML selection coefficients were significantly different ($F_{1,1845}$=27.0, P<0.001) with FL selection coefficients positive and significantly different from zero (t_{1845}=9.2, P<0.001). Selection coefficients under ML selection were not significantly different from zero (t_{1845}=1.8, P=0.07).

Supplementary information

Supplementary methods

Hemiclonal haplotypes

We sequenced the genomes of the set of most antagonistic genotypes identified in a previous study[1]. 5 male-beneficial/female-detrimental (MB) genotypes and 4 female-beneficial/male-detrimental (FB) genotypes (a fifth FB had been lost prior to this study) had been maintained as 'hemiclones', i.e., intact haploid sets of chromosomes X, II and III, by back-crossing males to 'clone-generator' females [C(1)DX, y, f; T(2;3) rdgC st in ri pP bwD]. In order to obtain sequences of the haploid hemiclonal genomes, we expressed them in three different genetic backgrounds that could then be used to infer the genome sequence of interest with confidence (see section 'Identification of informative SNPs'). The three genotypes sequenced for each hemiclonal genome were (i) females in which the hemiclonal genome was complemented with chromosomes from the *Drosophila melanogaster* reference strain iso-1 (y[1]; Gr22b[iso-1] Gr22d[iso-1] cn[1] CG33964[iso-1] bw[1] sp[1]; MstProx[iso-1] GstD5[iso-1] Rh6[1]), (ii) females in which the hemiclonal genome was complemented with chromosomes from the Canton-S strain, and (iii) males in which the hemiclonal genome was complemented with chromosomes from the clone-generator strain.

DNA extraction, sequencing and initial processing of haplotype sequencing reads

10-25 individuals were collected for each of the 27 samples in March 2012 (9 hemiclones, 3 genetic backgrounds per hemiclone). We extracted total genomic DNA using DNeasy® Blood and Tissue Kit (Qiagen) according to the manufacturer's protocol and purified DNA samples using Agencourt® AMPure XP® beads (Beckman Coulter). Paired-end Nextera libraries were prepared using a Nextera DNA sample preparation kit (Illumina). Completed libraries were pooled in equimolar amounts and fragments between 450 and 650bp were collected using the Pippin Prep DNA size selection system (Sage Scientific). Size-selected pools were purified using Agencourt® AMPure XP beads. Sequencing was performed on 4 lanes of an Illumina HiSeq 2000 at the Center for Genomic Research, University of Liverpool.

Basecalling and de-multiplexing of the indexed sequencing reads was performed using CASAVA version 1.8.2 (Illumina). The raw fastq files were first trimmed using Cutadapt version 1.2.1[2] to remove Illumina adapter sequences. The reads were then further trimmed to remove low quality bases (minimum window score of 20) and very short reads (<10bp) with Sickle version 1.200 (github.com/najoshi/sickle). Read pairs were then aligned to the *D. melanogaster* reference sequence (BDGP 5.5), obtained from the FlyBase online database (http://flybase.org/), using Bowtie2 version 2.1.0[3] in '--local' alignment mode. To avoid false positive SNP calls resulting from misalignment around indels, reads were locally realigned using the Genome Analysis Toolkit (GATK) version 2.1.13[4,5]. Duplicate reads, arising from PCR amplification during library construction were removed using Picard version 1.85 (http:// sourceforge.net/projects/picard/). The alignment results were visually inspected using the Integrative Genomics Viewer[6] SNP calling was performed using the GATK 'UnifiedGenotyper' package with heterozygosity set to 0.14 (to more closely reflect heterozygosity of *D. melanogaster*). Identified SNPs were then filtered to remove SNPs with low confidence and low coverage (<10 reads) using the GATK 'VariantFiltration' package.

Identification of informative and candidate SNPs

To identify informative SNPs from the hemiclonal genomes, we exploited the fact that their DNA was present in all three crosses and that in one of the crosses, they were complemented with chromosomes of the reference strain. Thus, for a given chromosomal position with sufficient coverage in the sequencing of all three crosses, we inferred the hemiclonal variant to be an alternative (non-reference) variant if an identical alternative variant was detected in all three crosses. The variant was assumed to be the reference variant if this condition was not fulfilled.

Once variants for all nine hemiclonal genomes were identified in this way, we identified a set of informative SNPs where variants were polymorphic across the nine hemiclonal genomes.

Candidates for sexually antagonistic SNPs were then defined as those SNPs where all MB hemiclonal genomes were fixed for one variant and all FB hemiclonal genomes were fixed for the other variant ('perfectly segregating').

Computational validation of candidate SNPs

We implemented two independent tests to determine whether the number of candidate SNPs detected with our approach was greater than expected by chance, (i) a resampling test based on population allele frequencies, and (ii) a permutation test that shuffled hemiclonal haplotypes among fitness classes (MB and FB). In the first approach, we merged our dataset with high quality genome-wide allele frequency data generated from the same LH_M population (see section 'LH_M, DGRP and DPGP3 population genomic data' below). We then used binomial sampling to generate allelic samples for each SNP according to the population allele frequencies and recalculated the number of perfectly segregating sites ('pseudo-candidate SNPs'). This procedure was repeated 1,000 times and the null distribution of the number of pseudo-candidate SNPs was compared to the observed number of true antagonistic SNPs (Fig. 1b). Significance was determined by calculating the proportion of instances where the number of pseudo-candidates was greater than or equal to the number of true antagonistic SNPs (none of the simulation runs returned a value equal to or higher than the observed number of antagonistic SNPs, so $P<0.001$). In the second approach, we produced all possible permutations of hemiclonal haplotypes among the two fitness classes and for each recalculated the numbers of perfectly segregating sites. We then presented these as a function of the number of genotypes that had been swapped between classes (Fig. 1c).

Clustering of candidate SNPs

To examine the clustering of candidate SNPs across chromosome arms we calculated the median distance between all pairs of adjacent candidate SNPs (ignoring interspersed non-candidate SNPs). We did this separately for the autosomes and X chromosome, to accommodate for the lower SNP density on the X chromosome. We then designed a permutation test to determine the significance of the observed clustering. Candidate/non-candidate status was permuted among all informative SNPs, distances recalculated as before between adjacent SNPs labelled as 'candidates' after permutation and the median distance recorded. This process was repeated 1,000 times in order to generate a null distribution of median distances. Significance was calculated as the proportion of median distances in the null distribution that were lower than or equal to the true median distance. None of the permutated medians were lower or equal to the observed value, so $P<0.001$.

We clustered candidate SNPs based on linkage disequilibrium information in order to (i) reduce the number of false positives in our dataset, and (ii) examine the haplotypic structure of antagonistic SNPs (see also the section on 'Patterns of segregation across antagonistic clusters'). The aim was to adjust the level of clustering in such a way that SNPs grouped in the same cluster showed appreciable levels of linkage disequilibrium (LD; defined here as $r^2>0.25$), while SNPs grouped in different clusters showed low levels of LD. To achieve this, we used Ward's hierarchical clustering method based on Euclidean distances between candidate SNPs. We varied the parameter k, which describes the number of clusters formed (a larger k thus results in a larger number of smaller clusters). For each run, we calculated LD (r^2) between the first and last SNP in each cluster and fitted a model that describes cluster-wide LD as a declining exponential function of chromosomal distance. We then selected the smallest value of k (i.e. the smallest number of clusters) where the predicted LD for the longest cluster was greater than $r^2=0.25$ (Supp. fig. 2). This procedure was performed for each autosomal chromosome arm in turn, and consequently accounts for chromosomal differences in the degree of LD. Due to the paucity of antagonistic sites on the X chromosome, fitting an exponential model was not feasible for this chromosome. Instead, we chose a cut-off value ($k = 30$) which led to the inclusion of three long-range haplotypes (see Supp. fig. 2).

We used the above-defined clusters to class candidate SNPs as either clustered (i.e., more than one candidate SNP per cluster) or isolated ('cluster' consisting of a single candidate SNP). We examined any significant differences in their population genetic properties by comparing their heterozygosities using Wilcoxon Rank-Sum tests (Supp. fig. 3).

Genomic distribution of antagonistic SNPs

To examine the relative representation of antagonistic SNPs on autosomes and the X chromosome, we compared the proportion of antagonistic SNPs to the proportion of all informative SNPs mapping to each chromosomal compartment. We did this for each chromosomal compartment in turn, using Z-tests. The under- or over-representation of antagonistic SNPs in each compartment therefore accounted for general differences in SNP density between chromosome arms and, in particular, lower diversity on the X chromosome.

LH_M, DGRP and DPGP3 population genomic data

We used population genomic data derived from pooled sequencing of 165 adult female individuals from the LH_M population[7] to estimate polymorphism in the source population. We also used publicly available population genomic data from wild *D. melanogaster* populations in North America and Zambia. The North American data, from the *Drosophila* Genetic Reference Panel (DGRP[8,9]), consists of 205 whole-genome sequences derived from inbred lines established from flies caught in North Carolina (USA). The Zambian data, from phase 3 of the *Drosophila* Population Genomics Project (DPGP3[10]), consists of 197 whole-genome sequences derived from haploid embryos sampled from flies caught in a wild Zambian population. The two populations represent a range of divergence times from the LH_M population. While DGRP flies, like those of LH_M, are descendants of recent (~150 years[11]) colonisation of the USA, the DPGP3 was sampled in the ancestral range of *D. melanogaster*, and is separated from the American populations by ~10,000-200,000 years[11–13]. Genome sequences for the DGRP and DPGP3 were downloaded as FASTA files from http://www.johnpool.net/genomes.html. These sequence files have undergone standardised alignment and have been quality filtered[10]. The files were then converted to vcf with *snp-sites*[14]. We further excluded sites that were covered in <100 genomes sequences and those that segregated for more than two variants or variants that did not match those detected among the antagonistic LH_M hemiclones.

Balancing selection estimates in LH_M, DGRP and DPGP3

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, S), was compared for windows containing (i) one or more antagonistic SNPs and (ii) only non-antagonistic SNPs (Fig. 2b). In the LH_M, Tajima's D was calculated using *popoolation* (v. 1.2.2[15]; pool size=330, minimum quality=20, minimum count=2, minimum coverage=100, maximum coverage=290). In the DGRP and DPGP3, Tajima's D was calculated using the *PopGenome* package[16]. We incorporated estimates of linked selection[17] (estimated in 1,000bp sliding windows) to account for genomic correlations between populations owing to factors unrelated to sexual antagonism, such as local recombination rate variation and proximity to functional sequences. We used general linear models (GLMs) with Gaussian error structure to assess the effect of window class (antagonistic vs. non-antagonistic) on Tajima's D, with estimates of linked selection included as covariates (Tajima's D estimates displayed normally distributed residuals, justifying the normality assumption). Estimates of linked selection were not available on the X chromosome, so we used estimates of recombination rate[21] as covariates instead.

For the DGRP and DPGP3, we implemented SNP-level analyses by comparing expected heterozygosity between antagonistic and non-antagonistic SNPs (Fig. 2a). Expected heterozygosity was calculated as twice the product of the frequencies of the two alleles at a given SNP, and heterozygosities were compared between the two classes of SNPs using a Wilcoxon Rank-Sum test.

We tested whether balancing selection at the level of windows was associated with reduced population differentiation. Measures such as F_{ST} are sensitive to heterozygosity[18] and are less suitable for a window-based sequence comparison because they ignore the genetic distance between segregating haplotypic sequences. We therefore used the AMOVA framework[19] to calculate Φ_{ST}, a measure of population differentiation that takes into account genetic distances (i.e., numbers of substitutions) between sequences when quantifying the proportion of genetic variation that occurs between vs within populations. Φ_{ST} was calculated between the DGRP and the DPGP3 over sliding windows as above (1000bp windows, 500bp step) using the *poppr*[20] package. We did not include LH_M in this analysis because polymorphism data for this population is derived from pooled sequencing data

where haplotype information is not available. Wilcoxon Rank-Sum tests were used to test for differences in Φ_{ST} between windows with one or more, and windows with only non-antagonistic SNPs (Fig. 2c). To consider effects due to linked selection we re-ran Wilcoxon Rank-Sum tests and compared the residual Φ_{ST} between window classes, once linked selection estimates had been accounted for.

Patterns of segregation across antagonistic clusters

We quantified the extent to which antagonistic haplotypes are selectively maintained by investigating whether clusters of antagonistic SNPs commonly segregate as male- and female-benefit haplotypes (as inferred from LH_M) in the DPGP3, the population that is most distant from LH_M and where such a signal would therefore be expected to be weakest. For a given cluster, we assigned each sequence in the DPGP3 a sex-specific fitness score, counting +1 for each SNP variant defined as male-beneficial in the LH_M, -1 for each variant defined as female-beneficial in the LH_M, and 0 for each variant where fitness consequence was unknown (allele neither male-beneficial, nor female-beneficial). Thus, a sequence carrying four male-beneficial variants across a given cluster of antagonistic SNPs would be assigned a score of +4, while a sequence carrying two female-beneficial and two male-beneficial variants would obtain a score of 0. We could then calculate the variance in fitness scores across sequences within a population. Because this variance is expected to increase with the length L of the cluster (i.e., the number of antagonistic SNPs per cluster), we standardized fitness variances by dividing by L^2, which is the maximum fitness variance attainable (i.e., the variance in a population where one half of the sequences are composed exclusively of male-beneficial variants, resulting in scores of +L, and the other half exclusively of female-beneficial variants, resulting in scores of -L). A standardised fitness variance close to 1 then indicated that antagonistic haplotypes segregate as linked groups of relatively homogeneous fitness effects and close to even frequencies of male- and female-beneficial haplotypes. Conversely, a standardised fitness variance close to 0 indicated that haplotypes are composed of a mixture of male- and female-beneficial variants and/or segregate with uneven frequencies of more male- and more female-beneficial haplotypes.

To determine if antagonistic haplotypes explain variation in polymorphism and differentiation inferred from sliding windows analyses, we correlated fitness variance with signatures of balancing selection. We correlated window-wide estimates of Tajima's D and Φ_{ST} (see section 'Balancing selection estimates in LH_M, DGRP and DPGP3') with cluster-wide fitness variance using Gaussian distributed GLMs (Tajima's D) and Spearman rank-sum correlations (Φ_{ST}) respectively (Supp. fig. 4). We also re-calculated Φ_{ST} over clusters (calculated using *poppr*, but run over antagonistic SNPs rather than windows), and correlated it with cluster-wide fitness variance using Spearman rank-sum correlations (Fig. 2e).

Functional analyses of antagonistic loci

We used the variant effect predictor (Ensembl VEP[21]) to map SNPs to genes and infer their genetic consequences. In accordance with the VEP default settings, we included extended gene regions (+/- 5kb of gene coordinates) in our gene definition. To gain preliminary insights into the functions of antagonistic genes we used the Gorilla[22] gene ontology tool, and applied false discovery rate (FDR) to correct for multiple testing across many GO terms. Here all genes covered in our entire 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[23] to annotate genes covered in our sequencing as having either sex-biased or unbiased expression profiles, based on the meta-class identifier. We then used a χ^2 test to compare the sex-biased expression status of antagonistic and non-antagonistic genes.

To assess the degree of overlap between antagonistic genes identified here and those associated with sexually antagonistic expression patterns in a previous study[1], 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[24]). 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.

To compare the number of antagonistic and non-antagonistic transcription factor genes, we annotated all genes in our data that encode for transcription factors using the supervised regulatory network data from Marbach *et al.*[25], which defines 617 putative transcription factors.

To assess antagonistic genes for signatures of shared regulation we used the Motif Enrichment Tool[26] with default settings, to search regions 5kb upstream of the transcription start sites for enriched motifs. Here again, all genes covered in our SNP dataset were used for the background set. We applied FDR corrections to P-values to account for the testing of multiple motifs.

To test for motif enrichment close to antagonistic SNPs (+/- 20bp), we used the Centrimo tool from MEMEsuite[27] with default settings. We generated the input FASTA files using bedtools[28] and extracted the sequences (+/-20bp) flanking the SNPs in our dataset from the *D. melanogaster* reference genome (BDGP 5.5). We used flanking sequences for all informative SNPs to generate the background model and the flanking sequences of antagonistic SNPs as the target set. Finally, genes covered in our experiment were annotated according to the 'Sex differentiation' GO category (GO:0007548) from FlyBase. We used this information to test specifically for overrepresentation of sexual differentiation genes among candidate genes using a χ^2 test.

Experimental evolution

To evolve populations under sex-limited selection, we used a modified middle-class-neighborhood design[29]. This design removes fitness variance in a population by constraining reproductive output of all individuals to the same number of offspring, thereby effectively eliminating selection. By applying this design asymmetrically to one sex, it is possible to create populations where only males or only females experience selection (e.g.[30]). As counter-selection from the unselected sex is removed in this design, sexually antagonistic alleles are free to accumulate and/or rapidly increase in frequency. We applied sex-limited selection to both males and females in this way, establishing two replicates for each selection treatment. We allowed all four experimental populations to evolve for 3 generations. The number of replicate populations chosen was a compromise between statistical power and logistical constraints.

Each population of the male-limited selection (ML) treatment was composed of 120 mating groups of 4 males and 1 female (600 flies in total), sampled as virgins from the LH_M stock population. After interacting for two days in 'adult competition' vials, each group of flies was transferred to new 'oviposition vials' for 18h before being discarded. After 9 days of development, 5 virgin adults (4:1) were collected from each vial, ensuring that each female in the population had produced an identical number of offspring (4 males, 1 female) while male fitness varied according to mating success. After aging for 3 days the virgin flies of the next generation were pooled and re-distributed among a new set of 120 adult competition vials, maintaining the 4:1 ratio. This procedure was repeated for two more generations. Evolution under female-limited selection was performed in an equivalent way, using mating groups of 4 females and 1 male. Both the male- and female-limited selection (FL) regimes were designed to closely follow the standard rearing regime of the LH_M base population[31]. The number of flies per vial was chosen to maximise the selective pressure (with a more skewed sex ratio allowing for stronger selection) whilst still being feasible in terms of virgin collections and mortality of experimental flies (mainly owing to male harassment of females flies in the ML treatment) throughout the experiment.

DNA extraction and sequencing of the experimental populations

We collected flies from each experimental population at the start (all 600 founder parents in the first generation) and the end of the experiment (2000+ offspring collected after the last round of selection), yielding a total of 8 pooled samples. We extracted total genomic DNA from each of these samples using a CTAB/DTAB extraction method, and further purified the DNA samples using Agencourt® AMPure XP beads (Beckman Coulter). Paired-end Nextera libraries were prepared using a Nextera DNA sample preparation kit (Illumina) and sequenced on an Illumina HiSeq 2000 at the UCL Institute of Child Health.

Basecalling and de-multiplexing of the indexed sequencing reads was performed using CASAVA version 1.8.2 (Illumina), producing fastq files from the sequenced libraries. The raw fastq

files were first trimmed using Cutadapt version 1.2.1[2] to remove Illumina adapter sequences. The reads were then further trimmed to remove low quality bases (minimum window score of 20) and very short reads (<10bp) with Sickle version 1.200 (github.com/najoshi/sickle). Read pairs were then aligned to the *Drosophila* melanogaster reference sequence (BDGP 5.5), obtained from the FlyBase online database (http://flybase.org/), using Bowtie2 version 2.1.0[3].To avoid false positive SNP calls resulting from misalignment around indels, reads were locally realigned using the Genome Analysis Toolkit (GATK) version 2.1.13[4,5]. Duplicate reads, arising from PCR amplification during library construction were removed using Picard version 1.85 (http:// sourceforge.net/projects/picard/). The alignments were visually inspected using the Integrative Genomics Viewer[6].

We called SNPs separately for each population using SNVer[32]. SNver tests for the significant presence of an alternative allele rather than polymorphism *per se*. In order to correct for this, we re-ran the tool against a modified reference genome where the reference allele had been swapped for the alternative allele detected in the initial run. We considered a site to be a SNP if both tests were significant. All P-values were corrected for false discovery rate using the q-value package (http://www.bioconductor.org). We then used the filtered SNVer output (minimum read depth = 100, maximum read depth = 2x mean coverage) to ascertain allele frequencies at SNPs identified across the 8 samples.

Inferring selection coefficients from the evolve-and-resequence data

To estimate selection coefficients for antagonistic SNPs we used ApproxWF[33], which approximates the Wright-Fisher process using a Markov model. We analysed allele frequency data for all autosomal antagonistic SNPs covered in the sequencing of the experimental evolution experiment (N=2,721). We excluded X-linked SNPs from the analysis because the lack of father-to-son transmission of the X makes male-limited selection less efficient on this chromosome and the lower census size of X-chromosomes (which are hemizygous in males) generally reduces the efficacy of selection on X-linked variants. We arbitrarily polarised the data for female-beneficial alleles, meaning that a positive selection coefficient indicates selection for the female-beneficial allele/against the male-beneficial allele and a negative selection coefficient indicates selection for the male-beneficial allele/against the female-beneficial allele. The inference of selection requires information about the intensity of genetic drift, and hence the effective population size N_e. We parameterised the ApproxWF runs with the theoretical expectation for N_e based on our experimental design. We calculated N_e from the numbers of males and females and the (co)variances in male and female mating success using Equation 2 of Nomura[34]:

$$\frac{1}{N_e} = \frac{1}{16N_m}\left(2 + \sigma_{mm}^2 + 2\frac{N_m}{N_f}\sigma_{mm,mf} + \left(\frac{N_m}{N_f}\right)^2 \sigma_{mf}^2\right) + \frac{1}{16N_f}\left(2 + \sigma_{ff}^2 + 2\frac{N_f}{N_m}\sigma_{fm,ff} + \left(\frac{N_f}{N_m}\right)^2 \sigma_{fm}^2\right),$$

where N_m and N_f are the number of males and females, σ_{ij}^2 is the variance in the number of offspring of sex j produced by a parent of sex i and $\sigma_{ij,ik}$is the covariance between the number of offspring of sexes j and k produced by a parent of sex i. All variances are equal to zero for the sex that is not under selection, as all individuals of that sex produce the same number of male and female offspring. Reproductive variances for the selected sex are non-zero and N_e increases with decreasing variances, i.e., with a more even sharing of reproduction between competitors. In order to be conservative, we calculated N_e for the worst-case scenario where reproductive variance is maximal and only a single individual of the selected sex reproduces in each mating group. In this case, $\sigma_{ij}^2 = \frac{1}{4}(4-1)^2 + \frac{3}{4}(0-1)^2$ for i=j (e.g., paternal variance in the production of sons in the male-limited treatment) and $\sigma_{ij}^2 = \frac{1}{4}\left(1-\frac{1}{4}\right)^2 + \frac{3}{4}\left(0-\frac{1}{4}\right)^2$ for i≠j and $\sigma_{ij,ik} = \frac{1}{4}(4-1)\left(1-\frac{1}{4}\right) + \frac{3}{4}(0-1)\left(0-\frac{1}{4}\right)$ for the covariance between offspring of sexes j and k in the selected sex i=j. The theoretical effective population size in this case is N_e=291.

To compare selection coefficients for female-beneficial alleles under male- and female-limited selection we first summed the selection coefficients estimated at individual antagonistic SNPs over all SNPs of a cluster and fitted a GLM with Gaussian error structure to model variation in selection coefficients explained by selection regime (Fig. 4). The distribution of summed selection coefficients across clusters (N=923) was very close to normal, although slightly heavy-tailed. However, this did not affect our conclusions as we also tested for differences between the two treatments using a non-parametric Wilcoxon Rank-Sum test, with qualitatively identical results.

Statistical software

All statistical analyses were carried out in R (version 3.2.3, R Core Development Team 2016).

Data availability

Pooled sequencing data from the LH_M population is available at http://www.ebi.ac.uk/ena/data/view/PRJEB12822 under accession SAMEA3881048. Population genomic data from the DGRP is available at http://dgrp2.gnets.ncsu.edu/data.html. Population genomic data from the DPGP3 is available at http://www.johnpool.net/genomes.html. Sequence data from this study will be made available at the European Nucleotide archive.

Code availability

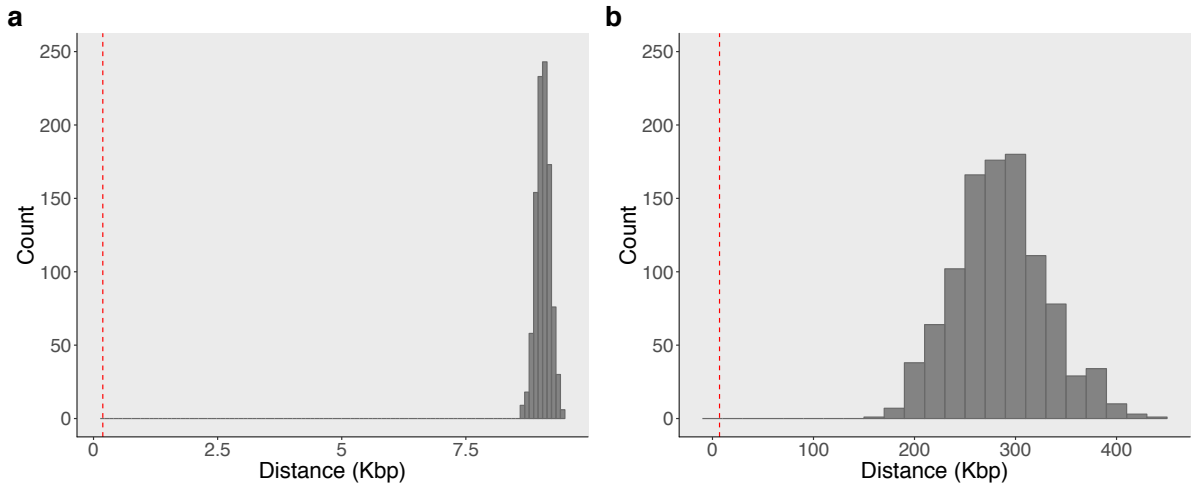
Analysis code is available on request.

References

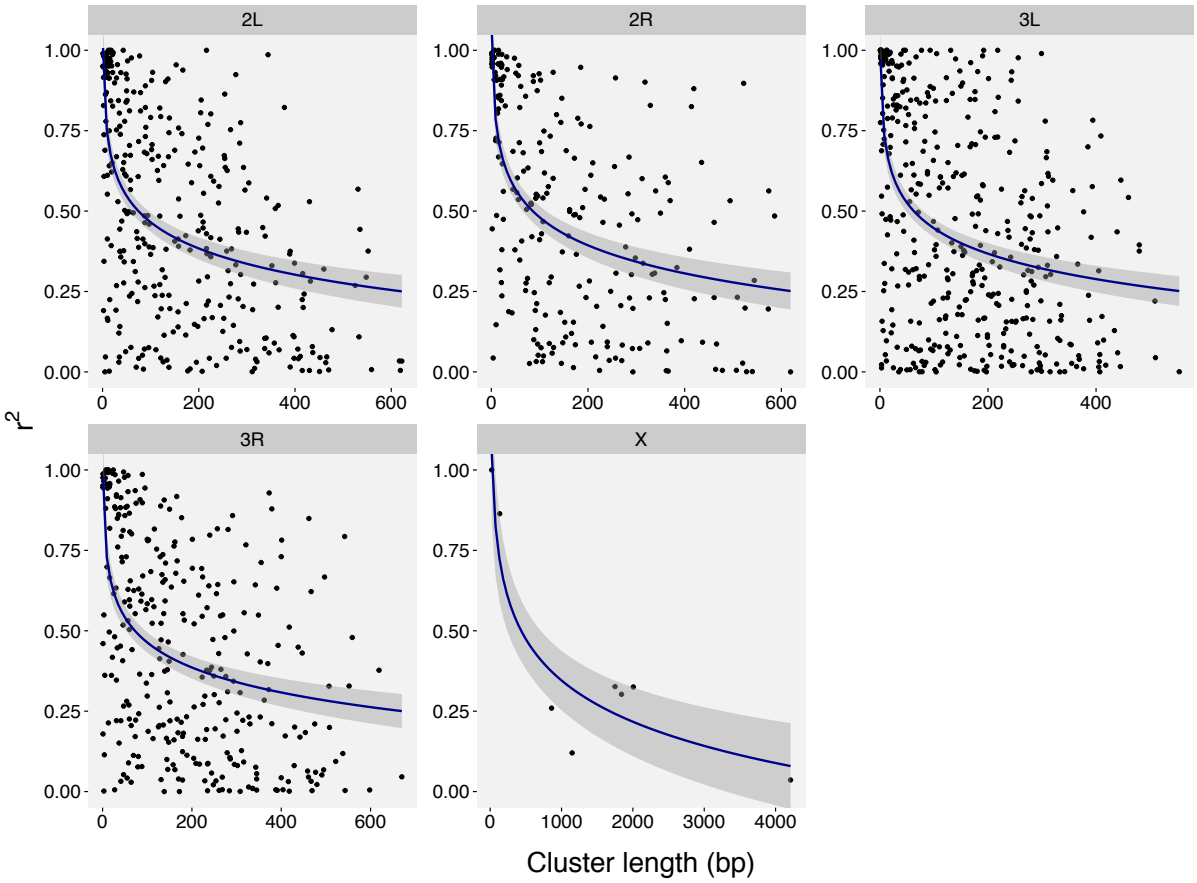
1. Innocenti, P. & Morrow, E. H. The sexually antagonistic genes of *Drosophila melanogaster*. *PLoS Biol*. **8,** e1000335 (2010).
2. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17,** 10 (2011).
3. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–359 (2012).
4. McKenna, A. *et al*. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. **20,** 1297–303 (2010).
5. DePristo, M. A. *et al*. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet*. **43,** 491–498 (2011).
6. Robinson, J. T. *et al*. Integrative genomics viewer. *Nat. Biotechnol*. **29,** 24–26 (2011).
7. Collet, J. M. *et al*. Rapid evolution of the intersexual genetic correlation for fitness in *Drosophila melanogaster*. *Evolution* **70,** 781–795 (2016).
8. Mackay, T. F. C. *et al*. The *Drosophila melanogaster* Genetic Reference Panel. *Nature* **482,** 173–8 (2012).
9. Huang, W. *et al*. Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. *Genome Res*. **24,** 1193–1208 (2014).
10. 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).
11. Duchen, P., Zivkovic, 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).
12. Lachaise, D. *et al*. Historical biogeography of the *Drosophila melanogaster* species subgroup. *Evol. Biol*. **22,** 159–225 (1988).
13. Pool, J. E. *et al*. Population Genomics of Sub-Saharan *Drosophila melanogaster*: African Diversity and Non-African Admixture. *PLoS Genet*. **8,** e1003080 (2012).
14. Keane, J. A. *et al*. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb. Genomics* **2,** (2016).
15. Kofler, R. *et al*. PoPoolation: A Toolbox for Population Genetic Analysis of Next Generation Sequencing Data from Pooled Individuals. *PLoS One* **6,** e15925 (2011).
16. 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).
17. Elyashiv, E. *et al.* A genomic map of the effects of linked selection in *Drosophila*. *PLoS Genet.* **12,** e1006130 (2016).
18. 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).
19. Excoffier, L., Smouse, P. E. & Quattro, J. M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131,** 479–91 (1992).
20. Kamvar, Z. N., Tabima, J. F. & Grünwald, N. J. *Poppr* : an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2,** e281 (2014).
21. McLaren, W. *et al.* Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* **26,** 2069–2070 (2010).
22. 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).
23. Gnad, F. & Parsch, J. Sebida: a database for the functional and evolutionary analysis of genes with sex-biased expression. *Bioinformatics* **22,** 2577–2579 (2006).
24. 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).
25. Marbach, D. *et al.* Predictive regulatory models in *Drosophila melanogaster* by integrative inference of transcriptional networks. *Genome Res.* **22,** 1334–1349 (2012).
26. Blatti, C. & Sinha, S. Motif Enrichment Tool. *Nucleic Acids Res.* **42,** W20–W25 (2014).
27. Bailey, T. L. *et al.* MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37,** W202–W208 (2009).
28. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26,** 841–842 (2010).
29. Moorad, J. A. & Hall, D. W. Mutation Accumulation, Soft Selection and the Middle-Class Neighborhood. *Genetics* **182,** 1387–1389 (2009).
30. Morrow, E. H., Stewart, A. D. & Rice, W. R. Assessing the extent of genome-wide intralocus sexual conflict via experimentally enforced gender-limited selection. *J. Evol. Biol.* **21,** 1046–1054 (2008).
31. 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.* 6527–34 (2005).
32. Wei, Z., Wang, W., Hu, P., Lyon, G. J. & Hakonarson, H. SNVer: a statistical tool for variant calling in analysis of pooled or individual next-generation sequencing data. *Nucleic Acids Res.* **39,** e132–e132 (2011).
33. Ferrer-Admetlla, A., Leuenberger, C., Jensen, J. D. & Wegmann, D. An Approximate Markov Model for the Wright-Fisher Diffusion and its Application to Time Series Data. *Genetics* (2016).
34. Nomura, T. Effective size of populations with unequal sex ratio and variation in mating success. *J. Anim. Breed. Genet.* **119,** 297–310 (2002).

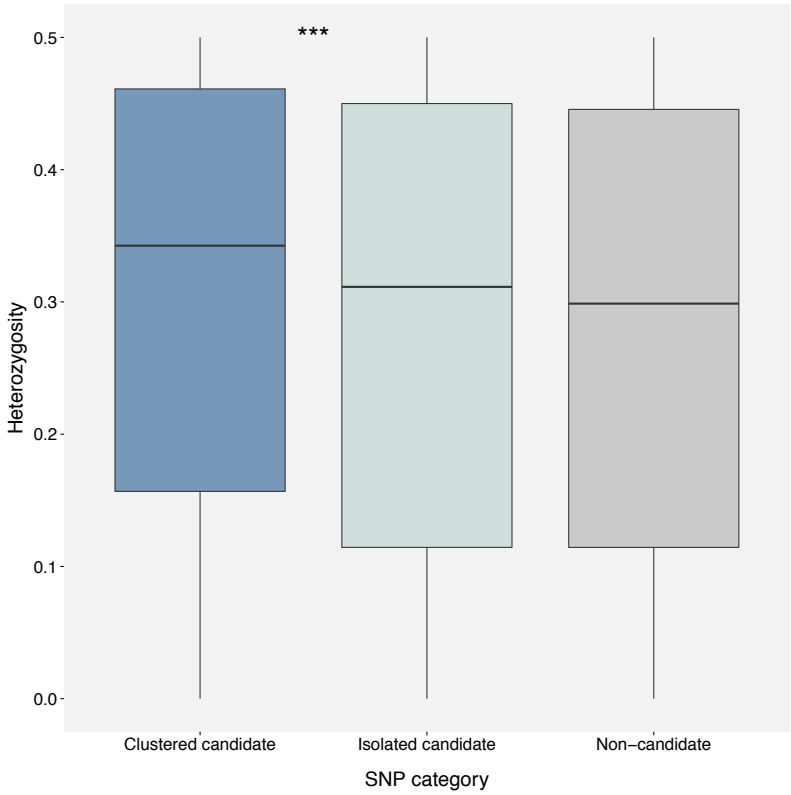
Supplementary tables and figures



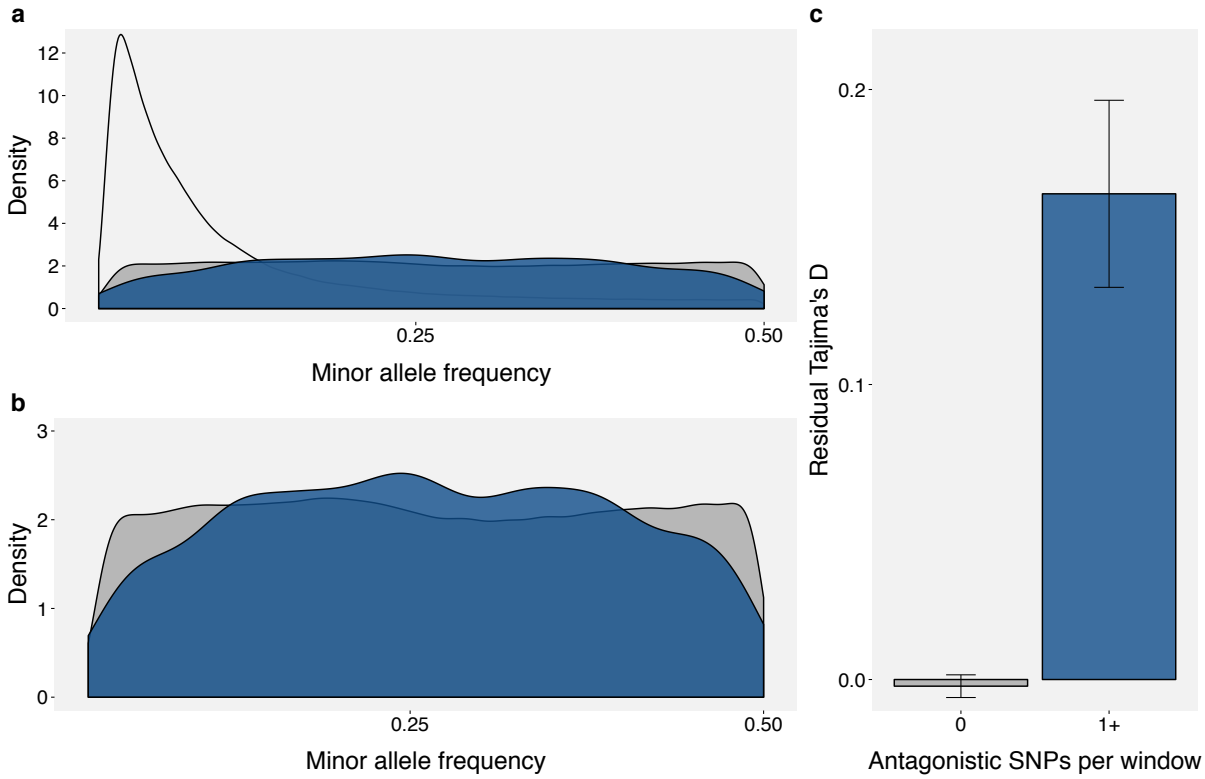
Supplementary figure 1 | Clustering analysis of candidate antagonistic SNPs along chromosomes. a, Null distribution of genetic distances between candidate SNPs on autosomal chromosome arms (median distance between pseudo-candidates = 9,055bp). The dashed red line indicates the observed median distance between autosomal candidate SNPs (187bp). **b,** Null distribution of genetic distances between candidate SNPs on the X chromosome (median distance between pseudo-candidates = 285,036bp). The dashed red line indicates the observed median distance between X-linked candidate SNPs (6,980bp). For both autosomal and X-linked primary candidate SNPs the observed median distance between candidate SNPs was significantly lower than expected ($P<0.001$).



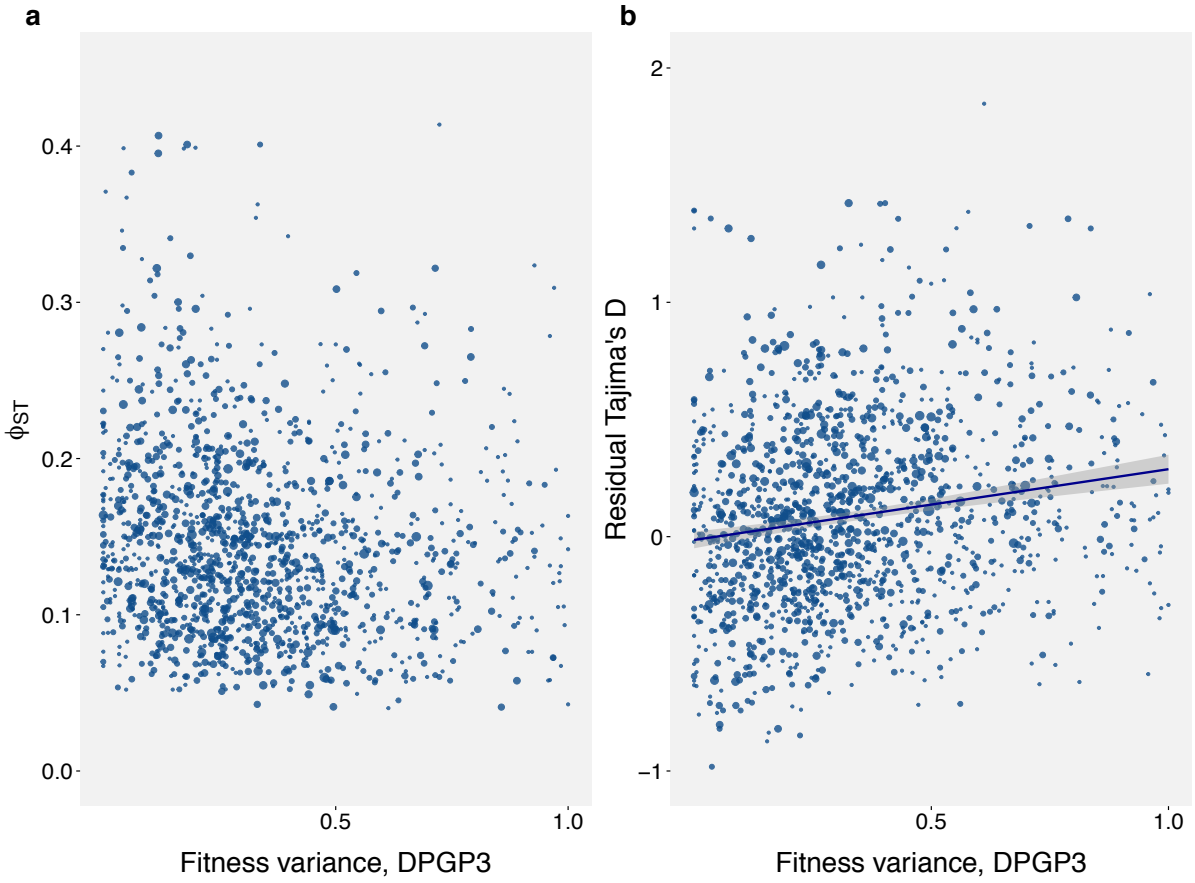
Supplementary figure 2 | Linkage disequilibrium-informed clustering of antagonistic SNPs. Linkage disequilibrium (r^2, calculated based on DGRP genomes) between the first and last SNP in each cluster, modelled as a declining exponential function of cluster length (i.e., chromosomal distance between the first and last SNP in each cluster). This relationship was modelled for varying numbers of clusters (k, see Supplementary methods). We chose a final k as the smallest possible k where the value of r^2 predicted by the model for the longest cluster exceeded 0.25. Plots shown here depict the observed relationship between r^2 and cluster length for that final value of k for each chromosome. The paucity of SNPs on the X chromosome precluded meaningful model fitting and we chose a value of k=30 (Supplementary methods).



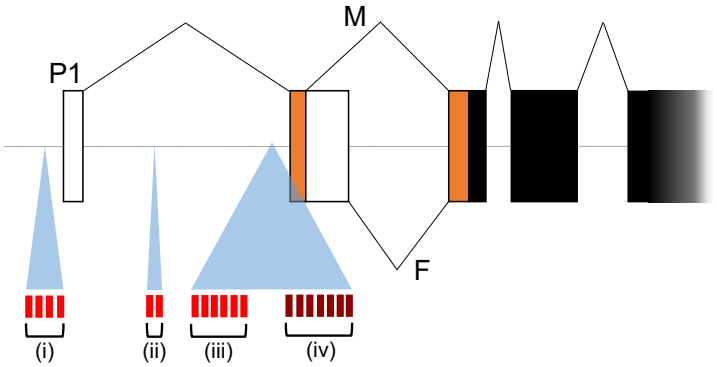
Supplementary figure 3 | Heterozygosity in the DPGP3 at clustered candidate SNPs, isolated candidate SNPs, and non-candidate SNPs. Isolated candidate SNPs resemble non-candidate SNPs more than clustered candidate SNPs. Expected heterozygosity in the DPGP3 at isolated candidate SNPs is not significantly different from expected heterozygosity at non-candidate SNPs (Wilcoxon test, P=0.312), but is significantly different from heterozygosity of clustered candidate SNPs (Wilcoxon test, P<0.001).



Supplementary figure 4 | Population genetic parameters in the source population, LH_M. a, Site frequency spectrum for highly covered (depth ≥100) SNPs in LH_M (from ref. 23 of the main text), for non-antagonistic SNPs (grey density surface, N=595,618) and antagonistic SNPs (blue density surface, N=3,146). For comparison, we also show the frequency spectrum of SNPs which were not polymorphic among the nine hemiclonal genomes and hence appear as neither antagonistic nor non-antagonistic in the current study (empty density surface, N=1,159,677). SNPs which are present in the haplotype sequencing (blue and grey surfaces) have comparatively intermediate minor allele frequencies. This is due to ascertainment bias where SNPs detected as polymorphic among the sample of 9 genomes tend to be those with elevated minor allele frequencies. **b,** Same as **a,** including only SNPs which appear in the haplotype sequencing. These two categories of SNP do not differ significantly in terms of their heterozygosity (Wilcoxon test, P=0.740). **c,** Mean±s.e.m. of Tajima's D, corrected for the effects of linked selection, in LH_M. Values are calculated over sliding windows (window size=1000bp, step size=500bp). Blue bars indicate windows containing one or more antagonistic SNPs (N=1,339), grey bars indicate windows with only non-antagonistic SNPs (N=97,821). Antagonistic windows display significantly elevated Tajima's D ($F_{1,99158}$=21.82, P<0.001).



Supplementary figure 5 | Relationship between fitness variance and window-wide estimates of balancing selection. The figures shown here relate fitness variance calculated per cluster to metrics calculated across corresponding chromosomal windows. They thereby make the link between window-based results on Tajima's D and Φ_{ST} presented in Figs. 2b and 2c of the main text with the cluster based results on the relationship between fitness variance and Φ_{ST} shown in Fig. 2e of the main text. **a,** Negative correlation between fitness variance, calculated over clusters of antagonistic SNPs, and Φ_{ST}, calculated over 1000bp sliding windows containing those clusters (Spearman rank-sum correlation, ρ=-0.144, P<0.001, N=1,461). Each point depicts a cluster of antagonistic SNPs; relative sizes denote the number of SNPs per cluster. Clusters where fitness variance is high tend to be at intermediate frequencies in both populations, thus lowering the proportion of variance explained by population (Φ_{ST}). **b,** Positive correlation between fitness variance, calculated over clusters of antagonistic SNPs, and Tajima's D, calculated over 1000bp sliding windows and corrected for the effects of linked selection ($F_{1,1460}$=45.54, β=0.302, P<0.001, N=1,461).



Supplementary figure 6 | Clusters of antagonistic SNPs in *fruitless*. The figure shows 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 (ref. 39 of the main text). 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). Antagonistic SNPs (red bars) are arranged in four clusters (i-iv). Cluster (i) spans 61bp and sits in vicinity of the P1 promoter. Clusters (ii), (iii), and (iv), spanning 16bp, 46bp, and 57bp, respectively, are close to the male-specific exon. (reference[39] in the main text). The antagonistic SNPs of cluster (iv), shown in dark red, were used to construct the *fruitless* haplotype network shown in Figure 2d of the main text.

Supplementary table 1 | Enriched motifs in regions spanning 5kb upstream of antagonistic candidate genes.

Motif name	Associated gene name	FDR q-value
AbdB_SOLEXA_FBgn0000015	*Abdominal B*	0.001200533
Achi_Cell_FBgn0033749	*achintya*	0.001200533
run_Bgb_NBT_FBgn0013753	*Big brother*	0.001200533
Exd_SOLEXA_FBgn0000611	*extradenticle*	0.001200533
Hth_SOLEXA_FBgn0001235	*homothorax*	0.001200533
run_Bgb_NBT_FBgn0003300	*runt*	0.001200533
Vis_SOLEXA_FBgn0033748	*vismay*	0.001200533
CG12236_SANGER_10_FBgn0029822	*CG12236*	0.009510096
z_FlyReg_FBgn0004050	*zeste*	0.013981706
dsf_SANGER_5_FBgn0015381	*dissatisfaction*	0.017227399
Unc4_SOLEXA_FBgn0024184	*unc-4*	0.017227399
zen2_SOLEXA_2_FBgn0004054	*zerknullt-related*	0.017227399
ato_da_SANGER_5_3_FBgn0010433	*atonal*	0.018548856
br-Z4_FlyReg_FBgn0000210	*broad*	0.018548856
ato_da_SANGER_5_3_FBgn0000413	*daughterless*	0.018548856
ems_FlyReg_FBgn0000576	*empty spiracles*	0.018548856
En_Cell_FBgn0000577	*engrailed*	0.018548856
HLH54F_da_SANGER_5_FBgn0022740	*HLH54F*	0.018548856
CG5669_SOLEXA_5_FBgn0039169	*Sp1-like factor for pairing sensitive-silencing*	0.018548856
tll_NAR_FBgn0003720	*tailless*	0.018548856
wor_SANGER_2.5_FBgn0001983	*worniu*	0.018548856
ase_da_SANGER_10_FBgn0000137	*asense*	0.022469907
l1sc_da_SANGER_5_FBgn0002561	*lethal of scute*	0.022469907
nau_da_SANGER_5_FBgn0002922	*Nautilus*	0.022469907
nub_NAR_FBgn0085424	*nubbin*	0.022469907
pad_SANGER_5_FBgn0038418	*poils au dos*	0.022469907
Vnd_SOLEXA_FBgn0003986	*ventral nervous system defective*	0.022469907
Ap_Cell_FBgn0000099	*apterous*	0.030615515
D_NAR_FBgn0000411	*Dichaete*	0.030615515
CG17181_SOLEXA_5_FBgn0035144	*Kahuli*	0.030615515
net_da_SANGER_10_FBgn0002931	*net*	0.030615515
Scr_Cell_FBgn0003339	*Sex combs reduced*	0.030615515
Six4_Cell_FBgn0027364	*Six4*	0.030615515
byn_FlyReg_FBgn0011723	*brachyenteron*	0.042770853
D19A_F10-12_SANGER_5_FBgn0022935	*D19A*	0.042770853
H2.0_SOLEXA_FBgn0001170	*Homeodomain protein 2.0*	0.042770853
lola-PT_SANGER_5_FBgn0005630	*longitudinals lacking*	0.042770853
Met_SANGER_5_FBgn0002723	*Methoprene-tolerant*	0.042770853

Supplementary table 2 | Enriched gene ontology process terms for antagonistic candidate genes.

GO Term	Description	FDR q-value
GO:0040003	chitin-based cuticle development	2.71E-06
GO:0010171	body morphogenesis	3.90E-06
GO:0042335	cuticle development	3.10E-06
GO:0048856	anatomical structure development	3.16E-04
GO:0006030	chitin metabolic process	9.16E-04
GO:1901071	glucosamine-containing compound metabolic process	2.44E-03
GO:0006040	amino sugar metabolic process	2.45E-03
GO:0009888	tissue development	6.09E-03
GO:0006022	aminoglycan metabolic process	6.45E-03
GO:0035295	tube development	1.14E-02
GO:0042659	regulation of cell fate specification	1.22E-02
GO:0010453	regulation of cell fate commitment	2.90E-02
GO:0009653	anatomical structure morphogenesis	3.33E-02
GO:0045861	negative regulation of proteolysis	3.09E-02
GO:0048634	regulation of muscle organ development	3.10E-02
GO:0001708	cell fate specification	3.18E-02
GO:1901135	carbohydrate derivative metabolic process	3.40E-02

Supplementary table 3 | Enriched gene ontology function terms for antagonistic candidate genes.

GO Term	Description	FDR q-value
GO:0005214	structural constituent of chitin-based cuticle	2.65E-04
GO:0042302	structural constituent of cuticle	4.41E-04
GO:0008061	chitin binding	4.10E-03
GO:0004046	aminoacylase activity	1.04E-02

Supplementary table 4 | Enriched gene ontology component terms for antagonistic candidate genes.

GO Term	Description	FDR q-value
GO:0005578	proteinaceous extracellular matrix	3.69E-03
GO:0031012	extracellular matrix	2.69E-02