Rapid evolution of the intersexual genetic correlation for fitness in *Drosophila melanogaster*

Julie M. Collet,[1,2,*] Sara Fuentes,[1,*] Jack Hesketh,[1] Mark S. Hill,[1] Paolo Innocenti,[3] Edward H. Morrow,[3,4] Kevin Fowler,[1] and Max Reuter[1,5]

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

[2]Current Address: School of Biological Sciences, University of Queensland, St. Lucia, QLD, Australia

[3]Department of Animal Ecology, Uppsala University, Uppsala, Sweden

[4]Current Address: School of Life Sciences, University of Sussex, Brighton, United Kingdom

[5]E-mail: m.reuter@ucl.ac.uk

Sexual antagonism (SA) arises when male and female phenotypes are under opposing selection, yet genetically correlated. Until resolved, antagonism limits evolution toward optimal sex-specific phenotypes. Despite its importance for sex-specific adaptation and existing theory, the dynamics of SA resolution are not well understood empirically. Here, we present data from *Drosophila melanogaster*, compatible with a resolution of SA. We compared two independent replicates of the "LH$_M$" population in which SA had previously been described. Both had been maintained under identical, controlled conditions, and separated for around 200 generations. Although heritabilities of male and female fitness were similar, the intersexual genetic correlation differed significantly, being negative in one replicate (indicating SA) but close to zero in the other. Using population sequencing, we show that phenotypic differences were associated with population divergence in allele frequencies at nonrandom loci across the genome. Large frequency changes were more prevalent in the population without SA and were enriched at loci mapping to genes previously shown to have sexually antagonistic relationships between expression and fitness. Our data suggest that rapid evolution toward SA resolution has occurred in one of the populations and open avenues toward studying the genetics of SA and its resolution.

Due to their different reproductive roles, male and female adults are often selected for different optimal phenotypes. However, the response to this divergent selection is limited by the fact that the sexes share a large part of their genome and, thus, new mutations frequently affect the phenotype of males and females in a similar way. The combination of genetically correlated male and female phenotypes and divergent selection on the sexes sets the scene for intralocus sexual conflict or sexual antagonism (SA), where some alleles increase the fitness in one sex at the expense of the fitness in the other sex (Rice 1984; Bondurinsky and Chenoweth 2009;

Van Doorn 2009; Connallon and Clark 2014). Sexually antagonistic genetic variation has been shown to segregate in natural and laboratory populations of a wide range of organisms, including insects (Chippindale et al. 2001; Fedorka and Mousseau 2004; Lewis et al. 2011; Berg and Maklakov 2012), vertebrates (Brommer et al. 2007; Foerster et al. 2007; Mainguy et al. 2009; Mokkonen et al. 2011), and plants (Kohorn 1994; Scotti and Delph 2006; Delph et al. 2011). This growing body of evidence demonstrates that the common genetic basis of male and female phenotypes limits the adaptive evolution of sex-specific traits. The adaptive trade-offs inherent in sexually antagonistic allelic variation prevent both sexes from attaining their sex-specific optima

*These authors contributed equally to this work.

and generate balancing selection that can maintain genetic variation at antagonistic loci. For this reason, sexual antagonism is also a powerful agent for the maintenance of genetic variation for fitness (Kidwell et al. 1977).

Despite being recognized as an important evolutionary force, the study of SA has so far focused mostly on characterizing antagonism as snapshots of particular populations at particular time-points and relatively little is known about its long-term dynamics (Stewart et al. 2010; Dean et al. 2012; Pennell and Morrow 2013). More specifically, it is currently unclear to what degree the intensity of antagonism changes over time, and over what timescale such changes occur. Similarly, there is scant information on whether individual antagonistic loci remain polymorphic over long periods of time and, if not, what mechanisms are involved in the fixation of one or the other allele. Answering these questions is vital for our understanding of ongoing conflicts over adaptation between the sexes, and of the evolution of sexual dimorphism.

Changes in the extent of SA can occur in response to variation in a population's environment as well as in its genetic composition. Experimental work in fruit flies has shown that the extent of SA varies between environmental conditions due to genotype-by-environment interactions (Delcourt et al. 2009; Punzalan et al. 2014). Large shifts in the environment might also eliminate antagonism (Long et al. 2012; Connallon and Clark 2014; Punzalan et al. 2014). This occurs if the change in conditions is large enough for the new selective optima of two sexes to both fall above (or below) their current trait values. This would then put both sexes under concordant directional selection and result in a positive genetic correlation between male and female fitness across genotypes.

The extent of SA can also vary in response to changes in the genetic composition of a population through mutation, drift, and selection. Mutations can increase the degree of antagonism if it generates similar phenotypic effects in the two sexes and hence increases their genetic correlation for traits under opposing selection (Connallon and Clark 2014). Genetic drift, in contrast, will result in a loss of genetic polymorphism at antagonistic loci, and hence can reduce the degree of SA observable within populations (Connallon and Clark 2012; Mullon et al. 2012; Hesketh et al. 2013). The rate at which this loss occurs depends on the effective population size at antagonistic loci, and thus will be faster in small populations, populations with large reproductive skew, and at loci under increased drift due to effects of chromosome dose (such as X or Z sex chromosomes) or selective interference between loci (Vicoso and Charlesworth 2009; Connallon and Clark 2012; Mullon et al. 2012).

Finally, and most importantly from an evolutionary point of view, the intensity of SA can be reduced through adaptive
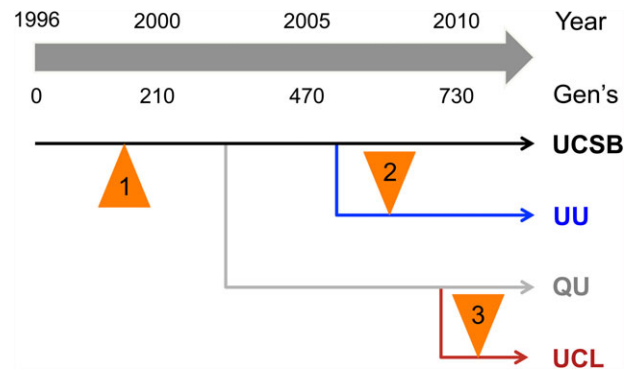
Figure 1. Schematic history of LH_M populations. The relationship between the LH_M populations used here, and the ancestral populations from which they were derived. The timeline is represented in calendar years and generations (approximate). For more details, refer to the Methods section of the main text. Orange triangles denote quantitative genetic studies of sex-specific fitness, 1: Chippindale et al. (2001), 2: Innocenti and Morrow (2010), 3: this study.

evolution. The trade-off between male and female fitness that underlies antagonism will create a selection pressure for mechanisms that diminish the deleterious fitness effects of the allele whenever it resides in the disfavored sex. These mechanisms are not well characterized empirically but could include sex-specific modifier loci (Rice 1984), sex-specific dominance (Kidwell et al. 1977; Barson et al. 2015), gene duplication followed by the evolution of sex-specific gene expression and adaptation in the two paralogs (Connallon and Clark 2011; Parsch and Ellegren 2013), genetic imprinting (Iwasa and Pomiankowski 1999, 2001; Day and Bondurianski 2004) or sex-specific splicing (Pennell and Morrow 2013). Such adaptations will result in the long-term resolution of antagonism and allow both sexes to approach their phenotypic optima, thereby increasing the degree of sexual dimorphism (Lande 1980; Ellegren and Parsch 2007; Bondurianski and Chenoweth 2009; Parsch and Ellegren 2013).

Understanding how antagonism evolves requires repeated measurements of SA at several time points. This can be achieved by monitoring SA in a single population through time or, alternatively, by measuring the extent of SA in different populations derived from a common ancestral population. Here, we present the results of such a comparative study investigating SA in two recently diverged replicate populations of the laboratory-adapted *D. melanogaster* stock LH_M. Both replicates were derived from the original LH_M population maintained by William Rice at the University of California, Santa Barbara, in which sexual antagonism had previously been documented (Chippindale et al. 2001). At the time of our analysis, they had been separated for about 200 generations but maintained according to the same strictly imposed rearing regime (see Methods, Fig. 1). Combining existing

data on the genetic architecture of male and female fitness in one population (LH$_M$-UU, Innocenti and Morrow 2010) with newly collected data on the other (LH$_M$-UCL), we show that the intersexual genetic correlation for fitness has diverged significantly between the two populations since they have been separated. While one population shows a negative genetic correlation between male and female fitness, indicative of antagonism, male and female fitness are not significantly correlated in the other, despite similar levels of heritable variation for fitness in the two sexes of both populations. Using a population genomic approach to compare allele frequencies at genome-wide SNP loci, we identify islands of significant differentiation between the populations, both on the X chromosome and the autosomes. We show that patterns of differentiation are biased toward frequency change in the population with reduced SA and enriched in genes with sexually antagonistic expression patterns. We argue that these results are unlikely to be due to environmental effects and compatible with the alleviation or resolution of sexual antagonism in one of the study populations. Given the increasing recognition of SA as a significant evolutionary force, our findings provide insights into the long-term evolutionary dynamics of SA and indicate future ways to elucidate the mechanistic basis of antagonism.

Materials and Methods

STUDY POPULATIONS

This study uses two replicate populations derived from the outbred laboratory stock LH$_M$ (here LH$_M$-UCSB), maintained by W. Rice at the University of California, Santa Barbara (Fig. 1). LH$_M$-UCL was established from a duplicate of LH$_M$-UCSB that was taken to Queen's University by A. Chippindale in February 2002 and then transferred to the Reuter group in May 2009. Independently, a replicate of LH$_M$-UCSB was taken to the Morrow group, University of Uppsala, in December 2005 to establish the other population, LH$_M$-UU.

Starting with the establishment of LH$_M$-UCSB in 1996, all LH$_M$ populations have been continuously maintained at a constant adult population size of 1792 individuals (896 males and 896 females) and under the same strictly regimented 14-day rearing regime (described in Rice et al. 2005). Therefore, neither LH$_M$-UCL nor LH$_M$-UU had experienced population bottlenecks or more than subtle environmental shifts. The rigorous two-week cycle further allowed us to estimate the number of generations of separation between the populations, dating back to the split between LH$_M$-QU and in February 2002 (Fig. 1). At the time of sampling genotypes for fitness measurements in October 2007, LH$_M$-UU had undergone approximately 145 generations from that branching point. In LH$_M$-UCL, genotypes were sampled in June 2010, approximately 215 generations since the original split.

HEMICLONAL ANALYSIS OF MALE AND FEMALE FITNESS

We used hemiclonal analysis to measure the effects of haploid genomes on male and female fitness (see Abbott and Morrow 2011 for a review of the approach). Hemiclonal individuals share a common copy of chromosomes X, 2, and 3, which amount to 99.5% of an identical haplotype (all genes except for the 0.5% of the genome located on the "dot" fourth chromosome). The quantitative genetic analysis captures additive genetic effects of the hemiclonal X, 2, and 3 chromosomes (as well as additive effects of epistatic interactions between alleles within the hemiclonal haplotype) but averages the effects of epistasis or dominance between the hemiclonal haplotype and the genetic background (Rice et al. 2005).

The quantitative genetic analysis of fitness used here was closely modeled on previous studies (Chippindale et al. 2001; Pischedda and Chippindale 2006; Innocenti and Morrow 2010; see Supplementary Material and the previously cited studies for details). To measure fitness in hemiclones, crosses were performed to generate males and females that carry an identical hemiclonal genome complemented with random genetic material from the corresponding source population. The fitness of these flies was then measured under conditions that closely mimic the LH$_M$ rearing regime and in competition with a standard competitor stock. The competitors provide a point of reference from which to calculate the relative fitness of different experimental genotypes. Their exact identity and genetic composition is not important for the results generated, as long as their fitness is similar to those of the experimental flies. The competitor flies used here carried a homozygous *brown* (*bw*) mutation in a variable, outbred LH$_M$ background, ensuring competitiveness while allowing us to assign paternity to experimental and competitor males. The competitor stock was maintained following exactly the same regime as the wild-type LH$_M$ population.

Fitness measurements for the LH$_M$-UCL population

LH$_M$-UCL hemiclonal lines were established in June 2010 and their fitness was measured between July 2010 and September 2011. For all lines, fitness was assayed three times in each sex in blocks that included one replicate of each hemiclonal line and alternatingly assayed male and female fitness. Complete fitness data were obtained for 113 lines. In order to assess potential genotype-by-laboratory effects, we also assessed the fitness of nine of the 10 most sexually antagonistic lines created in the LH$_M$-UU population (Fig. 1 in Innocenti and Morrow 2010; see also section "Comparing fitness measures across laboratories" below and Fig. 2 and Fig. S1).

Fitness assays for LH$_M$-UCL were conducted in groups of 60 flies per vial, including 10 focal flies, 20 standard *bw* competitors of the same sex and 30 standard *bw* flies of the opposite sex.

For male assays, the flies were allowed to interact for 66 hours (days 11–14, interaction and oviposition phases of the 14-day rearing cycle) and fitness was measured as the proportion of offspring produced by the females of an assay vial that were sired by the focal hemiclonal males. For female assays, flies were allowed to interact for the 48 hours of the competition phase of the rearing regime (days 11 and 12 of the 14-day cycle) and fitness was measured as the average number of eggs laid by the focal hemiclonal females over the following 19.5 hours (the oviposition phase of the rearing regime). The average fecundity of *bw* competitors was also measured and included in standardized fitness measures (see section "Transformation of fitness data").

Fitness measurements for the LH$_M$-UU population

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

STATISTICAL ANALYSIS OF FITNESS

Transformation of fitness data

Before analysis, each set of fitnesses was standardized to remove block and vial effects. Thus, we defined fitness as the residuals of a linear model that decomposed raw individual fitness scores into a fixed effect of experimental block and random residual error (which includes the genotypic effect on fitness). For female fitness measures obtained for LH$_M$-UCL, we further included a fixed effect describing the productivity of individual competition vials, as measured by the average number of offspring produced by *bw* competitor females. The residuals of these models were Z-transformed separately for each sex and population in order to obtain metrics of fitness that were comparable across sexes and populations.

Comparing fitness measures across laboratories

We used nine hemiclones from the LH$_M$-UU sample to verify that fitness across populations was measured in a repeatable and comparable way. These hemiclones constitute the extremes of the fitness distribution of the LH$_M$-UU sample (four with extremely female-beneficial/male detrimental and five with extremely male-

beneficial/female-detrimental effects) and had been maintained in the Uppsala laboratory. Complementing their fitness measures originally obtained in Uppsala, the fitness of these hemiclones was assayed again at UCL alongside those of the LH$_M$-UCL sample, using exactly the same experimental protocol as that used for all other UCL hemiclones. To compare the fitness scores across laboratories, we performed correlation analyses, separately for each sex, between the standardized fitness scores obtained in the original analysis by Innocenti and Morrow (2010) and those obtained in the experiments at UCL. In addition, we applied analyses of variance, again separately for each sex, to the standardized fitness data from both laboratories, modeling fitness as a function of hemiclone (G), laboratory (E), and their interaction (GxE).

Estimation and comparison of genetic variance components

We used mixed models to estimate the contribution of additive genetic effects of hemiclones to the variation in male and female fitness, and the covariance between these genetic effects on fitness in males and females. Prior to analysis, we removed one outlier hemiclone from the UCL dataset that had a very low male and female fitness (Fig. 2), compatible with the effects of a strongly deleterious mutation affecting both males and females. Removing this outlier was conservative because, if included, it artificially increased the estimates of heritabilities and the intersexual genetic correlation in the UCL population.

We used WOMBAT (Meyer 2007) to fit the following multivariate animal model:

$$Y = Zu_Y + e_Y, \qquad (1)$$

where Y is the vector of standardized fitness scores (hemiclones, sexes, and populations concatenated), Z is the incidence matrix defining the population-specific combination of sex and genotype for each fitness value, u_Y is the vector of sex- and population-specific additive genetic effects for fitness and e_Y is the vector of residual effects (Meyer 1991; Lynch and Walsh 1998).

We estimated a cross-population genetic (co)variance matrix

$$G = \begin{pmatrix} G_{mf_{UCL}} & 0 \\ 0 & G_{mf_{UU}} \end{pmatrix}.$$

In this matrix, $G_{mf_{UCL}}$ and $G_{mf_{UU}}$ are the population-specific additive genetic variance-covariance matrices,

$$G_{mf} = \begin{pmatrix} \sigma^2_{a,m} & cov_{mf} \\ cov_{mf} & \sigma^2_{a,f} \end{pmatrix},$$

where $\sigma^2_{a,m}$ and $\sigma^2_{a,f}$ are the additive genetic variances for fitness in males and females and cov_{mf} is the intersexual additive genetic covariance. Due to the standardization of our fitness data to a mean fitness of 0 and a standard deviation in fitness of 1 within each
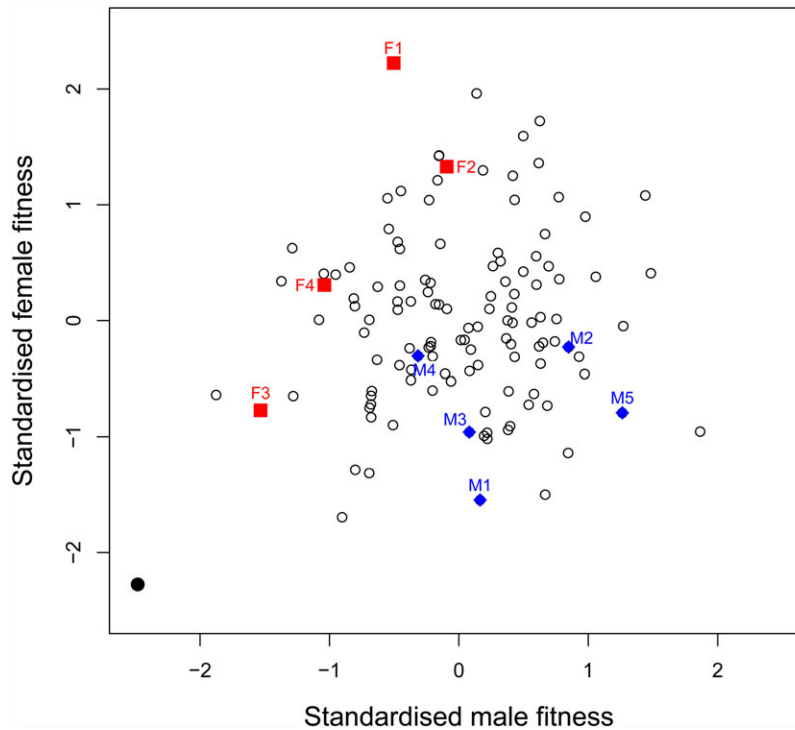
Figure 2. Male and female adult fitness across genotypes in the LHM-UCL population. Average male and female fitness across 113 hemiclonal lines randomly extracted from LH_M-UCL (open circles). One line (filled black circle) showed extremely low fitness in both sexes and was removed from further analyses. Fitness measures obtained at UCL for a set of hemiclones from LH_M-UU previously assayed as part of Innocenti and Morrow (2010) are also shown. The blue diamonds and red squares show the UCL fitness estimates of the hemiclones from this set that were classed as male beneficial/female detrimental and female beneficial/male detrimental fitness lines, respectively, in that previous study. Labels identify individual hemiclones for comparison with their fitness values in the previous study, shown in Fig. S1.

sex and population, heritabilities are directly given by additive genetic variances, $h_m^2 = \sigma_{a,m}^2$ and $h_f^2 = \sigma_{a,f}^2$. The intersexual genetic correlation for fitness can be calculated from the elements of the genetic variance-covariance matrix as $r_{mf} = \frac{cov_{mf}}{\sqrt{\sigma_{a,m}^2}\sqrt{\sigma_{a,f}^2}}$.

Estimation of genetic effects in the animal model relies on the numerator relationship matrix A that describes shared additive genetic effects between pairs of individuals. When defining this matrix, we considered hemiclonal males and females as full-sibs ($A_{ij} = 1/2$ for individuals i and j that are part of the same hemiclonal line) and all other pairs as unrelated ($A_{ij} = 0$ for pairs from different hemiclonal lines). The model was fitted using Restricted Maximum Likelihood (REML) and parameters and their approximate sampling errors (sensu Meyer and Houle 2013; Houle and Meyer 2015) were estimated.

The significance of parameter estimates and their differences between populations were tested using Log-likelihood Ratio Tests (LRTs) based on the X^2 distribution. These compared the full model (eq. 1) in which all parameters were freely estimated to simpler nested models in which specific parameters had been fixed to appropriate values (function FIXVAR in WOMBAT). Specifically, we tested for within-population differences between

the heritabilities h_m^2 and h_f^2 of fitness by comparing the full model to a model where genetic variances $\sigma_{a,m}^2$ and $\sigma_{a,f}^2$ were fixed to the average of their values estimated in the full model. To test whether the genetic covariance between male and female fitness cov_{mf} within a population was significantly different from 0, we compared the full model to a model where the covariance term for that population was fixed to 0. Similarly, to test for between-population differences in either the male or female fitness heritability or the covariance between male and female fitness, we compared the full model to models in which the values of these parameters in each population were fixed to their average between the UU and UCL estimates obtained in the full model.

DNA extraction and sequencing
We sampled 165 female adult flies from each of the populations in March 2012, about 260 generations after the separation of the two populations and about 46 and 115 generations after hemiclonal genomes had been sampled from LH_M-UCL and LH_M-UU, respectively. We extracted total genomic DNA from homogenized flies pooled by population using DNeasyBlood and Tissue Kit (Qiagen) and purified it using Agencourt AMPure XP

beads (Beckman Coulter). One paired-end Illumina library (insert size <500 bp) was made from each pool using the Nextera DNA Sample Preparation Kit (Illumina). Libraries were sequenced on an Illumina HiSeq2000 machine at the Centre for Genomic Research, University of Liverpool. Sequencing reads were extracted using CASAVA (version 1.82). Paired reads were trimmed using Sickle (version 1.2, default settings) and deduplicated using Picard (version 1.77, http://picard.sourceforge.net) before being aligned to the *D. melanogaster* reference genome (BDGP5.25.60) using Bowtie2 (version2.0.0-beta7, option -X 500). We removed regions flanking indels (±5bp) with Popoolation (version 1.2.2, Kofler et al. 2011a) and used RepeatMasker (Smit et al. 2006, version 4.0.5 with option species = drosophila and flag no_is) to mask interspersed repeats and low complexity regions. Finally, we applied a minimum read depth filter of 100 to ensure adequate precision of estimated allele frequencies, and a maximum read depth filter of 290 (about twice the average read depth of our sequencing runs, see Table S2), to avoid false-positive SNPs due to duplicated genomic regions.

SNP detection and analysis

We called SNPs in each population separately using SNVer (Wei et al. 2011, version 0.4.1 release 4, function for pooled sequencing SNVerPool.jar with minimum read and mapping quality cutoffs mq = 20 and bq = 20, haploid pool size $n = 330$ and no filtering by minor allele frequency $t = 0$). SNVer detects the significant presence of reads with the alternative allele (rather than polymorphism), so we reran the program on sites with high frequencies of the alternative allele, but using a reference genome sequence in which the corresponding positions were flipped to the alternative allele (i.e., testing for significant presence of the reference allele). We considered that a SNP was present if both polymorphism tests were significant. All *P*-values were corrected for false discovery rate (FDR < 0.05) using the package *q*-value (http://bioconductor.org). For all SNPs identified in this way, allele frequencies were extracted from the SNVer output. Furthermore, we calculated F_{ST} as a measure of genetic differentiation between LH_M-UU and LH_M-UCL using PoPoolation2 (Kofler et al. 2011b).

In order to summarize the chromosomal distribution of candidate SNPs with significantly elevated F_{ST} (see Results), we calculated the median distance between all pairs of adjacent candidate SNPs, separately for the X chromosome and the autosomes. We tested for significant clustering of candidate SNP by randomly permuting "candidate" and "noncandidate" labels among SNP loci and recalculating the median distance among SNPs labeled "candidate." Null distributions for the median distances between candidate SNPs were generated from 1000 such permutations (again, separately for X-linked, and autosomal markers). A *P*-value for significant clustering was calculated as the proportion

of median distances in the null distribution that was smaller or equal to the observed median distance.

Functional characterization of selected SNPs

We used the Variant Effect Predictor tool from Ensembl (McLaren et al. 2010) to map all SNPs to annotated genes and infer the consequences of variants (such as synonymous or nonsynonymous coding sequence changes, splice variants, etc.). The tool uses extended gene regions that span 5 Kb up- and downstream of the gene coordinates. In line with this default setting, we considered genes with candidate SNPs in that range as candidate genes. Analyses ignoring up- and downstream variants in these extended regions provided qualitatively identical results.

To assess the overlap between candidate genes in our study and previously described genes with sexually antagonistic expression patterns, we matched the identifiers of candidate genes to the corresponding Affymetrix *Drosophila* 2 probeset IDs. We then used the FlyAtlas data (Chintapalli et al. 2007) to consider only probesets that were expressed in adults, conservatively defined as those detected as "present" in at least one library of one FlyAtlas adult tissue sample. The list of adult-expressed probesets covered by our study was then matched to the list of genes with antagonistic expression from Innocenti and Morrow (2010, Table S1), retaining only those probesets in their study that were covered by our SNP data, and matches were back-translated into FlyBase gene identifiers.

We used Gene Ontology (GO) analysis implemented in DAVID (Huang et al. 2009a,b) to obtain insights into the functions of candidate genes. For enrichment analyses, we used all genes covered in our SNP dataset as the background set and applied a Benjamini–Hochberg correction for multiple testing to the *P*-values of individual tests. We further applied DAVID's Functional Annotation Clustering (using the "high" stringency setting) to categorize enriched individual GO terms into groups of related annotations.

Results
HEMICLONAL FITNESS CAN BE MEASURED RELIABLY ACROSS LABORATORIES

A meaningful comparison between the genetic architecture of fitness across populations requires that the fitness of different genotypes be measured in a reliable and comparable way and in the absence of significant genotype-by-environment effects. Analyses of the fitness data from our nine reference hemiclones suggest that this is true here. First, we observed significant positive correlations between the sex-specific fitness measures obtained in the original experiments at the University of Uppsala and those obtained on the same hemiclones at UCL (females: $r = 0.81, t_7 = 3.67, P = 0.008$; males: $r = 0.73, t_7 = 2.79, P = 0.027$, Fig. S2). Second, ANOVAs

performed on male and female fitness returned highly significant differences between hemiclones across laboratories, but no hemiclone-by-laboratory effects (females: hemiclone—$F_{8,45}$ = 11.42, $P < 0.0001$, hemiclone-by-laboratory—$F_{8,45}$ = 1.19, $P = 0.33$; males: hemiclone—$F_{8,63}$ = 11.34, $P < 0.0001$, hemiclone-by-laboratory—$F_{8,45}$ = 1.54, $P = 0.16$; laboratory term nonsignificant in both analyses, as expected for Z-transformed fitness data). Based on the (additive) Sums of Squares, these analyses also show that, in line with the nonsignificant effect, the total variance in the data attributable to hemiclone-by-laboratory interactions is low (females: $SS_{hemiclone-by-laboratory}$/$SS_{total}$ = 0.065; males: $SS_{hemiclone-by-laboratory}$/$SS_{total}$ = 0.073; see Table S1 for full ANOVA tables). These results indicate that the fitness variation between genotypes detected in our experiments is overwhelmingly due to genetic differences, rather than the interaction between genotypes and the specific assay and laboratory environments.

THE INTERSEXUAL GENETIC CORRELATION FOR FITNESS DIFFERS BETWEEN POPULATIONS

We evaluated whether the genetic architecture of fitness, and in particular the intersexual genetic correlation for fitness, differed between the two replicate LH_M populations. Fitness heritabilities for LH_M-UU were slightly higher than the estimates obtained in the earlier analysis of the UU dataset (Table 1 in Innocenti and Morrow 2010), in line with the fact that our approach accounted for the environmental effects of experimental assays. Our estimates confirmed that the heritability of male and female fitness differed in the LH_M-UU population (X^2 = 12.2, $P = 0.0005$, Table 1) while showing that LH_M-UCL male and female heritabilities did not differ significantly ($X^2 = 0.7, P = 0.39$). Furthermore, male and female fitness heritabilities did not significantly differ across populations (X^2 = 2.6, $P = 0.11$ and $X^2 = 0.4$, P = 0.52, respectively).

While both populations featured ample and comparable heritable fitness variance, they differed in their intersexual genetic correlation. The point estimate of $r_{mf_{UU}}$ = –0.41 was significantly negative (X^2 = 4.5, $P = 0.03$) while the genetic correlation in LH_M-UCL was positive ($r_{mf_{UCL}}$ = 0.21) and not significantly dif-

ferent from 0 ($X^2 = 1.2, P = 0.27$). Further, the intersexual genetic correlations differed significantly between populations ($X^2 = 5.1$, $P = 0.02$, Table 1, Fig. 2 and Fig. S1). This indicates that the sexual antagonism that was present in LH_M-UU was absent in LH_M-UCL.

POPULATION GENOMICS REVEALED REGIONS OF SIGNIFICANT GENETIC DIVERGENCE BETWEEN THE POPULATIONS

We performed genome-wide pooled sequencing of flies from LH_M-UU and LH_M-UCL in order to identify SNP loci with significant allele frequency differences between the populations (see Table S2 for general sequencing statistics). These loci would be candidates for regions potentially functionally related to the change in the genetics of fitness observed in the quantitative genetic analysis. Our sequencing approach covered the entire genome and for completeness we present results for all chromosomes, even though the contribution of the small 4th chromosome to fitness variation was not measured in our phenotypic assays.

Population sequencing identified more than 680,000 high-quality SNPs with significant allelic variation in at least one of the two populations. The density of SNP loci varied between chromosome arms ($X_5{}^2$ = 44997.1, $P < 0.0001$). Chromosome arms 2L, 2R, and 3L were enriched for SNP polymorphism, chromosome arm 3R had slightly fewer SNPs than expected, and chromosomes X and 4 were severely depleted for polymorphic sites (Table S3). The lower SNP densities on chromosomes X and 4 are expected based on the lower effective population sizes of these chromosomes, caused by the lower numerical population size of the X chromosome relative to the autosomes (Charlesworth et al. 1987; Mank et al. 2010) and selective interference along the virtually nonrecombining chromosome 4 (Jensen et al. 2002; Haddrill et al. 2007; Betancourt et al. 2009; Charlesworth et al. 2010).

Analysis of the SNP allele frequencies showed that LH_M-UU was weakly, but significantly, more genetically diverse than LH_M-UCL. This difference was reflected at two levels. First, the percentage of SNP loci that were variable in LH_M-UU but fixed in LH_M-UCL (12.3%; Table S3) was greater than the percentage that only segregated in LH_M-UCL (8.8%; Proportion test:

Table 1. Heritabilities and genetic correlations for fitness in LH_M-UU and LH_M-UCL.

Population	Female h^2	Male h^2	r_{mf}
LH_M-UU	0.71 (0.15)	0.19 (0.07)	–0.41 (0.18)[1]
LH_M-UCL	0.58 (0.12)	0.41 (0.12)	0.21 (0.19)[2]
Difference between populations	$X_1^2 = 0.4, P = 0.52$	$X_1^2 = 2.6, P = 0.11$	$X_1^2 = 5.1, P = 0.02$

The table provides estimates and, in parentheses, the approximate sampling error for the heritabilities of male and female fitness and the intersexual genetic correlation (r_{mf}) for fitness in the two populations, as well as the results of likelihood ratio tests comparing estimates between populations (see Methods).
[1]Likelihood ratio test comparing the estimate to zero (see Methods): $X_1^2 = 4.5, P = 0.03$.
[2]Likelihood ratio test comparing the estimate to zero (see Methods): $X_1^2 = 1.2, P = 0.27$.
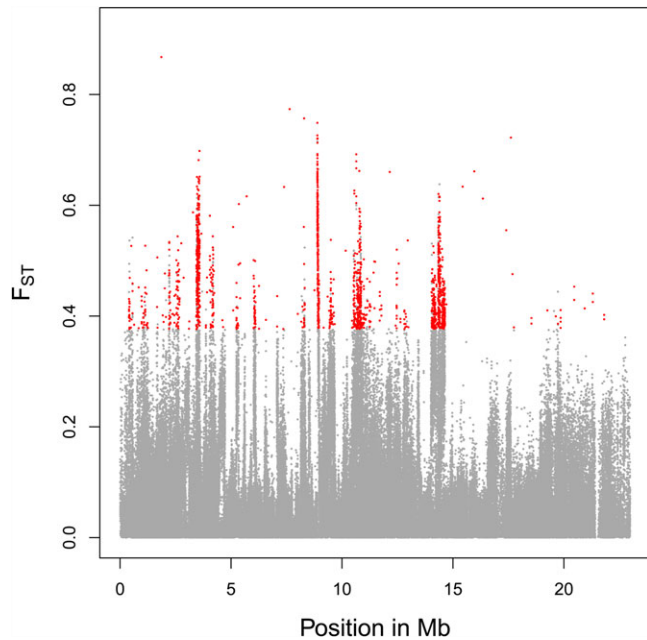
Figure 3. F_{ST} variation along SNPs in chromosome arm 2L. Gray dots indicate noncandidate loci, red dots candidate loci.

Table 2. Distribution of candidate and noncandidate SNP loci across chromosomes.

| Chr | Candidate | | Noncandidate | |
	Count	Percentage	Count	Percentage
X	717	*16.00*[+]	68,587	10.11
2L	889	*19.85*[−]	155,525	22.93
2R	634	*14.16*[−]	144,559	21.31
3L	493	*11.01*[−]	155,422	22.91
3R	1746	*38.98*[+]	153,456	22.62
4	0	0	804	0.12
Autosomes	3762	*84.00*[−]	609,766	89.89
Total	4479	100	678,353	100

Percentages of candidate SNPs on each chromosome were tested for significant over- or underrepresentation relative to noncandidate SNPs using one-sample *Z*-tests and *P*-values were Bonferroni-corrected for multiple testing. Fractions in *italics* are significantly over- (+) or underrepresented (−).

$X_1^2 = 4575.2$, $P < 0.0001$). Second, expected heterozygosity (H_e) was slightly higher in LH_M-UU than LH_M-UCL (Fig. S3), both for autosomal SNPs ($H_{e\text{-auto,UU}} = 0.295 \pm 0.165$ mean $\pm$ SD; $H_{e\text{-auto,UCL}} = 0.278 \pm 0.173$; paired *t*-test: $t_{61,3527} = 73.378$, $P < 0.0001$) and for X-linked SNPs ($H_{e\text{-X,UU}} = 0.284 \pm 0.166$; $H_{e\text{-X,UCL}} = 0.264 \pm 0.176$; paired *t*-test: $t_{69,303} = 26.350$, $P < 0.0001$; note that these results are robust to corrections for possible nonindependence between sites, e.g., including only every 10th, 20th, or 50th SNP). These patterns also show again that X-linked variation was smaller than autosomal variation in both populations.

We estimated the genetic differentiation between the populations at each SNP locus by calculating the fixation index F_{ST} (Fig. 3 for chromosome arm 2L as an example, Fig. S4 for all chromosomes). The average level of differentiation across autosomal SNP loci was $F_{ST} = 0.054 \pm 0.073$, and that on the X chromosome was $F_{ST} = 0.071 \pm 0.099$. In order to identify SNP loci where genetic differentiation significantly exceeded the level expected from random genetic drift, we used the 105,851 SNPs causing synonymous variation (amino acid and stop codons). Synonymous allelic variation can be considered nearly neutral and used to establish an empirical null distribution of neutral background differentiation, against which sites under potential selection can be compared (e.g., Mcdonald and Kreitman 1991; Smith and Eyre-Walker 2002; Andolfatto 2005; Haddrill et al. 2010). We defined cut-off values for selective divergence between populations as F_{ST} values that exceeded the 99th quantile of the F_{ST} distribution across synonymous sites (Fig. S3). We did so sep-

arately for the X chromosome and the autosomes, in order to accommodate the different intensities of genetic drift acting on these two genomic compartments. The cut-offs used were $F_{ST} > 0.371$ for autosomes and $F_{ST} > 0.453$ for the X chromosome. Applying these cut-offs to our entire SNP dataset (excluding the synonymous SNPs, which were used to define the cut-offs), 3762 autosomal and 717 X-linked SNP loci showed above-threshold levels of population differentiation (Fig. S5). The average absolute frequency difference between the two populations at these loci was 0.666 ± 0.091 (median absolute difference 0.658). Independent verification using allele counts confirmed our F_{ST}-based definition of population differentiation, as allele frequencies differed significantly between the populations at all candidate loci, even when using a stringent Bonferroni correction (Fisher's Exact test on counts of reference and alternative alleles, $P < 0.05/4479 = 1.12 \, 10^{-5}$ for all loci).

The distribution of the candidate SNPs was nonrandom across the chromosome arms and differed from that of noncandidate SNPs ($X_5^2 = 1082.1$, $P < 0.0001$). Specifically, candidate SNPs were overrepresented on the X chromosome and chromosome arm 3R and underrepresented elsewhere (Table 2). In addition to showing uneven distributions between chromosomes, candidate loci showed a clustered distribution along the chromosome arms and a large proportion of them fell within 100 bp from each other (Fig. S6). A permutation test (see Methods) confirmed statistically significant clustering and showed that the observed distance between candidate SNPs was much smaller than expected by chance (autosomes: observed median distance = 770 bp, range of median distances among 1000 sets of permuted loci = [12,998 bp–14,978 bp], $P < 0.001$; X chromosome: observed median distance = 1292 bp, range of permuted median distances = [15,592 bp–20,567 bp], $P < 0.001$).

The local clustering of candidate SNPs could indicate population differentiation in the frequency of chromosomal inversions, leading to parallel frequency changes of large numbers of alleles linked within the inverted part of the chromosome. Such effects have been observed in frequency clines among populations (Fabian et al. 2012) and frequency changes in response to laboratory selection (Kapun et al. 2014). In order to assess this possibility, we inspected patterns of polymorphism for diagnostic SNP alleles linked to seven cosmopolitan *D. melanogaster* inversions identified by Kapun et al. (2014). All but a few of these marker positions were well covered with high-quality reads in both samples (Table S4), but the data indicate that none of the inversions segregate in our populations. Only two of the inversion markers showed significant polymorphism in our samples (Table S4) and in both these cases the marker allele is the reference allele, suggesting either homoplasy (where the site on the inversion has convergently mutated back to the ancestral state) or an error in the inference of the marker allele.

CANDIDATE SNPS SHOW BIASED PATTERNS OF ALLELE FREQUENCY CHANGE

Population differentiation and elevated F_{ST} at candidate sites could arise due to allele frequency change in LH_M-UCL, in LH_M-UU or in both. In order to distinguish between these possibilities and infer the directionality of evolution at candidate SNPs, we used genotype data from the Drosophila Genetics Reference Panel (DGRP, Mackay et al. 2012; Huang et al. 2014). The DGRP constitutes an independent sample of genetic diversity derived, like the LH_M, from a wild North American population of *D. melanogaster*. In the absence of data on the ancestral LH_M population, it therefore provides a suitable external reference point to polarize frequency changes between LH_M-UCL and LH_M-UU. Focusing on the 636,924 SNPs that were shared and biallelic across the DGRP and LH_M samples, we found that levels of genetic differentiation between each of the LH_M populations and the DGRP differed markedly between noncandidate and candidate SNPs. For noncandidate SNPs, F_{ST} values between LH_M-UCL and the DGRP were of similar magnitude—although marginally (and significantly) higher—than those between LH_M-UU and the DGRP (autosomes: $F_{ST, UCL-DGRP} = 0.099 \pm 0.127$ mean $\pm$ SD, $F_{ST, UU-DGRP} = 0.091 \pm 0.118$, mean pairwise difference $= 0.0073$, CI $= [0.0070, 0.0076]$, $t_{568070} = 47.739$, $P < 0.0001$; X chromosome: $F_{ST, UCL-DGRP} = 0.105 \pm 0.126$, $F_{ST, UU-DGRP} = 0.092 \pm 0.114$, mean pairwise difference $= 0.0128$, CI $= [0.0118, 0.0138]$, $t_{64464} = 24.904$, $P < 0.0001$). For candidate SNPs, in contrast, we observed a large excess of differentiation between LH_M-UCL and the DGRP, compared to that between LH_M-UU and the DGRP (autosomes: $F_{ST, UCL-DGRP} = 0.270 \pm 0.224$, $F_{ST, UU-DGRP} = 0.145 \pm 0.167$, mean pairwise difference $= 0.124$, CI $= [0.113, 0.136]$, $t_{3634} = 21.348$, $P < 0.0001$; X chromosome: $F_{ST, UCL-DGRP} =$ 0.315 $\pm$ 0.235, $F_{ST, UU-DGRP} = 0.142 \pm 0.154$, mean pairwise difference $= 0.173$, CI $= [0.146, 0.201]$, $t_{707} = 12.548$, $P < 0.0001$). Echoing these quantitative differences, a disproportionally large number of candidate loci showed greater F_{ST} between LH_M-UCL and the DGRP than between LH_M-UU and the DGRP (autosomes: 2327 of 3635 loci with $F_{ST, UCL-DGRP} > F_{ST, UU-DGRP}$, $X_1^2 = 202.660$, $P < 0.0001$; X chromosome: 458 of 708 loci, $X_1^2 = 34.168$, $P < 0.0001$). These results show that differentiation between the two LH_M populations at candidate sites is disproportionally driven by allele frequency change in LH_M-UCL.

CANDIDATE SNPS HAVE NONRANDOM FUNCTIONS

To understand the functional relevance of the candidate SNPs, we used the Variant Effect Predictor tool to examine the genetic consequences of variants segregating at candidate and noncandidate sites. When compared to noncandidate SNPs, candidate SNPs were overrepresented among nonsynonymous coding polymorphisms (amino acid changes, loss and addition of translation start sites, stop codons; Table 3). In contrast, we observed an underrepresentation of candidate SNPs in intergenic regions (i.e., those further than 5 Kb away from any annotated gene). These nonrandom patterns of enrichment and depletion are indicative of a functional role for candidate polymorphisms.

Candidate SNPs mapped to a total of 1131 genes, 939 on the autosomes and 192 on the X chromosome. We performed GO term enrichment analyses to gain information on the biological processes in which these genes are involved and their molecular function. Term-by-term analyses revealed a strong association of candidate genes with biological processes related to growth, development, and differentiation (Table S5A). These trends were further highlighted in subsequent clustering of enriched GO terms, where eight out of the ten most enriched term groupings were related to development (Table S6). In terms of molecular function, we found significant enrichment for two terms, both related to transcription regulation (Table S5B; no clustering was performed due to the small number of significant terms).

To relate our results to previous analyses of sexual antagonism, we compared our list of candidate genes to genes previously shown to have sexually antagonistic expression patterns. Innocenti and Morrow (2010) used a combination of phenotypic fitness assays and whole-fly microarray expression analysis to identify genes that showed sex-differences in the relationship between expression level and fitness, mostly due to opposing associations of expression levels with male and female fitness (positive correlation between expression level and female fitness but negative correlation between expression level and male fitness, or vice versa). We observed a 30% excess of overlap between genes with candidate SNPs in our study and genes with such antagonistic expression patterns (147 overlapping genes to 112 expected, $X_1^2 = 13.396$, $P = 0.0003$). No excess overlap was seen between

Table 3. Comparison of genomic feature distributions between all SNPs, candidate SNPs, and noncandidate SNPs.

Functional category	All		Candidate		Noncandidate	
	Count	Percentage	Count	Percentage	Count	Percentage
UTR	42,706	6.25	312	6.97	42,394	7.41
Nonsynonymous	33,483	4.90	312	6.97^+	33,171	5.79
Synonymous	105,851	15.50	0	–	105,851	–
Splice site	6484	0.95	61	1.36	6423	1.12
Intron/noncoding exon	289,792	42.44	2286	51.04	287,506	50.22
Up-/downstream	135,364	19.82	1110	24.78	134,254	23.45
Intergenic	69,152	10.13	398	8.89^-	68,754	12.01

Proportional representations of functional categories were tested for significant over- or underrepresentation of candidate SNPs compared to noncandidate SNPs using one-sample *Z*-tests and *P*-values were Bonferroni-corrected for multiple testing. Categories in *italics* are significantly over- (+) or underrepresented (−). Synonymous polymorphisms were not considered while performing these tests, as they had been used to identify candidate SNPs and hence not tested here. For easier comparison, the synonymous variants were also excluded when calculating the percentages of candidate and noncandidate variants that fall into the different categories.

our candidates and Innocenti and Morrow's lists of genes show-ing expression levels that were only associated with male fitness ($X_1^2 = 0.515$, $P = 0.473$) or only with female fitness ($X_1^2 = 1.821$, $P = 0.177$). Thus, genes with antagonistic expression patterns were more likely to show large population differentiation in our study, but this association was not due to fitness-related genes being generally enriched among candidates in our study.

Finally, we investigated patterns of sex-biased gene expres-sion among our candidate genes. Sex-biased expression is thought to be involved in the resolution of sexual antagonism. Accord-ingly, genes that show sex-biased expression should be less prone to antagonistic fitness effects and, inversely, genes with sexually antagonistic variation should have lower than average sex-biased expression. We compared our dataset to the Sebida database (Gnad and Parsch 2006) that classifies *D. melanogaster* genes as male-, female-, or unbiased, based on the integration of multiple expres-sion datasets. Our candidate genes show an underrepresentation of sex-biased gene expression, relative to noncandidate genes. This is true when pooling male- and female-biased genes ($X_1^2 = 22.241$, $P < 0.0001$), and also when analyzing male- and female-biased genes separately (male-biased: $X_1^2 = 6.397$, $P = 0.0114$; female-biased: $X_1^2 = 27.627$, $P < 0.0001$). This suggests that candidate genes tend to be those where, if present, antagonism could be expected to be more pronounced, as it is not tempered by sex-biased expression.

Discussion

Despite the well-documented prevalence of sexual antagonism in plant and animal populations (e.g., Kohorn 1994; Chippindale et al. 2001; Fedorka and Mousseau 2004; Scotti and Delph 2006; Brommer et al. 2007; Foerster et al. 2007; Delph et al. 2011; Lewis et al. 2011; Mokkonen et al. 2011; Berg and Maklakov 2012) we

know relatively little about its evolutionary dynamics (Cox and Calsbeek 2009; Stewart et al. 2010; Dean et al. 2012; Pennell and Morrow 2013). In this study, we document differences in intersex-ual genetic correlations of fitness (r_{MF}) between two populations recently derived from the outbred laboratory stock LH_M in which sexual antagonism had originally been documented. While we found a significantly negative genetic correlation between male and female fitness in LH_M-UU, consistent with ongoing SA in this population, r_{MF} was not different from zero and antagonism absent in LH_M-UCL. These differences in the genetic architecture of sex-specific fitness were associated with significant shifts in allele frequencies at SNPs across the genome. These frequency changes occurred more often than expected at loci linked to genes that had been previously and independently linked to SA and more strongly affected frequencies in LH_M-UCL, the population with reduced SA. We argue that our data are suggestive of rapid evolution of r_{MF} and are, at the very least, compatible with recent partial resolution of SA in LH_M-UCL.

An important first step to interpreting our data is to estab-lish whether evolutionary change has in fact occurred between the two populations. While this is straightforward at the genetic level, where we directly demonstrate differences in allele fre-quencies between the populations, caution is required when inter-preting the observed difference in the genetic fitness correlation r_{MF}. Gene-by-environment interactions are known to cause differ-ences between estimates of genetic (co-)variances for male and female fitness obtained under different environmental conditions (Delcourt et al. 2009; Punzalan et al. 2014). Such effects could generate a spurious difference in r_{MF} here, because fitness mea-sures for LH_M-UU and LH_M-UCL were obtained under conditions that—despite our efforts at standardization—cannot be assumed to have been completely identical. Importantly, however, our in-clusion of a reference set of genotypes in experiments at both

Uppsala and UCL allowed us to rule out substantive genotype-by-laboratory effects. Comparison of fitness values between laboratories showed that the measures obtained at UCL correspond closely to those originally measured in Uppsala. Furthermore, the proportion of variation in male and female fitness that could be attributed to genotype-by-laboratory interactions was very small and not statistically significant. It therefore appears implausible that the observed differences in r_{MF} arose due to genotype-by-laboratory effects alone and can instead be assumed to reflect at least some evolutionary divergence in the genetic basis of fitness.

Several factors could have driven the evolutionary divergence between the populations. An important distinction is between neutral selective forces. Changes at the phenotypic (r_{MF}) and allele frequency level can arise as a consequence of genetic drift during founding events or in populations of small effective size. Random fixation of antagonistic alleles can reduce the extent of SA or even render it undetectable. At the same time, however, such fixation events would result in a loss of heritable genetic variation in sex-specific fitness and a genetic homogenization at the sequence level. These hallmarks of genetic drift do not appear in LH_M-UCL, where antagonism was absent. Although it is difficult to compare absolute levels of quantitative genetic variation between populations, heritable variation in male and female fitness was present in LH_M-UCL and tended to be greater than in LH_M-UU. Similarly, our molecular data are incompatible with the action of strong genetic drift in LH_M-UCL. Although the average heterozygosity was slightly reduced in LH_M-UCL compared to LH_M-UU, both populations showed comparable levels of allelic variation. Most importantly, loci with particularly elevated frequency differences showed many aspects of nonrandomness in direction (divergence relative to the DGRP), position (clustering), and function (variant effects, association with antagonistic genes) that are incompatible with the action of random genetic drift. It therefore seems likely that the divergence between LH_M-UU and LH_M-UCL constitutes adaptive responses to selection pressures in one or both populations.

The most likely candidates for selection in these well-established laboratory populations are (i) differences in the microenvironment in which the populations were maintained, (ii) ongoing coevolutionary dynamics between the sexes (interlocus sexual conflict) occurring independently within each population, or (iii) differences in the response to selection to alleviate the gender load (Rice 1992) that is generated by the sex-specific deleterious effects of sexually antagonistic alleles segregating in the populations. Our results cannot unambiguously differentiate between these alternative adaptive scenarios but we can assess the plausibility of each selective force in the light of the data. Environmental adaptation appears the least plausible scenario. It is conceivable that different selective pressure in the two laboratories could drive genetic divergence and potentially shifts in r_{MF}, especially when considering pathogens as part of the environment. Even though LH_M stocks are superficially healthy, they are not maintained in sterile conditions and can be expected to undergo selection due to viral and/or bacterial pathogens. Interestingly, the evolution of immune resistance and disease tolerance shows sex-specific effects, where mutations altering resistance and tolerance can do so in opposing ways in males and females (Vincent and Sharp 2014). However, once again, divergent selection pressures deriving from any part of the environment in Uppsala and UCL would be expected to generate genotype-by-laboratory variation in fitness among our reference genotypes. In contrast, we found a positive correlation in their fitness across laboratories that implies that environmental selection pressures in the two laboratories were aligned, with the same genotypes (and hence phenotypes) having high (or low) fitness in both places. If environmental selection pressures occurred, their effect would have to be slight and unlikely to drive the population differentiation observed over the short timespans considered here.

Coevolutionary arms races between the sexes occur whenever males and females differ in their reproductive interests. The promiscuous mating system imposed by the LH_M maintenance regime is expected to generate antagonistic interactions between males and females ("interlocus sexual conflict" or simply "sexual conflict," Chapman et al. 2003; Arnqvist and Rowe 2005), often over mating rates (Pischedda and Rice 2012). The evolutionary arms race driven by sexual conflict can generate strong selection pressures, driving rapid divergence between populations (Rice and Holland 1997; Gavrilets 2000; Martin and Hosken 2003; Debelle et al. 2014). However, how this evolution would affect antagonistic variation is currently unclear. Potential links between male–female coevolution and sexual antagonism have been proposed (Pennell and Morrow 2013), but they await theoretical and empirical investigation. While we cannot rule out sexual conflict as a driver of evolution, certainly our data do not support that link. Many of the molecular phenotypes involved in the male contribution to the arms race (such as accessory gland proteins) are encoded by genes that show male-specific expression while our candidate SNPs are preferentially associated with genes that show unbiased expression in both sexes.

A more plausible scenario is that LH_M-UCL adapted to selection generated by sexual antagonism. Sexually antagonistic variation can be stably maintained in populations under divergent selection on male and female traits (Gavrilets and Rice 2006), but any new variant that relieves the deleterious effect of antagonistic alleles in one sex while maintaining the benefit in the other is selectively favored (Rice 1984). The invasion of such variants will contribute to the resolution of antagonism and the evolution of further sexual dimorphism. Several aspects of our data are compatible with such change occurring in the LH_M-UCL population. First, the evolutionary change that we observe at the

phenotypic and the genetic levels establish independent links to antagonism. This is obviously true for the phenotypes, where we directly assess SA. However, the population genomic analysis, which is unbiased with regard to the phenotypic measures, also invokes SA through the significant overlap of our candidates and genes previously shown to have antagonistic expression patterns (Innocenti and Morrow 2010). Second, the directionality of phenotypic and genetic evolution indicates adaptation in LH_M-UCL, rather than symmetric divergence between the populations. The presence of SA and a negative r_{MF} is the ancestral state in the LH_M population, demonstrated in the original stock population at UCSB (Chippindale et al. 2001) and then confirmed in LH_M-UU (Innocenti and Morrow 2010). The absence of antagonism ($r_{MF} \approx 0$) thus is a derived state toward which LH_M-UCL evolved. The same directionality is reflected at the genomic level. Our analysis of allele frequencies across the two LH_M populations and the DGRP show that frequency change at candidate loci consists disproportionately of LH_M-UCL frequencies shifting away from those in LH_M-UU. Phenotypic and genetic data thus paint mirror images of rapid change occurring mostly in LH_M-UCL.

Our evidence in favor of a possible recent alleviation of antagonism in LH_M-UCL contrasts with the repeated detections of antagonistic variation in LH_M (Chippindale et al. 2001; Innocenti and Morrow 2010). This suggests that the recent rapid evolution of SA occurred following a long period of stasis, during which antagonistic variation was stably maintained for many generations. It implies that the alleviation of SA was caused by a key innovation decoupling male and female phenotypes, most likely relying on a rare mutational event, such as several epistatic variants arising more or less simultaneously. It seems implausible that the many loci across the genome at which we observe significant population divergence are causally involved in the alleviation of antagonism. Instead, the frequency change we observe at most loci is more likely a consequence, rather than the cause, of alleviation. An emergent mechanism of SA resolution would alter the sex-specific fitness effects of antagonistic alleles previously maintained in balanced polymorphism, leading to shifts in frequencies and potentially the fixation of the allele with the greater average fitness across the sexes (Rice 1984; Connallon and Clark 2012; Mullon et al. 2012).

Analyses to identify the causal genetic changes underlying an alleviation of SA and the specific mechanism(s) of resolution that they exploit are beyond the scope of this study. However, with the proviso that most of our candidate genes are associated with antagonism, our results are informative about the genetic basis of SA. Our data suggests an enrichment of loci on the X chromosome, in line with previous theoretical predictions (Rice 1984). It further suggests that the disproportionate contribution of the X chromosome to quantitative genetic variation in fitness previously observed in LH_M-UCSB (Gibson et al. 2002) is at least in part

due to the number of X-linked antagonistic loci (rather than their phenotypic effects). Our data also provide a first glimpse of genes potentially involved in antagonism. At a general level, the prominent involvement in development of our candidate genes implies that many antagonistic fitness effects may be rooted in anatomical differences between the sexes. This fits with data showing how developmental processes allow the decoupling of male and female phenotypes (Gompel et al. 2005; Roberts et al. 2009; Khila et al. 2012). At the level of individual genes, our candidates include several genes known to be involved in male courtship behavior (*cacophony*, *period* and *Btk family kinase at 29A*) and loci related to sex determination and differentiation (*transformer 2*, *Wnt oncogene analog 2*, *doublesex-Mab related 93B*, *sister-of-Sex-lethal*, and *bric à brac*). While these manual screens are subjective, the presence of such genes raises the intriguing possibility that some of the adaptive trade-offs between males and females are caused by variation close to the top of the sex-specific regulatory cascade.

In conclusion, we have presented data indicating rapid evolutionary change in the genetic basis of fitness and consistent with a partial resolution of antagonism in LH_M-UCL. Due to the enormous scale of the phenotypic assays that we report, our study only compares two populations and relied on combining existing and new data that were collected at different time-points. While the inclusion of phenotypic results from LH_M-UCSC and the DGRP for genetic comparison add some depth and directionality, future work will need to expand our approach in order to establish patterns of SA and its resolution on a broader scale. Phenotypic measures will remain a limiting factor in this effort, due to the sheer amount of work they require and—in particular for wild populations—the uncertainty about how to measure fitness in a way that reflects the selection pressures under which such populations have evolved. However, a growing understanding of the genetic basis of SA will provide an interesting parallel avenue for enquiry, where patterns of diversity at antagonistic loci can be compared across populations to gain insights into the evolutionary dynamics of sexual antagonism. Building upon the results presented here, those efforts will allow us to gain a full understanding of the constraints acting on male and female traits and so elucidate the genomic innovations that allow populations to overcome constraints.

DATA ARCHIVING

LITERATURE CITED

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

Andolfatto, P. 2005. Adaptive evolution of non-coding DNA in *Drosophila*. Nature 437:1149–1152.

Arnqvist, G., and L. Rowe. 2005. Sexual conflict. Princeton Univ. Press, Princeton, NJ.

Barson, N. J., T. Aykanat, K. Hindar, M. Baranski, G. H. Bolstad, P. Fiske, C. Jacq, A. J. Jensen, S. E. Johnston, S. Karlsson, et al. 2015. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. Nature, 528:405–408.

Berg, E. C., and A. A. Maklakov. 2012. Sexes suffer from suboptimal lifespan because of genetic conflict in a seed beetle. Proc. R Soc. B 279:4296–4302.

Betancourt, A. J., J. J. Welch, and B. Charlesworth. 2009. Reduced effectiveness of selection caused by a lack of recombination. Curr. Biol. 19:655–660.

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

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

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

Charlesworth, B., A. J. Betancourt, V. B. Kaiser, and I. Gordo. 2010. Genetic recombination and molecular evolution. Cold Spring Harb. Symp. Quant. Biol. 74:177–186.

Charlesworth, B., J. A. Coyne, and N. H. Barton. 1987. The relative rates of evolution of sex chromosomes and autosomes. Am. Nat. 130:113–146.

Chintapalli, V. R., J. Wang, and J. A. T. Dow. 2007. Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. Nat. Genet. 39:715–720.

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

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

———. 2012. A general population genetic framework for antagonistic selection that accounts for demography and recurrent mutation. Genetics 190:1477–1489.

———. 2014. Evolutionary inevitability of sexual antagonism. Proc. R Soc. B 281:20132123.

Cox, R. M., and R. Calsbeek. 2009. Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. Am. Nat. 173:176–187.

Day, T., and R. Bonduriansky. 2004. Intralocus sexual conflict can drive the evolution of genomic imprinting. Genetics 167:1537–1546.

Dean, R., J. C. Perry, T. Pizzari, J. E. Mank, and S. Wigby. 2012. Experimental evolution of a novel sexually antagonistic allele. PLoS Genet. 8:e1002917.

Debelle, A., M. G. Ritchie, and R. R. Snook. 2014. Evolution of divergent female mating preference in response to experimental sexual selection. Evolution 68:2524–2533.

Delcourt, M., M. W. Blows, and H. D. Rundle. 2009. Sexually antagonistic genetic variance for fitness in an ancestral and a novel environment. Proc. R Soc. B 276:2009–2014.

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

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

Fabian, D. K., M. Kapun, V. Nolte, R. Kofler, P. S. Schmidt, C. Schlotterer, and T. Flatt. 2012. Genome-wide patterns of latitudinal differentiation among populations of *Drosophila melanogaster* from North America. Mol. Ecol. 21:4748–4769.

Fedorka, K. M., and T. A. Mousseau. 2004. Female mating bias results in conflicting sex-specific offspring fitness. Nature 429:65–67.

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

Gavrilets, S. 2000. Rapid evolution of reproductive barriers driven by sexual conflict. Nature 403:886–889.

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

Gibson, J. R., A. K. Chippindale, and W. R. Rice. 2002. The X chromosome is a hot spot for sexually antagonistic fitness variation. Proc. R. Soc. Lond. B 269:499–505.

Gnad, F., and J. Parsch. 2006. Sebida: a database for the functional and evolutionary analysis of genes with sex-biased expression. Bioinformatics 22:2577–2579.

Gompel, N., B. Prud'homme, P. J. Wittkopp, V. A. Kassner, and S. B. Carroll. 2005. Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. Nature 433:481–487.

Haddrill, P. R., D. L. Halligan, D. Tomaras, and B. Charlesworth. 2007. Reduced efficacy of selection in regions of the *Drosophila* genome that lack crossing over. Genome Biol. 8:R18.

Haddrill, P. R., L. Loewe, and B. Charlesworth. 2010. Estimating the parameters of selection on nonsynonymous mutations in *Drosophila pseudoobscura* and *D. miranda*. Genetics 185:1381–1396.

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

Houle, D., and K. Meyer. 2015. Estimating sampling error of evolutionary statistics based on genetic covariance matrices using maximum likelihood. J. Evol. Biol. 28:1542–1549.

Huang, D. W., B. T. Sherman, and R. A. Lempicki. 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37:1–13.

———. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4:44–57.

Huang, W., A. Massouras, Y. Inoue, J. Peiffer, M. Ramia, A. M. Tarone, L. Turlapati, T. Zichner, D. H. Zhu, R. F. Lyman, et al. 2014. Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. Genome Res. 24:1193–1208.

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

Iwasa, Y., and A. Pomiankowski. 1999. Sex specific X chromosome expression caused by genomic imprinting. J. Theor. Biol. 197:487–495.

———. 2001. The evolution of X-linked genomic imprinting. Genetics 158:1801–1809.

Jensen, M. A., B. Charlesworth, and M. Kreitman. 2002. Patterns of genetic variation at a chromosome 4 locus of *Drosophila melanogaster* and *D. simulans*. Genetics 160:493–507.

Kapun, M., H. Van Schalkwyk, B. Mcallister, T. Flatt, and C. Schlotterer. 2014. Inference of chromosomal inversion dynamics from Pool-Seq data in natural and laboratory populations of *Drosophila melanogaster*. Mol. Ecol. 23:1813–1827.

Khila, A., E. Abouheif, and L. Rowe. 2012. Function, developmental genetics, and fitness consequences of a sexually antagonistic trait. Science 336:585–589.

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

Kofler, R., P. Orozco-terWengel, N. De Maio, R. V. Pandey, V. Nolte, A. Futschik, C. Kosiol, and C. Schlotterer. 2011a. PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. PLoS One 6:e15925.

Kofler, R., R. V. Pandey, and C. Schlötterer. 2011b. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27:3435–3436.

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

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

Lewis, Z., N. Wedell, and J. Hunt. 2011. Evidence for strong intralocus sexual conflict in the Indian meal moth, *Plodia interpunctella*. Evolution 65:2085–2097.

Long, T. A. F., A. F. Agrawal, and L. Rowe. 2012. The effect of sexual selection on offspring fitness depends on the nature of genetic variation. Curr. Biol. 22:204–208.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, MA.

Mackay, T. F. C., S. Richards, E. A. Stone, A. Barbadilla, J. F. Ayroles, D. H. Zhu, S. Casillas, Y. Han, M. M. Magwire, J. M. Cridland, et al. 2012. The *Drosophila melanogaster* genetic reference panel. Nature 482:173–178.

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

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

Martin, O. Y., and D. J. Hosken. 2003. The evolution of reproductive isolation through sexual conflict. Nature 423:979–982.

Mcdonald, J. H., and M. Kreitman. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila*. Nature 351:652–654.

McLaren, W., B. Pritchard, D. Rios, Y. A. Chen, P. Flicek, and F. Cunningham. 2010. Deriving the consequences of genomic variants with the Ensembl API and SNP effect predictor. Bioinformatics 26:2069–2070.

Meyer, K. 1991. Estimating variances and covariances for multivariate animal-models by restricted maximum likelihood. Genet. Sel. Evol. 23:67–83.

———. 2007. WOMBAT—a tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). J. Zhejiang Univ. Sci. B 8:815–821.

Meyer, K., and D. Houle. 2013. Sampling based approximation of confidence intervals for functions of genetic covariance matrices. Proc. Assoc. Adv. Anim. Breed 20:523–527.

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

Mullon, C., A. Pomiankowski, and M. Reuter. 2012. The effects of selection and genetic drift on the genomic distribution of sexually antagonistic alleles. Evolution 66:3743–3753.

Parsch, J., and H. Ellegren. 2013. The evolutionary causes and consequences of sex-biased gene expression. Nat. Rev. Genet. 14:83–87.

Pennell, T. M., and E. H. Morrow. 2013. Two sexes, one genome: the evolutionary dynamics of intralocus sexual conflict. Ecol. Evol. 3:1819–1834.

Pischedda, A., and A. K. Chippindale. 2006. Intralocus sexual conflict diminishes the benefits of sexual selection. PLoS Biol. 4:2099–2103.

Pischedda, A., and W. R. Rice. 2012. Partitioning sexual selection into its mating success and fertilization success components. Proc. Natl. Acad. Sci. USA 109:2049–2053.

Punzalan, D., M. Delcourt, and H. D. Rundle. 2014. Comparing the intersex genetic correlation for fitness across novel environments in the fruit fly, *Drosophila serrata*. Heredity 112:143–148.

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

———. 1992. Sexually antagonistic genes: experimental evidence. Science 256:1436–1439.

Rice, W. R., and B. Holland. 1997. The enemies within: intergenomic conflict, interlocus contest evolution (ICE), and the intraspecific Red Queen. Behav. Ecol. Sociobiol. 41:1–10.

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

Roberts, R. B., J. R. Ser, and T. D. Kocher. 2009. Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes. Science 326:998–1001.

Scotti, I., and L. F. Delph. 2006. Selective trade-offs and sex-chromosome evolution in *Silene latifolia*. Evolution 60:1793–1800.

Smit, A. F. A., R. Hubley, and P. Green. 2013–2015. RepeatMasker Open-4.0. http://www.repeatmasker.org.

Smith, N. G. C., and A. Eyre-Walker. 2002. Adaptive protein evolution in *Drosophila*. Nature 415:1022–1024.

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

Van Doorn, G. S. 2009. Intralocus sexual conflict. Ann. NY Acad. Sci. 1168:52–71.

Vicoso, B., and B. Charlesworth. 2009. Effective population size and the faster-X effect: an extended model. Evolution 63:2413–2426.

Vincent, C. M., and N. P. Sharp. 2014. Sexual antagonism for resistance and tolerance to infection in *Drosophila melanogaster*. P R Soc. B 281:20140987.

Wei, Z., W. Wang, P. Z. Hu, G. J. Lyon, and H. Hakonarson. 2011. SNVer: a statistical tool for variant calling in analysis of pooled or individual next-generation sequencing data. Nucleic Acids Res. 39:e132.

Associate Editor: C. Wheat
Handling Editor: R. Shaw

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Table S1. ANOVA tables for analyses of fitness across laboratories.

Table S2. Basic information on sequencing data from LH_M-UU and LH_M-UCL.

Table S3. Chromosomal distribution of SNP loci.

Table S4. Inversion markers in LH_M-UU and LH_M-UCL.

Table S5. Results of a Gene Ontology analysis on genes with candidate SNPs.

Table S6. Results of a clustering analysis of Gene Ontology terms associated with candidate genes.

Figure S1. Male and female adult fitness across genotypes in the LH_M-UU population.

Figure S2. Relationship between fitness measures obtained across laboratories.

Figure S3. Difference in genetic polymorphism between LH_M-UU and LH_M-UCL.

Figure S4. F_{ST} at SNP loci along chromosome arms.

Figure S5. Distribution of F_{ST} among SNP loci.

Figure S6. Distribution of distances between neighbouring candidate SNPs.