**1 A conserved *trans* regulatory loop involving an odorant binding protein controls male
2 mating behavior in flies**

3 Pablo J Delclos^{†,1}, Kiran Adhikari^{†,1}, Oluwatomi Hassan¹, Alexander A Oderhowho¹, Vyshnika
4 Sriskantharajah^{1,2}, Tammie Trinh^{1,3}, Richard P Meisel¹

5 [†]Co-first author/equal contribution

6 Affiliation: ¹Department of Biology & Biochemistry, University of Houston, Houston, TX, USA

7 ²Current affiliation: University of Texas Health Science Center School of Biomedical
8 Informatics, Houston, TX, USA

9 ³Current affiliation: University of Texas School of Dentistry, Houston, TX, USA

10 ABSTRACT

11 A major goal in evolutionary biology is to understand how natural variation is maintained in
 12 sexually selected and sexually dimorphic traits. Hypotheses to explain genetic variation in
 13 sexually selected traits include context-dependent fitness effects, epistatic interactions, and
 14 pleiotropic constraints. The house fly, *Musca domestica*, is a promising system to investigate
 15 how these factors affect polymorphism in sexually selected traits. Two common Y chromosomes
 16 (Y^M and III^M) segregate as stable polymorphisms in natural house fly populations, appear to be
 17 locally adapted to different thermal habitats, and differentially affect male mating success. Here,
 18 we perform a meta-analysis of RNA-seq data which identifies genes encoding odorant binding
 19 proteins (in the *Obp56h* family) as differentially expressed between the heads of males carrying
 20 Y^M and III^M . Differential expression of *Obp56h* has been associated with variation in male
 21 mating behavior in *Drosophila melanogaster*. We find differences in male mating behavior
 22 between house flies carrying the Y chromosomes that are consistent with the relationship
 23 between male mating behavior and expression of *Obp56h* in *D. melanogaster*. We also find that
 24 male mating behaviors in house fly are affected by temperature, and the same temperature
 25 differentials further affect the expression of *Obp56h* genes. However, we show that
 26 temperature-dependent effects cannot explain the maintenance of genetic variation for male
 27 mating behavior in house fly. Using a network analysis and allele-specific expression
 28 measurements, we find evidence that the house fly III^M chromosome is a *trans* regulator of
 29 *Obp56h* gene expression. Moreover, we find that *Obp56h* disproportionately affects the
 30 expression of genes on the *D. melanogaster* chromosome that is homologous to the house fly
 31 III^M chromosome. This provides evidence for a conserved *trans* regulatory loop involving
 32 *Obp56h* expression that affects male mating behavior in flies. The complex regulatory
 33 architecture controlling *Obp56h* expression suggests that variation in male mating behavior
 34 could be maintained by epistasis or pleiotropic constraints.

35 INTRODUCTION

36 Sexual selection occurs when there is heritable variation in reproductive success that arises from
 37 competition for access to mates (Jones 2016). Sexual selection can shape phenotypic variation
 38 within and among species (Seehausen and van Alphen 1999; Greene et al. 2000; Brooks and
 39 Endler 2001; van Doorn and Weissing 2002; Rueffler et al. 2006). If a trait is under strong sexual
 40 selection, one may expect the fixation of alleles that increase reproductive success. In contrast to
 41 this expectation, substantial genetic variation can exist in strongly sexually selected traits (e.g.,
 42 the lek paradox), and explaining this variation has been a long-standing goal in evolutionary
 43 biology (Kirkpatrick and Ryan 1991; Tomkins et al. 2004). Addressing this goal includes
 44 determining the genetic basis of sexually selected traits, as well as the environmental drivers that
 45 affect variation in these traits (Jia et al. 2000; Ingleby et al. 2010).

46 Three potential, and non-exclusive, causes of variation in sexually selected traits are epistasis,
 47 pleiotropic constraints, and context-dependent fitness effects. Epistatic and pleiotropic
 48 constraints are signatures of a complex gene regulatory architecture underlying a trait, and both
 49 are predicted to be important genetic mechanisms in maintaining genetic variance, particularly
 50 for fitness-related traits (Fenster et al. 1997; Merilä and Sheldon 1999; Arnqvist et al. 2014). For
 51 example, although Y chromosome genotype is a major determinant of male fitness in *Drosophila*
 52 *melanogaster*, those fitness effects depend on the genetic background (Chippindale and Rice
 53 2001). Epistatic interactions between Y-linked and X-linked or autosomal alleles can therefore
 54 reduce heritable variation for male fitness. Pleiotropy creates genetic covariation amongst
 55 unrelated traits, which can similarly reduce the response to selection and thereby allow genetic
 56 variation in sexually selected traits to be maintained (Fitzpatrick 2004; Chenoweth and
 57 McGuigan 2010). For instance, *Drosophila* cuticular hydrocarbons (CHCs) are lipid compounds
 58 used for both chemical communication and resistance to various environmental stressors,
 59 including desiccation (Chung and Carroll 2015). Disrupted expression of individual genes
 60 responsible for CHC production in *D. melanogaster* can result in significant alterations to both
 61 mating behaviors and ecologically relevant CHCs (Marcillac et al. 2005; Shorter et al. 2016).
 62 Lastly, context-dependence, in the form of genotype-by-environment interactions, may also
 63 maintain variation when environments fluctuate substantially (Kokko and Heubel 2008).
 64 However, the combined effects of ecological and genetic variation on sexually selected and
 65 sexually dimorphic traits is especially unresolved, with conflicting evidence for the relative
 66 importance of sex-by-genotype-by-environment interactions (Delph et al. 2011; Connallon 2015;
 67 Allen et al. 2017; Lasne et al. 2018; Connallon et al. 2019; Ruzicka et al. 2020).

68 The house fly, *Musca domestica*, is a promising system to investigate the factors that maintain
 69 genetic variation for sexually selected traits. House fly has a polygenic sex determination system,
 70 in which multiple male and female determining loci segregate as polymorphisms in natural

71 populations (Hamm et al. 2015). This polymorphism has remained stable in natural house fly
72 populations since at least the mid-20th century (Kozielska et al. 2008; Meisel et al. 2016). Two
73 common proto-Y chromosomes (III^M and Y^M) are distributed along latitudinal clines across
74 multiple continents, with III^M most common in the south and Y^M most common in the north
75 (Franco et al. 1982; Tomita and Wada 1989; Hamm et al. 2005). This clinal distribution suggests
76 that geographically heterogeneous selection pressures are responsible for maintaining polygenic
77 sex determination (Levene 1953; Hedrick et al. 1976). Consistent with this prediction,
78 seasonality in temperature is the best predictor of the frequencies of the proto-Y chromosomes
79 across populations (Feldmeyer et al. 2008). The Y^M and III^M chromosomes also affect male
80 mating success, with III^M males exhibiting an advantage over Y^M males (Hamm et al. 2009).
81 Together, the geographic distribution of the proto-Y chromosomes and their effects on male
82 mating behavior raise the possibility that polymorphic house fly proto-Y chromosomes affect
83 sexually selected traits in a context-dependent (i.e., environment-specific) manner.

84 Here, we utilized the house fly system to test the potential roles of context-dependence and
85 genetic architecture (epistasis and pleiotropy) in the maintenance of genetic variation of a
86 sexually selected trait. To those ends, we tested for genotype-by-temperature interactions
87 affecting differences in male mating behavior between Y^M and III^M males. We also used
88 RNA-seq data to investigate the genetic basis of differences in mating performance between Y^M
89 and III^M males (Meisel et al. 2015; Son et al. 2019; Adhikari et al. 2021). Our results implicate
90 genes encoding odorant binding proteins (Obps) in an evolutionarily conserved *trans* regulatory
91 loop involving a chromosome that independently became sex-linked in both house fly and
92 *Drosophila*.

93 **METHODS**

94 ***RNA-seq differential gene expression analysis***

95 We analyzed published RNA-seq data from *M. domestica* male heads (NCBI Gene Expression
96 Omnibus accessions GSE67065, GSE126685, and GSE126689, shown in Table S1). The
97 RNA-seq data include nine Y^M and fifteen III^M samples (Meisel et al. 2015; Son et al. 2019). We
98 aligned RNA-Seq reads to house fly transcripts from genome assembly v2.0.2 and annotation
99 release 102 (Scott et al. 2014) using kallisto in single-end read mode (Bray et al. 2016). All
100 RNA-seq reads were single-end, and we set the average fragment length to 300 bp and standard
101 deviation to 30 bp for all samples.

102 We tested for differentially expressed (DE) genes between males with a Y^M chromosome and
103 males with a III^M chromosome using a combination of DESeq2 (Love et al. 2014), sva (Leek et
104 al. 2012), and limma (Ritchie et al. 2015). Read counts per gene that passed an initial threshold
105 filter (0.5 counts per million in at least 4 samples) were normalized by variance stabilizing
106 transformation in DESeq2. To remove batch effects across data sets, we used the sva package to

107 identify and estimate surrogate variables that adjust for latent sources of variation (e.g., batch
108 effects). To identify genes DE between Y^M and III^M males, we used the `lmFit()` function in `limma`
109 to fit a linear model comprised of male type (Y^M vs III^M) and our surrogate variables as fixed
110 effects, and read counts as the response variable. We then computed contrasts between male
111 types and calculated test statistics using the `eBayes()` function. Genes below a false discovery
112 rate (FDR) adjusted p value of 0.05 were categorized as DE (Benjamini and Hochberg 1995).

113 *Weighted gene co-expression network analysis*

114 We used weighted gene co-expression network analysis (WGCNA) to identify modules of house
115 fly genes whose expression correlates with male type (Y^M or III^M) on normalized read count data
116 that were adjusted for batch effects in `sva` (Langfelder and Horvath 2008). For all pairs of genes
117 with variable expression across samples, we calculated Pearson's correlation coefficient across
118 all samples. We created an unsigned correlation matrix and adjusted the soft-threshold value (β)
119 to which among-gene covariances are exponentially raised to ensure a scale-free topology (this
120 resulted in $\beta = 7$), thereby creating a weighted network of gene expression. An unsigned matrix
121 allows us to identify connected genes whose expression is either positively or negatively
122 correlated. Within this topological overlapping network (Li and Horvath 2007), genes were
123 hierarchically clustered, and modules were identified based on the degree of similarity among
124 genes. We used a merging threshold of 0.2, with a minimum module size of 30 and a mean
125 connectivity threshold of greater than or equal to 0.7. We used the default parameters of
126 WGCNA for the rest of the analyses. We then correlated module eigengene values for a given
127 module across samples via Pearson's correlation and identified modules differentially regulated
128 between male types at FDR-adjusted $p < 0.05$.

129 To visualize WGCNA genetic covariance results among modules significantly associated with
130 male type, we exported final co-expression networks to Cytoscape (Shannon et al. 2003). We
131 attached information on $\log_2$ fold-change in expression between III^M and Y^M males, as well as
132 chromosomal location, to the network as metadata so this information could be visualized. To
133 identify genes that may have more central functions within and across our significant modules,
134 we ranked genes in descending order based on intramodular connectivity (k_{Within}), calculated in
135 WGCNA. Hub genes identified by intramodular connectivity are generally functionally
136 important genes within a module (Langfelder et al. 2013).

137 We further analyzed among-gene connections involving a family of odorant-binding protein
138 genes (*Obp56h*). Specifically, to identify genes that may regulate or be regulated by genes within
139 the family, we calculated a "connection score" C_i for every gene i as follows:

140

$$C_i = \sum_{j=1}^n a_{ij} |F_j|,$$

141 where a_{ij} represents the adjacency (the Pearson correlation coefficient raised to the soft-
142 threshold power β) between gene i and *Obp56h* gene j , and F_j represents the $\log_2$ fold-change in
143 expression between III^M and Y^M males for *Obp56h* gene j . This weighted product ensured that
144 connections with *Obp56h* genes that are more differentially expressed between male types were
145 prioritized in calculating a gene's connection score. Genes were then ranked by C_i to identify
146 candidate genes that may be strongly tied to *Obp56h* gene expression. Genes with the 100
147 highest *Obp56h* connection scores were classified as “central genes”. We tested for chromosomal
148 enrichment among these central genes using Fisher's exact tests (comparing the number of
149 central and non-central genes on a focal chromosome with the number of central and non-central
150 genes on all other chromosomes) to determine whether the expression of *Obp56h* genes (which
151 are all located on the *M. domestica* chromosome V) might be involved in *trans* regulation with
152 genes located on the III^M proto-Y chromosome.

153 **Gene ontology enrichment analysis**

154 To identify gene ontology (GO) classes and molecular pathways that are enriched among DE
155 genes, across gene modules identified in WGCNA, or among central genes co-expressed with
156 *Obp56h* genes, we used the BiNGO plug-in within Cytoscape (Maere et al. 2005). We identified
157 *D. melanogaster* orthologs for each house fly gene within a given gene list via NCBI blastx best
158 hits (with default parameters) and used the *D. melanogaster* gene name as input. We identified
159 GO terms that are significantly enriched in BiNGO for biological processes, cellular
160 components, and molecular function.

161 **Allele-specific expression analysis**

162 We tested for differential expression of house fly chromosome III genes between the allele on the
163 III^M chromosome and the allele on the standard third chromosome in III^M males. To do so, we
164 followed methods as in previous studies (Meisel et al. 2017; Son and Meisel 2021), which use
165 the Genome Analysis Toolkit (GATK) best practices workflow for single nucleotide
166 polymorphism (SNP) calling to identify sequence variants in our RNA-Seq data (McKenna et al.
167 2010). We focused our analysis on libraries that were sequenced from head tissue of male house
168 flies that comprise a CS genetic background (Meisel et al. 2015; Son et al. 2019; Adhikari et al.
169 2021). We used STAR (Dobin et al. 2013) to align reads from a total of 30 head libraries (15 III^M
170 and 15 Y^M libraries) to the house fly reference genome (Musca_domestica-2.0.2). We then
171 followed the same methods and applied the same parameters as we have done previously to
172 identify SNPs and genotype individual strains (Meisel et al. 2017; Son and Meisel 2021). We

173 performed separate joint genotyping for each house fly strain within a given experiment (a total
174 of 4 III^M and 4 Y^M strain-by-experimental batch combinations).

175 We use the following approach to differentiate between III^M and standard chromosome III alleles.
176 We first identified SNPs in the exonic regions of the top “hub” genes within a WGCNA module
177 that mapped to house fly chromosome III. We selected SNPs in those genes that are
178 heterozygous in III^M males and homozygous in Y^M males. We used the genotype of these SNPs
179 in Y^M males (which possess two standard third chromosome alleles) to determine the standard
180 chromosome III allele. The allele not found in Y^M genotypes was assigned to the III^M
181 chromosome. We also identified positions where III^M males appear monoallelic for an allele not
182 found in Y^M males. These positions that exhibit a complete bias for a III^M allele are suggestive of
183 monoallelic expression of the III^M allele (i.e., no expression from the III allele).

184 We tested for differences in expression of III^M and standard chromosome III alleles by following
185 best practices for comparing allele-specific expression (Castel et al. 2015). First, for each
186 strain-by-experimental batch combination, we calculated the normalized read depth at each
187 variable site as the number of mapped reads at that site divided by the total number of mapped
188 reads throughout the genome. At each variable site, we used Wilcoxon rank sum tests to make
189 three different pairwise comparisons per site. First, we compared normalized read depths
190 between III^M and III alleles in III^M males (III^M-III). Second, we compared the read depths of the
191 III^M allele in III^M males with the normalized read depth of both III alleles in Y^M males (III^M-Y^M).
192 Third, we compared the read depths of the III allele in III^M males with the normalized read depth
193 of both III alleles in Y^M males (III-Y^M). We set a threshold of significance at $p < 0.05$ for all
194 comparisons.

195 ***Drosophila melanogaster* RNA-seq data analysis**

196 We analyzed RNA-seq results reported in a previous study (Shorter et al. 2016) to determine how
197 knockdown of *Obp56h* affects gene expression in *D. melanogaster*. Shorter et al. (2016)
198 identified DE genes between *Obp56h* knockdown and control samples. This analysis was done
199 separately in males and females, and in separate tissue samples within a given sex (head or the
200 remaining body). We conducted GO enrichment analysis, as described above, on the list of DE
201 genes in *D. melanogaster* male head tissue upon *Obp56h* knockdown.

202 We tested if an excess of DE genes (between *Obp56h* knockdown and controls) are found on the
203 *D. melanogaster* X chromosome, which is homologous to house fly chromosome III (Foster et
204 al. 1981; Weller and Foster 1993). This chromosome is known as Muller element A across flies
205 (Meisel and Scott 2018; Schaeffer 2018). *Obp56h* is located on *D. melanogaster* chromosome
206 2R (Muller element C), which is homologous to house fly chromosome V. We used Fisher’s
207 exact tests (comparing the number of X and non-X chromosome genes that are DE in a given

208 tissue within a given sex with the number of X and non-X chromosome genes that are not DE) to
 209 determine whether *Obp56h* knockdown in *D. melanogaster* results in the disproportionate
 210 differential expression of X chromosome genes in male heads, male bodies, female heads, or
 211 female bodies.

212 We also tested if the same genes are DE between III^M vs Y^M house flies and *Obp56h* knockdown
 213 vs control *D. melanogaster*. Using NCBI blastx best hits, we identified 20 *M. domestica*
 214 transcripts that are orthologous to *D. melanogaster* genes that are DE upon knockdown of
 215 *Obp56h* (11 matches to upregulated *D. melanogaster* genes, and 9 matches to downregulated *D.*
 216 *melanogaster* genes). We compared the mean log₂ fold-changes between Y^M and III^M house fly
 217 males for those 20 genes to 10,000 random subsets of log₂ fold-change values taken from our
 218 data (10,000 subsamples without replacement of 11 genes to test for an excess of positive log₂
 219 fold-change values, and 10,000 subsamples of 9 genes to test for an excess of negative log₂
 220 fold-change values; see Additional Files for R script). We assessed significance by calculating
 221 the proportion of replicated subsamples that generated a mean log₂ fold-change value more
 222 extreme than our observed mean.

223 ***Competitive mating assays***

224 We performed competitive mating experiments in which two different house fly males were
 225 combined with a single female, and we recorded the “winning” male (i.e., the one who
 226 successfully mated with the female), similar to what was done previously (Hamm et al. 2009). In
 227 these experiments, we used the same two house fly strains as in Hamm et al. (2009): a III^M strain
 228 called CS and a Y^M strain called IsoCS. These two strains have a common genetic background
 229 (CS), and only differ in which proto-Y chromosome they carry. Both strains are represented in
 230 the RNA-seq data we analyzed (Meisel et al. 2015; Son et al. 2019), and IsoCS was also
 231 included in a previous RNA-seq study comparing the effects of proto-Y chromosome and
 232 temperature on gene expression (Adhikari et al. 2021). Our experiment differed from previous
 233 work because we reared larvae from each strain at either 18°C and 29°C, whereas Hamm et al.
 234 (2009) worked with flies raised at 28°C. We used the same larval wheat bran diet as done
 235 previously, and we fed adults an *ad libitum* supply of water and an *ad libitum* 1:1 mixture of
 236 dry-milk:sugar. This is also the same diet and rearing protocol used for the flies in the RNA-seq
 237 datasets that we analyzed (Meisel et al. 2015; Son et al. 2019). Male flies were isolated from
 238 females within ~1 hour of eclosion, and each sex was kept separately to ensure that flies had not
 239 mated prior to the experiment.

240 We carried out two distinct competitive mating experiments: 1) inter-strain competition between
 241 males with different genotypes (i.e., Y^M vs III^M) that were reared at the same temperature (363
 242 successful mating trials out of 490 total attempts across 27 experimental batches); and 2)
 243 intra-strain mating between males with the same genotype that were reared at different

temperatures (104 successful mating trials out of 129 total attempts across 7 batches). When we competed flies with different genotypes raised at the same temperature, all males were aged 4-6 d post pupal emergence. When we compared flies with the same genotype raised at different temperatures, 29°C males were aged 4-5 d post emergence and 18°C males were aged 6-7 d post emergence. We aged flies from the colder temperature for more days than flies from the warmer temperature because developmental rate is positively correlated with developmental temperature in flies (Atkinson 1996). The ages we selected ensure that all males were physiologically capable of mating, while also sampling flies at similar physiological ages across experiments. Aging calculations are reported in Supporting Information 1.

The two males in each experiment were labeled using red and blue luminous powder (BioQuip) by shaking the flies in an 8 oz paper cup. The color assigned to males was switched in each successive batch (i.e., blue Y^M and red III^M in one batch, and then red Y^M and blue III^M in the next batch). In addition, we included the genotype or developmental temperature of the blue-colored male as a fixed effect in our statistical analysis (see below), which provides an additional control for color.

For each replicate of the competitive mating assay, we placed the two different males in a 32 oz transparent plastic container, along with a single virgin female. Each plastic container also contained a 1:1 mixture of dry milk:sugar in a 1 oz paper cup and water in a glass scintillation vial plugged with a cotton roll. Virgin females from the LPR strain (Scott et al. 1996) raised at 25°C were used for all combinations of males. The LPR strain has a different genetic background than the males used in the assay, minimizing any effects of co-adaptation between females and a particular subset of males. All flies were transferred into the mating containers using an aspirator and without anesthesia. All matings were performed in a 25°C incubator because copulation latency is too long for experimentally tractable measurement at lower temperatures. The color (i.e., genotype) of the first male to mate was recorded, as well as the time to mate.

We used the `glmer()` function in the `lme4` package in R (Bates et al. 2015) to test for the effects of genotype and temperature on male mating success. First, to test the effect of genotype, we constructed a logistic regression model with developmental temperature, genotype of the blue male, and their interaction as fixed effects. Experimental batch was modeled as a random effect, with the winning male (CS vs IsoCS) as a response variable. We then assessed significance of fixed effects (type II sum of squares) using the `Anova()` function in the `car` package in R (Fox et al. 2013). To test for the effect of temperature on mating success, we similarly constructed a logistic regression model with genotype and the developmental temperature of the blue male as fixed effects and experimental batch as a random effect, with the winning male (18°C vs 29°C) as a response variable.

279 *Single-choice mating assays*

280 We performed experiments to measure copulation latency, or the amount of time elapsed before
 281 successful mating, according to male type (Y^M or III^M). In these experiments, we used the same
 282 IsoCS (Y^M) and CS (III^M) strains as above and in Hamm et al. (2009). We also tested another
 283 strain from each genotype. CSrab (III^M) was created by backcrossing the III^M chromosome from
 284 the rspin strain isolated in New York onto the CS background (Shono and Scott 2003; Son et al.
 285 2019). CSaY (Y^M) was created by backcrossing the Y^M chromosome from the aabys genome
 286 reference strain onto the CS background (Scott et al. 2014; Meisel et al. 2015). Virgin females
 287 used in the assays were all from the LPR strain (Scott et al. 1996), which has a different genetic
 288 background than all males tested. In addition, we also assayed LPR males to determine how
 289 copulation latencies of III^M and Y^M males compare to those of males from the same genetic
 290 background as the females.

291 We first attempted to test flies reared at the same temperatures as in our competitive mating
 292 assays (18°C and 29°C), as well as at an intermediate developmental temperature (22°C).
 293 However, we did not generate enough flies at 18°C, and so we only have data for flies raised at
 294 22°C and 29°C. Our results demonstrate that 22°C is a sufficiently low temperature to detect
 295 effects of both genotype and developmental temperature on mating success (see below). All
 296 larvae from each male strain were reared in 32 oz plastic containers on the same wheat bran diet
 297 described above (Hamm et al. 2009). Upon emergence, unmated male and female progeny were
 298 separated and fed water and a 1:1 mixture of dry milk:sugar *ad libitum* until assays were
 299 conducted. Assays of males raised at 22°C were conducted 10-11 days after eclosion, while those
 300 of males raised at 29°C were conducted 6-7 days after eclosion. This ensures that males were
 301 assayed at similar physiological ages. Females were all raised at 25°C, and unmated females
 302 were aged 8-9 days after eclosion (see Supporting Information 1 for all accumulated degree day
 303 calculations).

304 We followed a similar protocol as in a previous experiment testing copulation latency in
 305 *D. melanogaster* (Shorter et al. 2016). Briefly, five males from a single strain were aspirated
 306 without anesthesia into an 8 oz container covered with a fine mesh cloth secured by rubber band.
 307 Five LPR females were similarly transferred into the container, marking the start of the mating
 308 assay. The house flies were then observed every ten minutes over the course of four hours.
 309 Copulation latency was determined in two ways. First, we measured the amount of time elapsed
 310 between the start of an assay and each successful mating within a container, defined as a male
 311 remaining attached to a female for at least 1 minute (Hamm et al. 2009). Male house flies
 312 typically remain attached to females for >60 minutes (Bryant 1980), making it unlikely, although
 313 possible, for us to miss matings within 10 minute intervals. Individuals who did not mate were
 314 excluded from this analysis. Second, we used a binary variable noting whether each male mated
 315 during the 4 hour assay. Although we were unable to distinguish between individual males in this

assay, we did not observe any males mate more than once within 4 hours in a pilot study conducted between one male and five females, suggesting that observed matings were by different males. All trials were conducted at 22–23°C.

To determine the effects of male type on the amount of time taken to mate, we used the `glmer()` function in the `lme4` package in R (Bates et al. 2015) to create a mixed effects model, including male genotype, developmental temperature, and their interaction as fixed effects and batch and strain as random effects. For the binary measure of copulation latency, we used a binomial logistic regression of the same model, with whether a fly mated as our dependent variable. We then assessed significance of fixed effects (type II sum of squares) using the `Anova()` function in the `car` package in R (Fox et al. 2013). Pairwise comparisons between male types (III^M , Y^M , and LPR) were conducted using Z-tests of proportions.

RESULTS AND DISCUSSION

Differential expression of odorant binding protein genes between III^M and Y^M males

We confirmed that the gene expression profiles of III^M and Y^M male heads are minimally differentiated (Meisel et al. 2015; Son et al. 2019). There are only 40 DE genes between heads of III^M and Y^M adult males (21 upregulated in Y^M males, 19 upregulated in III^M , Table S2). Gene ontology analysis revealed no significant biological process, molecular function, or cellular component terms enriched within the list of DE genes.

Within the list of DE genes, we identified one Obp gene (*LOC105261916*) upregulated in Y^M males. House fly Obp genes can be grouped into families corresponding to their *D. melanogaster* orthologs (Scott et al. 2014). The DE Obp gene in our analysis is orthologous to *Obp56h*. The *Obp56h* family, as well as other Obp families, was greatly expanded within muscids (house fly and close relatives, including stable fly and horn fly) compared to *D. melanogaster* (Scott et al. 2014; Olafson and Saski 2020; Olafson et al. 2021). In addition to *LOC105261916*, seven of the remaining eight house fly *Obp56h* genes for which we obtained RNA-seq count data showed similar trends of greater expression in Y^M than III^M males, with three of these showing significant DE ($p < 0.05$) before an FDR correction (Fig. 1). All but one of the *Obp56h* genes has higher expression in Y^M than III^M males (8/9, regardless of significance), which is significantly greater than the fraction of other genes with higher expression in Y^M males, regardless of significance, in the rest of the genome (Fisher's exact test, $p = 0.019$). Moreover, the expression levels of several house fly *Obp56h* genes are sensitive to developmental temperature or the interaction between temperature and male genotype (Fig S1).

Figure 1 - Neighbor-joining phylogenetic tree of the *Obp56h* gene family within *M. domestica* and *D. melanogaster* based on protein sequences constructed in MEGA X (Kumar et al. 2018). Amino acid sequences were aligned by MUSCLE (Edgar 2004). *M. domestica* *Obp56h* genes are identified based on gene IDs. The bootstrap consensus tree was inferred from 10,000 replicates. Branch lengths are scaled according to the number of amino acid substitutions per site. The phylogeny was arbitrarily rooted at *D. melanogaster Obp56h*. Batch-adjusted expression levels for each *M. domestica Obp56h* gene from each replicate are displayed at the branch tips (small circles). Large circles show the average across all replicates, with error bars denoting the standard error (unfilled stars: $p < 0.05$ before FDR correction for multiple comparisons; filled star: $p < 0.05$ after correction).

III^M confers a mating advantage that is robust to developmental temperature

Knockdown of *Obp56h* in *D. melanogaster* results in decreased male copulation latency, or the time it takes for a male to begin to mate with a female after they are first exposed to one another (Shorter et al. 2016). The *Obp56h* gene family is generally expressed higher in Y^M males relative to III^M males (Fig. 1). A previous study identified a competitive mating advantage of III^M over Y^M male house flies (Hamm et al. 2009), consistent with shorter copulation latency in III^M males because of lower expression of *Obp56h* genes. Segregating variation in such a strongly sexually selected trait likely requires some other factor to maintain the difference in copulation latency between Y^M and III^M males (Kirkpatrick and Ryan 1991; Ingleby et al. 2010). The clinal distribution of Y^M and III^M (Tomita and Wada 1989; Hamm et al. 2005; Kozielska et al. 2008) are suggestive that temperature may differentially affect males carrying these proto-Y chromosomes. In addition, two *Obp56h* genes are only upregulated in Y^M males at 29°C, but not at 18°C (Fig S1), suggesting the effect of Y^M on male mating success may be temperature-dependent. The previous experiment only compared mating performance of Y^M and III^M males at 28°C (Hamm et al. 2009). We therefore tested if the differences in mating success between Y^M and III^M males are sensitive to temperature and are thus context-dependent.

We performed competitive mating assays in which we allowed males carrying III^M or Y^M, reared at either 18°C or 29°C, to compete for a female of an unrelated strain. We found that III^M males were more successful at mating than Y^M males regardless of developmental temperature (ANOVA, $p = 6.53 \times 10^{-6}$, Fig. 2A). This mating advantage of III^M males is consistent with reduced copulation latency as a result of lower expression of *Obp56h* genes. However, our results suggest that there is not an effect of developmental temperature on the III^M male mating advantage.

Figure 2 - III^M chromosome and developmental temperature affect male mating success. A) Outcomes of competitive mating assays between III^M and Y^M males reared at 18°C or 29°C. Data points represent experimental batches. Horizontal lines denote the median across all batches. B) Outcomes of competitive mating assays conducted between males reared at 29 °C and 18 °C. Trials were conducted between males of the same proto-Y chromosome genotype (III^M or Y^M). Each data point represents ten replicate trials within a single batch. C) Outcomes of single-choice mating assays in males reared at 22°C. Data points refer to the percentage of males (five males within one replicate) that mated with females within 4 hours within each experimental trial. Horizontal lines denote means within male groups. All females used were from the LPR strain.

To further investigate if there are effects of developmental temperature on mating success, we tested whether males reared at different developmental temperatures, but with the same genotype, have a difference in mating success. We found that males reared at 18°C outcompeted males reared at 29°C (ANOVA, $p = 2.93 \times 10^{-4}$, Fig. 2B), regardless of genotype. This is consistent with decreased *Obp56h* expression in Y^M males reared at 18°C (Fig S1), which outcompete Y^M males raised at 29°C. In contrast, III^M males exhibited low *Obp56h* expression, regardless of developmental temperature (Fig. S1), yet III^M males raised at 18°C outcompeted III^M males reared at 29°C. Therefore, *Obp56h* expression levels alone cannot explain the effect of developmental temperature on mating success.

We conclude that there is substantial evidence that both proto-Y chromosome genotype and developmental temperature affect male mating success. Notably, the effect of temperature on mating success is consistent across Y^M and III^M male types, suggesting that context-dependence, in the form of a genotype-by-temperature interaction, cannot explain why genetic variation in male mating success is maintained in house fly. However, our mating assays were all performed at the same temperature (25°C), with male house flies that developed at two different temperatures. We therefore cannot rule out the possibility that differences in mating success are affected by courtship temperature (as opposed to developmental temperature).

III^M males have reduced copulation latency

Our competitive mating assays (Fig. 2A), as well as previously published results (Hamm et al. 2009), raise the possibility that III^M males have a reduced copulation latency relative to Y^M males, which could be explained by the lower expression of *Obp56h* genes in III^M males (Fig. 1). We therefore directly measured copulation latency in single-choice mating assays. To do so, we combined five males from a single strain raised at a single temperature with five females from the unrelated strain used in our competitive mating assays. Developmental temperature had a significant effect on copulation latency (ANOVA, $p = 9.40 \times 10^{-5}$), with males reared at 22°C mating faster than those reared at 29°C (Fig. S2A). In general, successful matings were rare for males that developed at 29°C regardless of whether they carry the Y^M or III^M chromosome (15 successful matings out of 75 males tested). These results are consistent with increased mating success of males raised at 18°C relative to those raised at 29°C in our competitive mating assays (Fig. 2B).

We found no significant effect of male genotype (ANOVA, $p = 0.89$), or the interaction between male genotype and developmental temperature (ANOVA, $p = 0.37$), on the time it takes males to mate. When we consider only males who mated within 4 hours, the time to mate for Y^M and III^M males did not significantly differ at either 22°C (Tukey's post-hoc, $p = 0.99$) or 29°C (Tukey's post-hoc, $p = 0.78$). However, this analysis is problematic because we have no measure of copulation latency for males that did not mate within the 4 hour experimental window, which amounts to >70% of males in three of the four genotype-by-temperature combinations. It is therefore possible that copulation latency does indeed differ between Y^M and III^M males in a way that is not detected in our censored data.

To overcome the problems associated with censored data, we next treated copulation latency as a binary variable by calculating the proportion of the five males per trial that mated within the 4 hour assay. We observed significant effects of male genotype (ANOVA, $p = 6.18 \times 10^{-3}$) and developmental temperature (ANOVA, $p = 9.04 \times 10^{-4}$) on the proportion of males that mated. The effect of developmental temperature was largely a result of very few matings for males that developed at 29°C relative to 22°C (Fig. S2B). The lower copulation latency for flies that

developed at 22°C is consistent with our competitive mating assays that showed males that developed at a lower temperature have higher mating success (Fig. 2B). In the 22°C treatment, a significantly greater proportion of III^M males mated within 4 hours than Y^M (61.7% v. 28.3%; Z-test of proportions, $p = 1.21 \times 10^{-4}$; Fig. 2C). This is evidence that III^M males have reduced copulation latency, which is consistent with their previously documented competitive mating advantage (Hamm et al. 2009), the competitive mating advantage that we observe (Fig. 2A), and the reduced expression of *Obp56h* genes (Fig. 1).

The Y^M and III^M males we and others used in mating experiments all share the CS genetic background that comes from a III^M strain (Hamm et al. 2009). This raises the possibility that III^M males perform better because they have a proto-Y chromosome that is co-adapted to its genetic background. To test this hypothesis, we measured copulation latency in Y^M males from the same strain (LPR) as the females in our experiments. We observed a greater proportion of III^M males mating within 4 hours when compared to the LPR Y^M males (Z-test of proportions, $p = 0.038$), although the copulation latency in LPR males was highly variable (Fig. 2C). Therefore, the reduced copulation latency conferred by the III^M chromosome overwhelms any potential effects of coadaptation of the proto-Y chromosome to male genetic background or male-female co-adaptation within strains. The reduced copulation latency of III^M males is only detectable when house flies develop at 22°C, suggesting that it is either temperature-dependent or we lack the resolution to detect it when males develop at warmer temperatures (because they take too long to mate). Our results also provide evidence that the effect of temperature on copulation latency is independent of genotype, suggesting that context-dependent effects of the proto-Y chromosomes cannot explain the maintenance of the polymorphism.

House fly chromosome III genes and Drosophila X chromosome genes have correlated expression with Obp56h genes

In order to identify the regulatory architecture underlying the differential expression of *Obp56h* genes between Y^M and III^M males, we identified 27 co-expression modules across house fly male heads. One of these modules (containing 122 genes, Table S3) is differentially expressed between Y^M and III^M males (FDR adjusted $p = 0.001$, Fig. 3). GO analysis revealed significant enrichment (FDR adjusted $p < 0.05$) of 15 biological process terms including those related to immune system processes (GO:0032501), responses to stress (GO:0006950), and response to external stimuli (GO:0009605) within this module (Table S4). This module is also enriched for house fly chromosome III genes (31 chromosome III genes versus 38 genes assigned to other chromosomes, Fisher's exact test $p < 1 \times 10^{-5}$, with 53 genes not assigned to a chromosome) and for DE genes between Y^M and III^M males (16 DE genes in this module versus 24 DE genes assigned to other modules, Fisher's exact test $p < 1 \times 10^{-5}$). We used the WGCNA measure of intramodular connectivity, kWithin, to identify hub genes within the module that likely have important roles in the regulation of gene expression. The top five hub genes are (with

471 *D. melanogaster* orthologs in parentheses): *LOC101887703* (CG8745), *LOC105262120*
 472 (CG10514), *LOC101894501* (*Gr98c*), *LOC101893264* (*gd*), and *LOC101893651* (CG2120) (Fig.
 473 3, S3).

474 **Figure 3** - Network visualization of the co-expression module that is differentially regulated between III^M
 475 and Y^M males. Each circle within the module is a gene, and *Obp56h* genes are indicated with purple fill.
 476 Lines represent edge connections between genes. Genes labeled with “+” are within the top 100 most
 477 strongly connected to *Obp56h* genes. Genes are ordered from top to bottom according to intramodular
 478 connectivity (kWithin), with genes of higher connectivity (i.e., hub genes) on top, and peripheral genes on
 479 the bottom. Borders around genes denote log₂ fold-change in expression between Y^M and III^M male
 480 heads, with darker blue borders denoting upregulation in Y^M, and darker red borders denoting
 481 upregulation in III^M. Chromosomal locations in house fly (*Mdom*) and *D. melanogaster* (*Dmel*) are shown
 482 for the 5 hub genes and *Obp56h*.

483 Three *Obp56h* genes that are DE between III^M and Y^M males (*LOC105261916*, *LOC101891822*,
 484 and *LOC101891651*) are all assigned to the co-expression module (Fig. 3). The *Obp56h* gene
 485 cluster itself is found on house fly chromosome V, which is unlikely to differ between the Y^M and
 486 III^M males in our experiments—the majority of males compared in the RNA-seq data and mating
 487 experiments have a common genetic background (including chromosome V) and differ only in
 488 whether they carry III^M or Y^M. Removing samples with a different background did not affect the
 489 general difference in *Obp56h* expression between III^M and Y^M males (see Supporting Information
 490 2 for a summary of these results). We therefore hypothesized that differential *Obp56h* expression
 491 is at least partly controlled by *trans* regulatory variation that maps to Y^M, III^M, or both.

492 If *Obp56h* gene expression is regulated by *trans* factors that map to chromosome III, then we
 493 expect the *Obp56h* gene family to be found within, or strongly connected to genes within, the
 494 gene module that is differentially regulated between III^M and Y^M males. Indeed, we find that the
 495 module is enriched for *Obp56h* genes relative to other Obp genes—three *Obp56h* genes and no
 496 other Obp genes were assigned to this module (Fisher’s exact test, $p = 5.1 \times 10^{-3}$, Fig. 3). This
 497 suggests that *Obp56h* expression is strongly correlated with, and is either regulated by or
 498 regulates, genes whose expression is affected by the III^M chromosome. We cannot perform the
 499 same analysis for the effect of Y^M because only 51 genes have been assigned to the house fly
 500 X/Y^M chromosome (Meisel and Scott 2018), limiting our power to detect an excess of genes.

501 To test for *trans* regulators of *Obp56h* gene expression, we identified house fly genes whose
 502 expression covaries with *Obp56h* genes (Table S5). We found significant enrichment for
 503 chromosome III genes within the 100 genes whose expression covaries most with *Obp56h* gene
 504 expression (corresponding to the top 0.55% covarying genes); of the 100 genes with the highest
 505 *Obp56h* connection scores, 26 are on chromosome III (Fisher’s exact test $p = 2.0 \times 10^{-4}$, Fig. 4A).
 506 This enrichment is robust to varying the threshold used to classify a gene as in the top covarying;
 507 considering genes with the top 1%, 5%, 10% covarying expression also resulted in significant
 508 enrichment of chromosome III genes (Fisher’s exact test, all $p < 0.05$). This supports the
 509 hypothesis that *trans* regulatory variants that differ between III^M and the standard chromosome
 510 III are at least partially responsible for DE *Obp56h* genes between III^M and Y^M house fly males.

511 Our network analysis does not ascribe directions to the edges connecting house fly genes, and it
 512 is therefore possible that *Obp56h* DE has *trans* regulatory effects on chromosome III expression.
 513 To test this hypothesis, we examined available RNA-seq data from an experiment comparing
 514 wild type *D. melanogaster* with flies in which *Obp56h* had been knocked down (Shorter et al.
 515 2016). *Obp56h* is on the right arm of the second chromosome in *D. melanogaster* (2R, or Muller
 516 element C), which is homologous to house fly chromosome V (Foster et al. 1981; Weller and
 517 Foster 1993). House fly chromosome III is homologous to the *D. melanogaster* X chromosome,
 518 which is known as Muller element A (Meisel and Scott 2018; Schaeffer 2018). The
 519 *D. melanogaster* males in the RNA-seq experiment all share the same X chromosome, and only
 520 differ in one copy of their second chromosome (which either carries a UAS-RNAi knockdown
 521 construct or does not). If *Obp56h* genes have *trans* regulatory effects on element A genes in
 522 males, we would expect an excess of DE *D. melanogaster* X chromosome genes *Obp56h*
 523 knockdown flies. Indeed, we found that *Obp56h* knockdown in *D. melanogaster* resulted in
 524 excess DE of X chromosome genes in male head (Fisher’s exact test, $p = 0.011$, Fig. 4B) and
 525 body ($p = 0.038$, Fig. S4), but not in either tissue sample in females (Fisher’s exact test, both $p >$
 526 0.49). These results suggest that there is male-specific *trans* regulatory control of *D.*
 527 *melanogaster* X-linked genes by *Obp56h*. This regulatory architecture associated with *Obp56h*
 528 expression may create epistatic or pleiotropic constraints that inhibit selection to reduce

expression of *Obp56h* genes, which should otherwise be favored because reduced expression shortens copulation latency.

Figure 4 - Percent of genes on each chromosome within (A) the top 100 genes with the strongest connections to the *Obp56h* family in house fly, and (B) genes differentially expressed (DE) between *Obp56h* knockdown and control *D. melanogaster* (black bars: males, grey bars: females). Asterisks indicate a significant difference between observed (bars) and expected (red lines) counts of genes on each chromosome compared to all other chromosomes (Fisher's exact test, $p < 0.05$).

Our house fly results suggest that alleles differing between III^M and the standard chromosome III have *trans* effects on *Obp56h* expression. The *D. melanogaster* data suggest that variation in *Obp56h* expression affects the expression of genes on the *D. melanogaster* X chromosome (which is homologous to house fly chromosome III). This raises the possibility that there is *trans* regulatory feedback from *Obp56h* genes that affects chromosome III gene expression in house fly. If that regulatory feedback is conserved between house fly and *D. melanogaster*, we expect that orthologous genes would be DE on Muller element A (house fly chromosome III and the *D. melanogaster* X chromosome) between III^M and Y^M house flies and between *Obp56h* knockdown and wild-type *D. melanogaster*. Consistent with this expectation, we found that genes that are downregulated upon knockdown of *Obp56h* in *D. melanogaster* have house fly orthologs that are more downregulated in III^M male house flies (i.e., lower log₂ fold-change) than expected by chance ($p = 5.60 \times 10^{-3}$, Fig. S5A). In contrast, genes that were upregulated upon *Obp56h* knockdown in *D. melanogaster* were not significantly differentially regulated between Y^M and III^M male genotypes, although the observed trend suggests that these genes may be more downregulated in III^M males than expected ($p = 0.103$, Fig. S5B). Our results therefore suggest that, genome-wide, the genes which vary in expression along with *Obp56h* (regardless of the direction of *trans* regulation) are evolutionarily conserved between *M. domestica* and *D. melanogaster*. Consistent with this hypothesis, we identified the GO term “response to stress” (GO:0033554) to be significantly enriched amongst genes with strong connection scores with *Obp56h* expression in *M. domestica* and in the list of DE genes in *D. melanogaster* upon *Obp56h*

knockdown (Table S6). This provides further evidence that there is an evolutionarily conserved *trans* regulatory feedback loop involving *Obp56h* expression in *Drosophila* and house fly through similar molecular functions.

Network analysis reveals candidate regulators of *Obp56h* expression

The house fly co-expression module contains candidate genes and pathways through which *Obp56h* genes, and likely male copulation latency, are regulated. For example, within the list of DE genes, we identified one gustatory receptor gene (*LOC101894501*, the ortholog of *D. melanogaster Gr98c*) upregulated in III^M males ($p_{\text{ADJ}} = 0.037$). Although Obps can have a variety of functions, they most typically interact with chemosensory receptors (odorant, ionotropic, and gustatory receptors) in the detection of chemical cues or signals (Zhou 2010; Benoit et al. 2017; Sun et al. 2018). If *Obp56h* serves a sensory detection role in male house fly heads, then *Gr98c* is a promising candidate gene with which it interacts. The *M. domestica* ortholog of *Gr98c* (*LOC101894501*, which we will refer to as *Md-Gr98c*) is a hub gene in the co-expression module containing *Obp56h* (Fig. 3). *Md-Gr98c* is the only chemosensory receptor assigned to this module, it is located on chromosome III, and its expression is negatively correlated with *Obp56h*. A negative correlation between the expression of a chemosensory receptor and its interacting binding protein has previously been reported in a pair of genes that modulate male *Drosophila* mating behavior (Park et al. 2006).

Figure 5 - Allele-specific expression (ASE) in A) *LOC101894501* (*Md-Gr98c*), B) *LOC101893264* (*Md-gd*), and C) *LOC101893651* (*CG2120* ortholog). The x-axis depicts base pair positions (scaffold coordinates) of the informative single nucleotide polymorphisms (SNPs) that differ between III^M and standard chromosome III alleles. The y-axis and data points depict the read depth of a given allele normalized by the total mapped reads for a given strain-by-experimental batch group combination (FPM = fragments per million). Lines depict mean read depths at each diagnostic site for III (turquoise) and III^M

(salmon) alleles in III^M males, and mean read depths at each site for III alleles in Y^M males (black). Tables under each graph mark significant differences (*: $p < 0.05$) in normalized read depths at each diagnostic site for each of three pairwise comparisons: III^M allele vs. III allele in III^M males (III^M-III), III^M allele in III^M males vs both III alleles in Y^M males (III^M-Y^M), III allele in III^M males vs. both III alleles in Y^M males (III-Y^M).

We tested if *Md-Gr98c* is differentially regulated between the III^M chromosome and standard chromosome III by comparing expression in III^M males (i.e., heterozygotes for III^M and a standard chromosome III) with males homozygous for the standard chromosome III. Differential expression of the III^M and III chromosome alleles would implicate this gene as having a causal effect on *Obp56h* expression. *Md-Gr98c* contains 4 exonic SNPs differentiating the III^M and III chromosomes. Within each III^M strain in each RNA-seq experiment, we observed significantly greater expression of the III^M allele than the standard chromosome III allele at two of the four diagnostic SNP sites (Fig. 5A). The other two SNPs showed the same pattern of III^M-biased expression but were not significant (both $p > 0.05$). The III^M allele is also expressed higher than both III alleles in Y^M males. Higher expression of the III^M allele is consistent with *cis* regulatory divergence between the III^M and standard chromosome III being partially responsible for elevated *Md-Gr98c* expression in III^M males. Furthermore, the standard chromosome III allele is expressed significantly higher in III^M males than Y^M males at two of the four diagnostic SNP sites (Fig. 5A); we observe the same pattern at the other two sites without significance ($p > 0.05$). Higher expression of the III allele in III^M males than Y^M males suggests that *trans* regulators further increase the expression of *Md-Gr98c* in III^M males. This combination of *cis* and *trans* regulatory affects on *Md-Gr98c* expression are consistent with the *trans*-regulatory loop we hypothesize between *Obp56h* and chromosome III that regulates male mating behavior. Future experiments could determine whether *Gr98c* and *Obp56h* do indeed interact and, if so, what pheromonal or other chemical compounds they detect.

A serine protease gene, *LOC101893264*, orthologous to *D. melanogaster gd* (Konrad et al. 1998), is also among the top 5 hub genes within the co-expression module (Fig. 3). This gene is predicted to encode a positive regulator of the Toll signaling pathway (Valanne et al. 2011), suggesting that the *M. domestica* ortholog of *gd* (*Md-gd*) could have an important gene regulatory function within the module via Toll signaling. *Md-gd* is located on chromosome III, and it is upregulated in III^M males (adj. $p = 0.022$). We identified seven sites where all RNA-seq reads were mapped to the III^M allele, while no reads were mapped to the standard chromosome III allele in III^M males (Fig. 5B). At all seven diagnostic SNP sites in this gene, the III^M allele is significantly more highly expressed than the III allele in III^M males (all $p = 0.021$), and it is more highly expressed than both III alleles in Y^M males at six of seven sites (all $p = 0.021$). The lack of expression of the III allele in either III^M or Y^M males is consistent with monoallelic gene

616 expression of the III^M allele, although further evidence is required to confirm this hypothesis (see
617 Supporting Information 3 for detailed discussion).

618 We identified similar evidence of monoallelic gene expression within another hub gene,
619 *LOC101893651*, which is orthologous to *D. melanogaster* *CG2120* (Fig. 5C). *LOC101893651* is
620 among the most central genes within the co-expression module (Fig. 3), and it is strongly
621 upregulated in III^M males ($\log_2$ fold-change: 1.33, adjusted $p = 0.033$). *LOC101893651* is found
622 on house fly chromosome III and is predicted to encode a transcription factor. At all four
623 diagnostic sites within *LOC101893651*, the III^M allele is significantly more highly expressed
624 than the III allele in III^M males (all $p = 0.021$), as well as both III alleles in Y^M males (all $p \leq$
625 0.027). Within the WGCNA module, *Obp56h* expression is most strongly correlated with
626 *LOC101893651*, suggesting that *LOC101893651* could encode the transcription factor that is
627 directly responsible for the repression of *Obp56h* expression in III^M males.

628 Based on their differential expression, allele-specific expression, centrality within the
629 co-expression module, and their location on chromosome III, *Md-Gr98c*, *Md-gd*, and
630 *LOC101893651* are all strong candidate genes that may directly or indirectly affect the
631 expression of house fly *Obp56h* genes. However, the expression of genes on other chromosomes
632 are also strongly correlated with *Obp56h* expression (Fig. 3), suggesting other chromosomes may
633 also be involved in the *trans* regulatory loop. In addition, many house fly genes are not yet
634 mapped to chromosomes (Meisel and Scott 2018), limiting our ability to infer the chromosomes
635 involved in regulating *Obp56h* expression. We describe two genes (*LOC105262120* and
636 *LOC101887703*) that may be located on other chromosomes below.

637 Expression of *LOC101893651* (the ortholog of *CG2120* that is predicted to encode a
638 transcription factor) is most strongly correlated with *LOC105262120* (an ortholog of *CG10514*),
639 which is predicted to encode an ecdysteroid kinase. Expression of *Obp56h* and *CG10514* are
640 correlated with the production of CHCs that are involved in social behavior in *D. melanogaster*
641 (Shorter et al. 2016) and *Drosophila serrata* (McGraw et al. 2011), respectively. CHCs are often
642 under strong sexual selection across insect systems, with individual or combinations of CHCs
643 serving as important mating cues (Thomas and Simmons 2009, Berson et al. 2019a,b). The
644 correlation of these genes with CHC profiles in *Drosophila* provides additional evidence that
645 *Obp56h* expression, and the house fly co-expression module more generally, are related to male
646 mating behavior, and possibly under sexual selection. However, cuticular hydrocarbons also
647 provide protection against biotic and abiotic stressors (Otte et al. 2018) and play a crucial role in
648 desiccation resistance (Lockey 1988). This dual role of CHCs in mating and stress resistance
649 suggests that sexual selection on *Obp56h* expression could be pleiotropically constrained by
650 trade-offs with stress response. Our GO enrichment analysis on both the house fly and *D.*
651 *melanogaster* RNA-seq data also revealed that *Obp56h* expression is correlated with the
652 expression of genes involved in general stress responses, supporting this hypothesis. Pleiotropic

constraints on *Obp56h* expression (because of correlated changes in CHCs) could therefore reduce the response to selection on male copulation latency, contributing to the maintenance of genetic variance. Future studies should aim at determining whether III^M and Y^M male house flies also differ in CHC profiles, as well as their resistance to desiccation and other environmental stressors.

Figure 6 - Correlations of gene expression between *Obp56h* (house fly *LOC105261916*) and *CG8745* (*LOC101887703*) in (A) house fly male head tissue, and (B) *D. melanogaster* male head tissue. Values for *D. melanogaster* are from count data as reported in Shorter et al. (2016). Linear regression models were used to determine 95% confidence intervals (shaded in grey) summarizing the effect of *Obp56h* expression on *CG8745* expression in each species.

LOC105262120 expression is most strongly correlated with the most central gene within the co-expression module, *LOC101887703*. *LOC105262120* and *LOC101887703* are both upregulated in III^M males (*LOC101887703*: log₂ fold-change: 2.21, adj. $p = 0.016$; *LOC105262120*: log₂ fold-change: 2.26, adj. $p = 0.007$). *LOC101887703* is orthologous to *D. melanogaster* *CG8745*, which is predicted to encode an ethanolamine-phosphate phospho-lyase and is broadly expressed in many *D. melanogaster* tissues (Chintapalli et al. 2007). In both *D. melanogaster* and house fly, *Obp56h* expression is significantly negatively correlated with the expression of *CG8745* or *LOC101887703*, respectively (Fig. 6).

LOC101887703 has a paralog (*LOC101890114*) that is predicted to be on chromosome III. The two transcripts encoded by these paralogs are <1% diverged in their nucleotide sequences, suggesting a recent duplication event. Gene duplication is hypothesized to be involved in the resolution of inter-sexual conflict (Connallon and Clark 2011; Gallach and Betrán 2011; Van Kuren and Long 2018), raising the possibility that at least one of the two house fly paralogs is partitioned into a sex-specific function. Moreover, broadly expressed genes often give rise to paralogs with sex-specific expression (Meisel et al. 2009). Notably, chromosome III is a proto-X chromosome (III^M is the proto-Y), and there is a general excess of gene duplication from X chromosomes to the autosomes across flies and other animals, possibly driven by selection on sex-specific functions (Betrán et al. 2002; Emerson et al. 2004; Meisel et al. 2009; Baker and

Wilkinson 2010). Future work could address a potential sexually dimorphic subfunctionalization of the two *CG8745* paralogs in the house fly genome.

Together, the hub genes in the co-expression module are likely to be members of an evolutionarily conserved *trans* regulatory feedback loop that controls and/or is controlled by *Obp56h* expression (Fig. 7A). Specifically, we infer that *Obp56h* regulates *CG8745* (*LOC101887703*) because knockdown of *Obp56h* causes an increase in *CG8745* expression in *D. melanogaster* (Fig. 6B). Based on their locations on house fly chromosome III, positions as hub genes in the house fly co-expression module (Fig. 3), and divergent expression between the III^M and standard III chromosomes (Fig. 5), we also hypothesize that *LOC101893651* (*CG2120*), *Md-gd*, and/or *Md-Gr98c* regulate *Obp56h* in house fly (Fig. 7B). *Md-Gr98c* is a particularly promising candidate, because chemosensory binding proteins and receptors are known to co-regulate one another (Park et al. 2006). Future manipulative experiments will help in further evaluating the direction of regulation of these co-expressed genes.

Figure 7 - Hypotheses on connections between *Obp56h* expression, proto-Y chromosome genotype, and male mating behavior based on house fly and *D. melanogaster* gene expression data. A) Summary of evidence for an evolutionarily conserved *trans* regulatory loop between *Obp56h* and Muller Element A (house fly chromosome III, and *D. melanogaster* X chromosome). Our hypothesis is based on differential expression between III^M vs. Y^M male house flies, *Obp56h* knockdown vs. control *D. melanogaster*, and network connectivity of *Obp56h* family gene expression within house fly. B) Summary of candidate genes implicated in conserved *trans* regulatory loop. Three of the top five hub genes of module A are located on house fly chromosome III, are negatively correlated with *Obp56h* expression, and exhibit either allele-specific expression (ASE) or show signs of monoallelic gene expression biased towards the III^M allele. Similar correlations between expression measures of *Obp56h* and *CG8745* (*LOC101887703*) in *D. melanogaster* and house fly male head tissue suggest that *Obp56h* regulates *CG8745*, which is the primary hub gene in the WGCNA module that is differentially expressed between III^M and Y^M male house

flies. Shared correlations between *Obp56h* expression and copulation latency in both house fly and *D. melanogaster* also suggest that *Obp56h* regulates male fly mating behavior.

CONCLUSIONS

In this study, we aimed to identify a genetic mechanism explaining phenotypic variation in male mating performance across male proto-Y chromosome genotypes in house fly (Hamm et al. 2009). An *Obp56h* gene is among our small list of genes differentially expressed between III^M and Y^M male heads (Fig. 1). In *D. melanogaster*, *Obp56h* expression affects male copulation latency (Shorter et al. 2016), and we identified a similar difference in copulation latency between male proto-Y genotypes in house fly (Fig. 2). We also identified an excess of house fly chromosome III (Muller element A) genes strongly co-expressed with the *Obp56h* gene family, suggesting *trans* regulation of *Obp56h* by genes from element A (Fig. 3, 4, 5). Similarly, we observe evidence for excess *trans* regulation of element A genes in *D. melanogaster* by *Obp56h* (Fig. 4). The inferred directions of *trans* regulation between *Obp56h* and element A are in opposite directions in house fly and *D. melanogaster* (Fig. 7A), but an excess of the same element A genes are correlated with *Obp56h* expression in both species. This suggests that there is an evolutionarily conserved *trans* regulatory loop affecting and affected by *Obp56h* expression (Fig. 7) between fly species that diverged >50M years ago (Wiegmann et al. 2011). The shared relationship between *Obp56h* expression and copulation latency in both species suggests that this *trans* regulatory loop has conserved effects on male mating behavior across distantly related flies.

House flies and *Drosophila* have independently acquired the same sex chromosome (Muller element A is house fly chromosome III, which is homologous to the *Drosophila* X). This raises the possibility that element A is primed to be recruited as a sex chromosome because of the *trans* regulatory connections with a gene (*Obp56h*) that has important effects on male mating behavior. Convergent sex-linkage of the same chromosomal region has been observed in vertebrates (O'Meally et al. 2012; Furman and Evans 2016; Ezaz et al. 2017), which could be explained by the same gene independently acquiring a sex determining allele in multiple independent lineages (Takehana et al. 2014). Genes with sex-specific effects (including sexually antagonistic variants) are also expected to be an important selective force in the formation of new sex chromosomes (van Doorn and Kirkpatrick 2007) and the subsequent evolution of X and Y chromosomes (Gibson et al. 2002; Charlesworth et al. 2005; Abbott et al. 2017). Our results suggest that an enrichment of genes that regulate sexually selected behaviors could promote the sex-linkage of the same chromosome in distantly related species without convergent evolution of a master sex determiner.

740 The apparently simple correlation between *Obp56h* expression and male mating behavior
 741 suggests that selection for reduced male copulation latency in house fly could easily be achieved
 742 by downregulation of *Obp56h* expression. Our results address three possible mechanisms for the
 743 paradoxical maintenance of genetic variation underlying this sexually selected trait:
 744 context-dependence, epistasis, and pleiotropy. To address context-dependence, we tested if the
 745 effects of Y^M and III^M on copulation latency vary across temperatures. Y^M and III^M are
 746 distributed across a latitudinal cline, suggesting that temperature differentially affects the fitness
 747 of males according to proto-Y chromosome genotype (Tomita and Wada 1989; Hamm et al.
 748 2005; Feldmeyer et al. 2008; Kozielska et al. 2008). We found that temperature does indeed
 749 affect male mating performance, but III^M males outperform Y^M males regardless of temperature
 750 (Fig. 2). In addition, the *Obp56h* gene with the strongest signal of differential expression
 751 (*LOC105261916*) is only differentially expressed in Y^M males across temperatures (Fig S1),
 752 suggesting that *Obp56h* expression levels alone cannot explain the effect of temperature on
 753 mating success. Therefore, we hypothesize that proto-Y chromosome genotype and
 754 developmental temperature have independent effects on male mating success and copulation
 755 latency. The lack of a genotype-by-environment interaction suggests that context-dependent
 756 effects cannot explain the maintenance of variation in male copulation latency in house flies.

757 The interaction between proto-Y chromosome genotype and *Obp56h* expression is suggestive
 758 that epistasis may reduce the response to selection on copulation latency in house fly. Similar to
 759 our results, Y chromosome genotype in *D. melanogaster* is a major determinant of male fitness,
 760 but this effect depends on genetic background (Chippindale and Rice 2001). This suggests that
 761 epistatic interactions reduce the efficacy of selection on male-beneficial Y-linked alleles in
 762 *D. melanogaster*. Our experiments were not explicitly designed to test for the effect of genetic
 763 background. Nonetheless, we found that LPR males exhibit similar delayed copulation latencies
 764 as the other Y^M males we tested (Fig. 2C), despite having a different genetic background. This
 765 suggests that epistatic interactions do not necessarily modulate the effect of Y^M and/or III^M on
 766 male mating. However, we are unable to tease apart the effects of the Y^M chromosome from
 767 those of the remaining genetic background, as LPR males also possess a different Y^M
 768 chromosome than the other Y^M males we tested. Regardless of this limitation, our results are
 769 consistent with the hypothesis that the effect of *Obp56h* expression on male copulation latency
 770 depends on male proto-Y chromosome genotype via *trans* regulation, which leaves open the
 771 possibility that epistasis constrains the response to selection for this trait.

772 Our results provide stronger evidence that pleiotropic constraints maintain genetic variation for
 773 male mating behavior in house fly. The *trans* regulatory loop between *Obp56h* expression and
 774 multiple genes on Muller element A may create pleiotropic constraints that weaken the efficacy
 775 of selection on this trait. For example, *trans* regulators are predicted to have pleiotropic effects,
 776 which could impede the response to selection on traits they affect (Carroll 2005). In addition, the
 777 genes implicated in the *Obp56h trans* regulatory loop may affect other phenotypes, which can

778 create correlations between traits and weaken the response to selection (Lande and Arnold 1983).
 779 Notably, the expression of both *Obp56h* and *CG10514* are associated with CHC profiles in
 780 *Drosophila* (McGraw et al. 2011; Shorter et al. 2016). CHCs serve dual functions for both
 781 chemical communication and protection against environmental stressors (Blomquist and
 782 Bagnères 2010). Selection on *Obp56h* expression and male mating behavior could thus be
 783 weakened by trade-offs between mating behavior and stress response. This is consistent with the
 784 general prediction that pleiotropy can maintain genetic variance for sexually selected traits in
 785 natural populations (Kirkpatrick and Ryan 1991; Turelli and Barton 2004; Johnston et al. 2013;
 786 Heinen-Kay et al. 2020). Our results therefore provide a potential example of how the regulatory
 787 architecture underlying a sexually selected trait can create pleiotropic constraints that could
 788 impede selection on the trait, maintaining genetic variation in spite of strong selection on the
 789 trait.

790 ACKNOWLEDGEMENTS

791 Rebecca Presley assisted with house fly mating experiments. This material is based upon work
 792 supported by the National Science Foundation under Grant No. DEB-1845686. Any opinions,
 793 findings, and conclusions or recommendations expressed in this material are those of the authors
 794 and do not necessarily reflect the views of the National Science Foundation.

795 REFERENCES

- 796 Abbott, J. K., A. K. Nordén, and B. Hansson. 2017. Sex chromosome evolution: historical insights and
797 future perspectives. *Proc. Biol. Sci.* 284.
- 798 Adhikari, K., J. H. Son, A. H. Rensink, J. Jaweria, D. Bopp, L. W. Beukeboom, and R. P. Meisel. 2021.
799 Temperature-dependent effects of house fly proto-Y chromosomes on gene expression.
- 800 Allen, S. L., R. Bonduriansky, C. M. Sgro, and S. F. Chenoweth. 2017. Sex-biased transcriptome
801 divergence along a latitudinal gradient. *Mol. Ecol.* 26:1256–1272.
- 802 Arnqvist, G., N. Vellnow, and L. Rowe. 2014. The effect of epistasis on sexually antagonistic genetic
803 variation. *Proc. Biol. Sci.* 281.
- 804 Atkinson, D. 1996. Ectotherm life-history responses to developmental temperature. Pp. 183–204 in A. F.
805 B. Ian A. Johnston, ed. *Animals and Temperature: Phenotypic and Evolutionary Adaptation*.
- 806 Baker, R. H., and G. S. Wilkinson. 2010. Comparative genomic hybridization (CGH) reveals a neo-X
807 chromosome and biased gene movement in stalk-eyed flies (genus *Teleopsis*). *PLoS Genetics*
808 6:e1001121.
- 809 Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4.
810 *Journal of Statistical Software* 67:1–48.
- 811 Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful
812 approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 57:289–300. [Royal Statistical
813 Society, Wiley].
- 814 Benoit, J. B., A. Vigneron, N. A. Broderick, Y. Wu, J. S. Sun, J. R. Carlson, S. Aksoy, and B. L. Weiss.
815 2017. Symbiont-induced odorant binding proteins mediate insect host hematopoiesis. *Elife* 6.
- 816 Berson, J. D., F. Garcia-Gonzalez, and L. W. Simmons. 2019a. Experimental evidence for the role of
817 sexual selection in the evolution of cuticular hydrocarbons in the dung beetle, *Onthophagus taurus*.
818 *J. Evol. Biol.* 32:1186–1193.

819 Berson, J. D., M. Zuk, and L. W. Simmons. 2019b. Natural and sexual selection on cuticular
820 hydrocarbons: a quantitative genetic analysis. *Proc. Biol. Sci.* 286:20190677.

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

823 Bliss, C. I. 1926. Temperature characteristics for prepupal development in *Drosophila melanogaster*. *J.*
824 *Gen. Physiol.* 9:467–495.

825 Blomquist, G. J., and A.-G. Bagnères. 2010. Insect Hydrocarbons: Biology, Biochemistry, and Chemical
826 Ecology. Cambridge University Press.

827 Bray, N. L., H. Pimentel, P. Melsted, and L. Pachter. 2016. Near-optimal probabilistic RNA-seq
828 quantification. *Nat. Biotechnol.* 34:525–527.

829 Brooks, R., and J. A. Endler. 2001. Female guppies agree to differ: phenotypic and genetic variation in
830 mate-choice behavior and the consequences for sexual selection. *Evolution* 55:1644–1655. Wiley
831 Online Library.

832 Bryant, E. H. 1980. Geographic variation in components of mating success of the housefly, *Musca*
833 *domestica* L., in the United States. *The American Naturalist* 116:655–669.

834 Carroll, S. B. 2005. Evolution at two levels: on genes and form. *PLoS Biol.* 3:e245.

835 Castel, S. E., A. Levy-Moonshine, P. Mohammadi, E. Banks, and T. Lappalainen. 2015. Tools and best
836 practices for data processing in allelic expression analysis. *Genome Biol.* 16:195.

837 Charlesworth, D., B. Charlesworth, and G. Marais. 2005. Steps in the evolution of heteromorphic sex
838 chromosomes. *Heredity* 95:118–128.

839 Chenoweth, S. F., and K. McGuigan. 2010. The genetic basis of sexually selected variation. *Annu. Rev.*
840 *Ecol. Evol. Syst.* 41:81–101. Annual Reviews.

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

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

845 Chung, H., and S. B. Carroll. 2015. Wax, sex and the origin of species: Dual roles of insect cuticular
846 hydrocarbons in adaptation and mating. Bioessays 37:822–830.

847 Connallon, T. 2015. The geography of sex-specific selection, local adaptation, and sexual dimorphism.
848 Evolution 69:2333–2344.

849 Connallon, T., and A. G. Clark. 2011. The resolution of sexual antagonism by gene duplication. Genetics
850 187:919–937.

851 Connallon, T., S. Sharma, and C. Olito. 2019. Evolutionary consequences of sex-specific selection in
852 variable environments: four simple models reveal diverse evolutionary outcomes. Am. Nat.
853 193:93–105.

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

856 Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, and T. R.
857 Gingeras. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21.

858 Edgar, R. C. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space
859 complexity. BMC Bioinformatics 5:113.

860 Emerson, J. J., H. Kaessmann, E. Betrán, and M. Long. 2004. Extensive gene traffic on the mammalian X
861 chromosome. Science 303:537–540.

862 Ezaz, T., K. Srikulnath, and J. A. M. Graves. 2017. Origin of amniote sex chromosomes: an ancestral
863 super-sex chromosome, or common requirements? J. Hered. 108:94–105.

864 Feldmeyer, B., M. Kozielska, and B. Kuijper. 2008. Climatic variation and the geographical distribution
865 of sex-determining mechanisms in the housefly. Evolutionary. evolutionary-ecology.com.

866 Fenster, C. B., L. F. Galloway, and L. Chao. 1997. Epistasis and its consequences for the evolution of

867 natural populations. Trends Ecol. Evol. 12:282–286.

868 Fitzpatrick, M. J. 2004. Pleiotropy and the genomic location of sexually selected genes. Am. Nat.

869 163:800–808.

870 Foster, G. G., M. J. Whitten, C. Konovalov, J. T. A. Arnold, and G. Maffi. 1981. Autosomal genetic maps

871 of the Australian Sheep Blowfly, *Lucilia cuprina dorsalis* R.-D. (Diptera: Calliphoridae), and

872 possible correlations with the linkage maps of *Musca domestica* L. and *Drosophila melanogaster*

873 (Mg.). Genet. Res. 37:55–69. Cambridge University Press.

874 Fox, J., M. Friendly, and S. Weisberg. 2013. Hypothesis tests for multivariate linear models using the car

875 package. The R Journal 5:39.

876 Franco, M. G., P. G. Rubini, and M. Vecchi. 1982. Sex-determinants and their distribution in various

877 populations of *Musca domestica* L. of Western Europe. Genet. Res. 40:279–293.

878 Furman, B. L. S., and B. J. Evans. 2016. Sequential turnovers of sex chromosomes in African clawed

879 frogs (*Xenopus*) suggest some genomic regions are good at sex determination. G3 6:3625–3633.

880 Gallach, M., and E. Betrán. 2011. Intralocus sexual conflict resolved through gene duplication. Trends

881 Ecol. Evol. 26:222–228.

882 Gibson, J. R., A. K. Chippindale, and W. R. Rice. 2002. The X chromosome is a hot spot for sexually

883 antagonistic fitness variation. Proceedings of the Royal Society of London B: Biological Sciences

884 269:499–505. The Royal Society.

885 Greene, E., B. E. Lyon, V. R. Muehter, L. Ratcliffe, S. J. Oliver, and P. T. Boag. 2000. Disruptive sexual

886 selection for plumage coloration in a passerine bird. Nature 407:1000–1003.

887 Hamm, R. L., J.-R. Gao, G. G.-H. Lin, and J. G. Scott. 2009. Selective advantage for III^M males over Y^M

888 males in cage competition, mating competition, and pupal emergence in *Musca domestica* L.

889 (Diptera: Muscidae). Environ. Entomol. 38:499–504.

890 Hamm, R. L., R. P. Meisel, and J. G. Scott. 2015. The evolving puzzle of autosomal versus Y-linked male

determination in *Musca domestica*. G3 5:371–384.

Hamm, R. L., T. Shono, and J. G. Scott. 2005. A cline in frequency of autosomal males is not associated with insecticide resistance in house fly (Diptera: Muscidae). J. Econ. Entomol. 98:171–176.

Hedrick, P. W., M. E. Ginevan, and E. P. Ewing. 1976. Genetic polymorphism in heterogeneous environments. Annu. Rev. Ecol. Syst. 7:1–32. Annual Reviews.

Heinen-Kay, J. L., R. E. Nichols, and M. Zuk. 2020. Sexual signal loss, pleiotropy, and maintenance of a male reproductive polymorphism in crickets. Evolution 74:1002–1009.

Ingleby, F. C., J. Hunt, and D. J. Hosken. 2010. The role of genotype-by-environment interactions in sexual selection. J. Evol. Biol. 23:2031–2045.

Jia, F. Y., M. D. Greenfield, and R. D. Collins. 2000. Genetic variance of sexually selected traits in waxmoths: maintenance by genotype x environment interaction. Evolution 54:953–967.

Johnston, S. E., J. Gratten, C. Berenos, J. G. Pilkington, T. H. Clutton-Brock, J. M. Pemberton, and J. Slate. 2013. Life history trade-offs at a single locus maintain sexually selected genetic variation. Nature 502:93–95.

Jones, A. G. 2016. Theory of sexual selection. Pp. 119–127 in R. M. Kliman, ed. Encyclopedia of Evolutionary Biology. Academic Press, Oxford.

Kirkpatrick, M., and M. J. Ryan. 1991. The evolution of mating preferences and the paradox of the lek. Nature 350:33–38.

Kokko, H., and K. Heubel. 2008. Condition-dependence, genotype-by-environment interactions and the lek paradox. Genetica 134:55–62.

Konrad, K. D., T. J. Goralski, A. P. Mahowald, and J. L. Marsh. 1998. The gastrulation defective gene of *Drosophila melanogaster* is a member of the serine protease superfamily. Proc. Natl. Acad. Sci. U. S. A. 95:6819–6824.

Košťál, V., J. Korbelová, T. Štětina, R. Poupardin, H. Colinet, H. Zahradníčková, I. Opekarová, M. Moos,

915 and P. Šimek. 2016. Physiological basis for low-temperature survival and storage of quiescent larvae
916 of the fruit fly *Drosophila melanogaster*. Sci. Rep. 6:32346.

917 Kozielska, M., B. Feldmeyer, I. Pen, F. J. Weissing, and L. W. Beukeboom. 2008. Are autosomal
918 sex-determining factors of the housefly (*Musca domestica*) spreading north? Genet. Res.
919 90:157–165. Cambridge University Press.

920 Kumar, S., G. Stecher, M. Li, C. Knyaz, and K. Tamura. 2018. MEGA X: Molecular evolutionary
921 genetics analysis across computing platforms. Mol. Biol. Evol. 35:1547–1549.

922 Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution
923 37:1210–1226.

924 Langfelder, P., and S. Horvath. 2008. WGCNA: an R package for weighted correlation network analysis.
925 BMC Bioinformatics 9:559.

926 Langfelder, P., P. S. Mischel, and S. Horvath. 2013. When is hub gene selection better than standard
927 meta-analysis? PLoS One 8:e61505.

928 Lasne, C., S. B. Hangartner, T. Connallon, and C. M. Sgrò. 2018. Cross-sex genetic correlations and the
929 evolution of sex-specific local adaptation: Insights from classical trait clines in *Drosophila*
930 *melanogaster*. Evolution 72:1317–1327. Wiley Online Library.

931 Leek, J. T., W. E. Johnson, H. S. Parker, A. E. Jaffe, and J. D. Storey. 2012. The sva package for
932 removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics
933 28:882–883.

934 Levene, H. 1953. Genetic equilibrium when more than one ecological niche is available. Am. Nat.
935 87:331–333. The University of Chicago Press.

936 Li, A., and S. Horvath. 2007. Network neighborhood analysis with the multi-node topological overlap
937 measure. Bioinformatics 23:222–231.

938 Lockey, K. H. 1988. Lipids of the insect cuticle: origin, composition and function. Comparative

939 Biochemistry and Physiology Part B: Comparative Biochemistry 89:595–645.

940 Loeb, J., and J. H. Northrop. 1917. On the influence of food and temperature upon the duration of life.

941 Journal of Biological Chemistry 32:103–121.

942 Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for

943 RNA-seq data with DESeq2. Genome Biol. 15:550.

944 Maere, S., K. Heymans, and M. Kuiper. 2005. BiNGO: a Cytoscape plugin to assess overrepresentation of

945 Gene Ontology categories in Biological Networks. Bioinformatics 21:3448–3449.

946 Marcillac, F., Y. Grosjean, and J.-F. Ferveur. 2005. A single mutation alters production and discrimination

947 of *Drosophila* sex pheromones. Proc. Biol. Sci. 272:303–309.

948 McGraw, E. A., Y. H. Ye, B. Foley, S. F. Chenoweth, M. Higgie, E. Hine, and M. W. Blows. 2011.

949 High-dimensional variance partitioning reveals the modular genetic basis of adaptive divergence in

950 gene expression during reproductive character displacement. Evolution 65:3126–3137.

951 McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D.

952 Altshuler, S. Gabriel, M. Daly, and M. A. DePristo. 2010. The Genome Analysis Toolkit: a

953 MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res.

954 20:1297–1303.

955 Meisel, R. P., T. Davey, J. H. Son, A. C. Gerry, T. Shono, and J. G. Scott. 2016. Is multifactorial sex

956 determination in the house fly, *Musca domestica* (L.), stable over time? J. Hered. 107:615–625.

957 Meisel, R. P., C. A. Gonzales, and H. Luu. 2017. The house fly Y Chromosome is young and minimally

958 differentiated from its ancient X Chromosome partner. Genome Res. 27:1417–1426.

959 Meisel, R. P., M. V. Han, and M. W. Hahn. 2009. A complex suite of forces drives gene traffic from

960 *Drosophila* X chromosomes. Genome Biol. Evol. 1:176–188.

961 Meisel, R. P., and J. G. Scott. 2018. Using genomic data to study insecticide resistance in the house fly,

962 *Musca domestica*. Pestic. Biochem. Physiol. 151:76–81.

963 Meisel, R. P., J. G. Scott, and A. G. Clark. 2015. Transcriptome differences between alternative sex
964 determining genotypes in the house fly, *Musca domestica*. *Genome Biol. Evol.* 7:2051–2061.

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

967 Olafson, P. U., S. Aksoy, G. M. Attardo, G. Buckmeier, X. Chen, C. J. Coates, M. Davis, J. Dykema, S. J.
968 Emrich, M. Friedrich, C. J. Holmes, P. Ioannidis, E. N. Jansen, E. C. Jennings, D. Lawson, E. O.
969 Martinson, G. L. Maslen, R. P. Meisel, T. D. Murphy, D. Nayduch, D. R. Nelson, K. J. Oyen, T. J.
970 Raszick, J. M. C. Ribeiro, H. M. Robertson, A. J. Rosendale, T. B. Sackton, P. Saelao, S. L. Swiger,
971 S.-H. Sze, A. M. Tarone, D. B. Taylor, W. C. Warren, R. M. Waterhouse, M. T. Weirauch, J. H.
972 Werren, R. K. Wilson, E. M. Zdobnov, and J. B. Benoit. 2021. The genome of the stable fly,
973 *Stomoxys calcitrans*, reveals potential mechanisms underlying reproduction, host interactions, and
974 novel targets for pest control. *BMC Biol.* 19. Springer Science and Business Media LLC.

975 Olafson, P. U., and C. A. Saski. 2020. Chemosensory-Related Gene Family Members of the Horn Fly,
976 (Diptera: Muscidae), Identified by Transcriptome Analysis. *Insects* 11.

977 O’Meally, D., T. Ezaz, A. Georges, S. D. Sarre, and J. A. M. Graves. 2012. Are some chromosomes
978 particularly good at sex? Insights from amniotes. *Chromosome Res.* 20:7–19.

979 Otte, T., M. Hilker, and S. Geiselhardt. 2018. Phenotypic plasticity of cuticular hydrocarbon profiles in
980 insects. *J. Chem. Ecol.* 44:235–247.

981 Park, S. K., K. J. Mann, H. Lin, E. Starostina, A. Kolski-Andreaco, and C. W. Pikielny. 2006. A
982 *Drosophila* protein specific to pheromone-sensing gustatory hairs delays males’ copulation attempts.
983 *Curr. Biol.* 16:1154–1159.

984 Ritchie, M. E., B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, and G. K. Smyth. 2015. limma powers
985 differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.*
986 43:e47.

987 Rueffler, C., T. J. M. Van Dooren, O. Leimar, and P. A. Abrams. 2006. Disruptive selection and then
988 what? Trends Ecol. Evol. 21:238–245.

989 Ruzicka, F., L. Dutoit, P. Czuppon, C. Y. Jordan, X. Li, C. Olito, A. Runemark, E. I. Svensson, H. P.
990 Yazdi, and T. Connallon. 2020. The search for sexually antagonistic genes: Practical insights from
991 studies of local adaptation and statistical genomics. Evolution Letters 57:289.

992 Schaeffer, S. W. 2018. Muller “elements” in *Drosophila*: How the search for the genetic basis for
993 speciation led to the birth of comparative genomics. Genetics 210:3–13.

994 Scott, J. G., P. Sridhar, and N. Liu. 1996. Adult specific expression and induction of cytochrome P450lpr
995 in house flies. Arch. Insect Biochem. Physiol. 31:313–323.

996 Scott, J. G., W. C. Warren, L. W. Beukeboom, D. Bopp, A. G. Clark, S. D. Giers, M. Hediger, A. K.
997 Jones, S. Kasai, C. A. Leichter, M. Li, R. P. Meisel, P. Minx, T. D. Murphy, D. R. Nelson, W. R.
998 Reid, F. D. Rinkevich, H. M. Robertson, T. B. Sackton, D. B. Sattelle, F. Thibaud-Nissen, C.
999 Tomlinson, L. van de Zande, K. K. O. Walden, R. K. Wilson, and N. Liu. 2014. Genome of the house
1000 fly, *Musca domestica* L., a global vector of diseases with adaptations to a septic environment.
1001 Genome Biol. 15:466.

1002 Seehausen, and van Alphen. 1999. Can sympatric speciation by disruptive sexual selection explain rapid
1003 evolution of cichlid diversity in Lake Victoria? Ecol. Lett. 2:262–271.

1004 Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and
1005 T. Ideker. 2003. Cytoscape: a software environment for integrated models of biomolecular
1006 interaction networks. Genome Res. 13:2498–2504.

1007 Shono, T., and J. G. Scott. 2003. Spinosad resistance in the housefly, *Musca domestica*, is due to a
1008 recessive factor on autosome 1. Pestic. Biochem. Physiol. 75:1–7.

1009 Shorter, J. R., L. M. Dembeck, L. J. Everett, T. V. Morozova, G. H. Arya, L. Turlapati, G. E. St Armour,
1010 C. Schal, T. F. C. Mackay, and R. R. H. Anholt. 2016. *Obp56h* modulates mating behavior in

1011 *Drosophila melanogaster*. G3 6:3335–3342.

1012 Son, J. H., T. Kohlbrenner, S. Heinze, L. W. Beukeboom, D. Bopp, and R. P. Meisel. 2019. Minimal
 1013 effects of proto-Y chromosomes on house fly gene expression in spite of evidence that selection
 1014 maintains stable polygenic sex determination. Genetics 213:313–327.

1015 Son, J. H., and R. P. Meisel. 2021. Gene-level, but not chromosome-wide, divergence between a very
 1016 young house fly proto-Y chromosome and its homologous proto-X chromosome. Mol. Biol. Evol.
 1017 38:606–618.

1018 Sun, J. S., S. Xiao, and J. R. Carlson. 2018. The diverse small proteins called odorant-binding proteins.
 1019 Open Biol. 8:180208.

1020 Takehana, Y., M. Matsuda, T. Myosho, M. L. Suster, K. Kawakami, T. Shin-I, Y. Kohara, Y. Kuroki, A.
 1021 Toyoda, A. Fujiyama, S. Hamaguchi, M. Sakaizumi, and K. Naruse. 2014. Co-option of *Sox3* as the
 1022 male-determining factor on the Y chromosome in the fish *Oryzias dancena*. Nat. Commun. 5:4157.

1023 Thomas, M. L., and L. W. Simmons. 2009. Sexual selection on cuticular hydrocarbons in the Australian
 1024 field cricket, *Teleogryllus oceanicus*.

1025 Tomita, T., and Y. Wada. 1989. Multifactorial sex determination in natural populations of the housefly
 1026 (*Musca domestica*) in Japan. The Japanese Journal of Genetics 64:373–382. The Genetics Society of
 1027 Japan.

1028 Tomkins, J. L., J. Radwan, J. S. Kotiaho, and T. Tregenza. 2004. Genic capture and resolving the lek
 1029 paradox. Trends Ecol. Evol. 19:323–328.

1030 Turelli, M., and N. H. Barton. 2004. Polygenic variation maintained by balancing selection: pleiotropy,
 1031 sex-dependent allelic effects and G x E interactions. Genetics 166:1053–1079.

1032 Valanne, S., J.-H. Wang, and M. Rämet. 2011. The *Drosophila* Toll signaling pathway. The Journal of
 1033 Immunology 186:649–656.

1034 van Doorn, G. S., and M. Kirkpatrick. 2007. Turnover of sex chromosomes induced by sexual conflict.

1035 Nature 449:909–912.

1036 van Doorn, G. S., and F. J. Weissing. 2002. Ecological versus sexual selection models of sympatric
1037 speciation: a synthesis. *Selection* 2:17–40. Akadémiai Kiadó.

1038 Van Kuren, N. W., and M. Long. 2018. Gene duplicates resolving sexual conflict rapidly evolved essential
1039 gametogenesis functions. *Nature Ecology & Evolution* 2:705–712.

1040 Weller, G. L., and G. G. Foster. 1993. Genetic maps of the sheep blowfly *Lucilia cuprina*: linkage-group
1041 correlations with other dipteran genera. *Genome* 36:495–506.

1042 Wiegmann, B. M., M. D. Trautwein, I. S. Winkler, N. B. Barr, J.-W. Kim, C. Lambkin, M. A. Bertone, B.
1043 K. Cassel, K. M. Bayless, A. M. Heimberg, B. M. Wheeler, K. J. Peterson, T. Pape, B. J. Sinclair, J.
1044 H. Skevington, V. Blagoderov, J. Caravas, S. N. Kutty, U. Schmidt-Ott, G. E. Kampmeier, F. C.
1045 Thompson, D. A. Grimaldi, A. T. Beckenbach, G. W. Courtney, M. Friedrich, R. Meier, and D. K.
1046 Yeates. 2011. Episodic radiations in the fly tree of life. *Proc. Natl. Acad. Sci. U. S. A.*
1047 108:5690–5695.

1048 Zhou, J.-J. 2010. Odorant-binding proteins in insects. *Vitamins & Hormones* 241–272.