

Male-biased gene expression resolves sexual conflict through the evolution of sex-specific genetic architecture

Alison E. Wright,^{1,2} Matteo Fumagalli,³ Christopher R. Cooney,¹ Natasha I. Bloch,⁴ Filipe G. Vieira,⁵ Severine D. Buechel,⁶ Niclas Kolm,⁶ and Judith E. Mank^{4,7}

¹Department of Animal and Plant Sciences, University of Sheffield, Sheffield, United Kingdom

²E-mail: a.e.wright@sheffield.ac.uk

³Department of Life Sciences, Silwood Park Campus, Imperial College London, London, United Kingdom

⁴Department of Genetics, Evolution and Environment, University College London, London, United Kingdom

⁵Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark

⁶Department of Zoology, Stockholm University, Stockholm, Sweden

⁷Department of Organismal Biology, Uppsala University, Uppsala, Sweden

Received December 4, 2017

Accepted January 9, 2018

Many genes are subject to contradictory selection pressures in males and females, and balancing selection resulting from sexual conflict has the potential to substantially increase standing genetic diversity in populations and thereby act as an important force in adaptation. However, the underlying causes of sexual conflict, and the potential for resolution, remains hotly debated. Using transcriptome-resequencing data from male and female guppies, we use a novel approach, combining patterns of genetic diversity and intersexual divergence in allele frequency, to distinguish the different scenarios that give rise to sexual conflict, and how this conflict may be resolved through regulatory evolution. We show that reproductive fitness is the main source of sexual conflict, and this is resolved via the evolution of male-biased expression. Furthermore, resolution of sexual conflict produces significant differences in genetic architecture between males and females, which in turn lead to specific alleles influencing sex-specific viability. Together, our findings suggest an important role for sexual conflict in shaping broad patterns of genome diversity, and show that regulatory evolution is a rapid and efficient route to the resolution of conflict.

KEY WORDS: Balancing selection, gene expression, population genetics, sexual conflict.

Impact Summary

Males and females are often subject to different and sometimes contradictory selection pressures, yet they share an almost identical set of genes. This can give rise to intra-locus sexual conflict, where an allele benefits one sex at the expense of the other. Sexual conflict can result from different types of selection pressures, including reproductive fitness and survival. However, the underlying causes of sexual conflict, and the potential for resolution, remain hotly debated. We combine several powerful population genetic approaches with gene expression analyses in the Trinidadian guppy, a classic ecological model species for the study of sexual selection, to study the

nature, locus and potential for resolution of sexual conflict across the genome. Our results suggest that the majority of sexual conflict is produced through conflicting selection over reproductive interests of males and females, and that sexual conflict has the potential to maintain genetic diversity through balancing selection. Furthermore, we find that regulatory evolution, the evolution of gene expression differences between males and females, is a rapid and efficient route to the resolution of conflict. This produces significant differences in genetic architecture between males and females, where alleles have phenotypic effects in one sex but not the other. Together, our findings suggest an important role for sexual

conflict in shaping broad patterns of genome diversity and evolution.

Males and females often experience different selection pressures, and when this occurs for traits with a shared genetic basis between the sexes, significant amounts of intralocus sexual conflict can result (Bonduriansky and Chenoweth 2009). As a consequence, intralocus sexual conflict is thought to be widespread across the genome (Chippindale et al. 2001), potentially affecting a large proportion of loci. Sexual conflict can result from different types of selection pressures, including reproductive fitness and viability, and it remains unclear which of these forces is the primary mechanism underlying sexual conflict.

Moreover, intralocus conflict can potentially be resolved, and this is often assumed to occur through the evolution of gene expression differences between females and males, ultimately leading to phenotypic dimorphism (Pointer et al. 2013; Hollis et al. 2014; Immonen et al. 2014; Mank 2017). The mechanisms by which sexual dimorphism in gene expression resolves sexual conflict within the genome has been the focus of considerable recent debate. Some work has suggested that the evolution of sex-biased expression may represent a footprint of resolved conflict between males and females (Innocenti and Morrow 2010), and there is increasing evidence that many loci exhibit sex differences in their phenotypic effects, otherwise defined as separate genetic architecture (Gilks et al. 2014; Karp et al. 2017), which would result from the effective resolution of conflict. However, other approaches (Cheng and Kirkpatrick 2016) have suggested that sexual conflict remains unresolved for a substantial proportion of sex-biased genes.

Sexual conflict leaves distinct population genetic signatures in sequence data, and patterns of genetic diversity and intersexual divergence in allele frequency offer complementary views into the mechanisms giving rise to sexual conflict. Intralocus sexual conflict leads to contrasting selection pressures depending on whether alleles are present in females or males, producing balancing selection. This in turn results in elevated genetic diversity, which can be measured with Tajima's D (Tajima 1989), an estimate of the relative proportion of variable sites in a given locus. Indeed, higher rates of balancing selection have been detected for partially sex-linked loci (Qiu et al. 2013; Guirao-Rico et al. 2017), consistent with sexual conflict theory predictions (Otto et al. 2011; Jordan and Charlesworth 2012; Kirkpatrick and Guerrero 2014). However, it is important to note that balancing selection can be caused by a number of different selective forces (Slate 2005; Mokkonen et al. 2011; Huang et al. 2014), in addition to sexual conflict, and that confounding population dynamics (e.g., bottlenecks) must be accounted for when estimating the strength of balancing selection (Hartl and Clark 2007).

Sexual conflict can arise from several forces, and the population genetic signature varies according to the type of sexual conflict (Table 1). Sexual conflict can result over reproductive fitness, where an allele increases the reproductive success of one sex at a cost to the other (Lonn et al. 2017). However, sexual conflict can also result when an allele has differential effects on viability, mortality or predation between males and females. Tajima's D alone cannot disentangle these mechanisms, and it is important to incorporate other population genetic approaches to determine the nature and mechanism of sexual conflict.

Intersexual F_{ST} , which measures the genetic difference between males and females in a population for a given locus, makes it possible to further differentiate these scenarios and is therefore an important complement to Tajima's D . We expect F_{ST} to deviate from neutrality only if loci influence viability, mortality, or predation differently between males and females (Lewontin and Krakauer 1973), but not for sexual conflict due to fecundity or reproductive fitness. This is because the allele frequencies are defined at the start of each generation by Hardy–Weinberg equilibrium before selection and are identical between the sexes at conception. Therefore, different allele frequencies in adults are assumed to be the result of sexual conflict over viability or survival.

Recent studies have employed intersexual F_{ST} to identify genes with sexually antagonistic fitness effects (Cheng and Kirkpatrick 2016; Lucotte et al. 2016). However, F_{ST} in isolation cannot distinguish loci subject to sexual conflict over viability or survival from loci where sexual conflict has been resolved through the evolution of separate genetic architectures (Table 1). In the latter case, a mutation can affect fitness in one sex, but have little or no effect in the other. It is increasingly clear that many traits, including somatic phenotypes, have distinct genetic architecture in males and females (Randall et al. 2013; Dapper and Wade 2016; Karp et al. 2017), and this has the potential to produce significant intersexual F_{ST} . However, this is not the result of sexual conflict, and will not produce signatures of balancing selection as measured by Tajima's D .

Comparisons between intersexual F_{ST} and Tajima's D therefore offer a powerful approach to investigate the underlying causes of sexual conflict, as well as the potential for sex-specific gene regulation to resolve this conflict (Gilks et al. 2014; Karp et al. 2017). We employed this novel, combined approach, deliberately choosing a closed, seminatural population of guppies to remove any biases due to sex-differences in predation or dispersal. This allows us to focus exclusively on reproductive fitness versus viability selection. We find male-biased expression resolves sexual conflict over reproductive fitness and that sex differences in viability are not due to intralocus sexual antagonism, rather loci only affecting viability in one sex due to sex-specific genetic architecture. Together, our results offer new insights into the mechanisms

Table 1. Distinguishing types of sexual conflict through contrasts between intersexual F_{ST} and Tajima's D.

Scenario	Cause	Tajima's D	Intersexual F_{ST}
I.	Sexual conflict due to differences in reproductive fitness	High	Low
II.	Sexual conflict due to differences in viability selection	High	High
III.	Sex-specific viability effects	Low	High

by which sexual conflict is resolved, and the fitness consequences of sex-biased gene expression.

Results

Balancing selection can be the result of different selective forces, including sexual conflict (Charlesworth 2006), heterozygote advantage (Slate 2005), spatially or temporally varying selection (Huang et al. 2014) and frequency-dependent selection (Mokkonen et al. 2011). To test our power to detect these forces in our dataset, we first measured Tajima's D for genes associated with immunity, which are known to exhibit high levels of heterozygote advantage in a broad array of animals (Stahl et al. 1999; Hedrick 2011; Ghosh et al. 2012). We used ANGSD (Korneliussen et al. 2014) to estimate population genomic statistics as it implements methods to account to sequencing uncertainty and is appropriate for uneven sequencing depth associated with transcriptome data. We detected significantly higher Tajima's D for immune genes compared to all other autosomal genes (Wilcoxon test $P = 0.015$, Fig. 1A), suggesting that we

have sufficient power to detect balancing selection in general. For all subsequent analyses, we removed these immune loci to reduce any potential confounding effects from heterozygote advantage. We also accounted for inbreeding and population structure across our population, as both factors can influence population genomic statistics (Supporting Information).

We next tested our ability to detect the signature of sexual conflict by assessing Tajima's D for loci on the pseudo-autosomal region (PAR) of the guppy sex chromosome. Partially sex-linked regions have been both predicted (Otto et al. 2011; Jordan and Charlesworth 2012; Kirkpatrick and Guerrero 2014) and shown (Qiu et al. 2013; Guirao-Rico et al. 2017) to exhibit higher levels of balancing selection due to sexual conflict. Using the PAR boundary that we previously identified in this population (Wright et al. 2017), we also detected significantly elevated Tajima's D for PAR loci compared to autosomal portions of the genome (Wilcoxon test $P = 0.033$, Fig. 1B). This and the analysis of immunity loci indicate that we have sufficient power to detect balancing selection and sexual conflict in our dataset. To remove any influence of

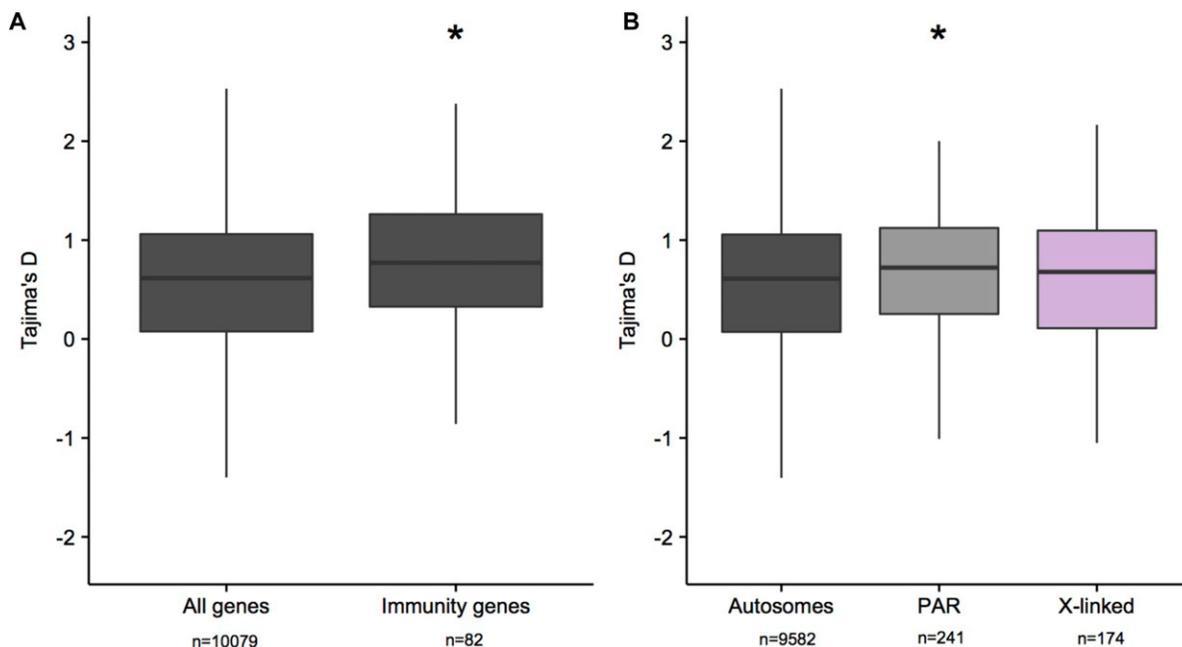


Figure 1. Distribution of Tajima's D across categories of genes. (A) Distribution of Tajima's D across immunity genes predicted to be under balancing selection. *Indicates a significantly elevated Tajima's D relative to all genes (Wilcoxon test $P < 0.03$). (B) Distribution of Tajima's D across categories of sex-linked genes after excluding immunity genes. *Indicates a significantly elevated Tajima's D relative to the autosomes (Wilcoxon test $P < 0.05$).

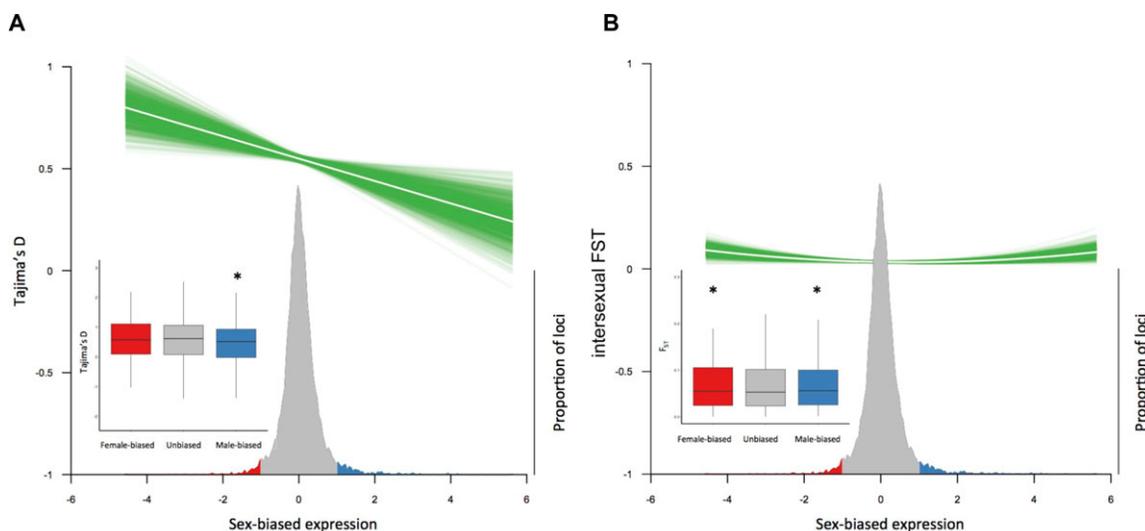


Figure 2. Sex-biased gene expression and sexual conflict. White line indicates the predicted relationship between two variables, and green indicates the probability distribution of the fitted line from 1000 bootstrap replicates. Density plot shows the distribution of sex-biased expression across all genes. Female-biased genes (\log_2 fold change < -1) are in red and male-biased genes (\log_2 fold change > 1) are in blue. (A) Relationship between Tajima's D and sex-bias in expression across all autosomes excluding immunity genes. Inset shows distribution of Tajima's D across categories of sex-biased genes. *Indicates a significantly different Tajima's D relative to unbiased genes (Wilcoxon test $P < 0.02$). (B) Relationship between F_{ST} and sex-bias in expression across all autosomes. Inset shows distribution of F_{ST} across categories of sex-biased genes. *Indicates a significantly different F_{ST} relative to unbiased genes (Wilcoxon test $P < 0.03$).

accumulated sexual conflict on the PAR, we removed sex-linked loci from the remainder of our analyses.

We next assessed Tajima's D for autosomal genes as a function of sex-biased expression. We measured male and female transcription in guppy tails, which we selected because it includes tissue related to both reproductive fitness and survival. Male coloration has been shown to be an important factor in female mate choice and male reproductive fitness in many populations of guppies (Houde and Endler 1990), including the population we use here (Corral-López et al. 2017), and genes transcribed in our samples include those related to coloration. Tail tissue also contains skin and somatic tissues that interface with the environment, including the lateral line, and therefore are important to viability and survival. We followed the approach taken by Cheng and Kirkpatrick (Cheng and Kirkpatrick 2016) to fit a parametric model to describe the relationship between Tajima's D and sex-biased expression for autosomal genes. The best model for the relationship between sex-bias and Tajima's D across the autosomes was linear (Fig. 2A, intercept = 0.549, slope = -0.055 , model statistics in Tables S2 and S3). The slope of our best-fit line was not due to an increase in Tajima's D for female-biased genes compared to unbiased genes (Wilcoxon test $P = 0.827$, Fig. 2A), rather a decrease for male-biased genes (Wilcoxon test $P = 0.011$, Fig. 2A). This suggests that male-biased gene expression largely resolves sexual conflict.

Furthermore, Tajima's D cannot differentiate sexual conflict resulting from reproductive fitness from conflict resulting from

viability selection. To investigate the relative roles of reproductive fitness or viability selection in generating sexual conflict, we assessed intersexual F_{ST} , which we would expect to deviate from neutrality if loci influence viability differently between males and females. Intersexual F_{ST} can also be influenced by sex differences in dispersal or predation; however, these forces are not factors in our closed population. We observed a 2nd degree polynomial pattern, otherwise known as a positive parabola, when we correlated intersexual F_{ST} and sex-biased expression (Fig. 2B, model statistics in Tables S4 and S5). F_{ST} is significantly elevated for both female-biased (Wilcoxon test $P = 0.026$, Fig. 2B) and male-biased genes (Wilcoxon test $P = 0.019$, Fig. 2B) relative to unbiased genes. This pattern suggests that sex-differences in viability exist in our population sample. However, the relationship between viability, as assessed by intersexual F_{ST} , and sex-bias is much less pronounced than the differences we observe for Tajima's D (Fig. 2, Table S6), suggesting that most sexual conflict in our population results from alleles that differentially affect male and female reproductive fitness. Furthermore, we conducted simulations (Supporting Information) to confirm that the patterns of Tajima's D and F_{ST} we observe are not the result of uneven sequencing depth for genes expressed differently between males and females, or differences in the number of samples of each sex. We simulated various scenarios where the sequencing depth varied between males and females, and found no effect of uneven coverage or sample number for estimating Tajima's D (minimum Kendall's correlation $\tau = 0.91$) or intersexual F_{ST} (minimum

Kendall's correlation $\tau = 0.78$) (Supporting Information). Finally, we conducted a power assessment for detecting outliers of summary statistics, and found that our experimental set up is sufficient to generate reliable predictions of outliers (Supporting Information).

To further investigate the drivers and resolution of sexual conflict across the genome, we mapped Tajima's D against intersexual F_{ST} for autosomal loci. It is important to remember that sexual conflict due to viability and reproductive fitness are not mutually exclusive, and our combined use of Tajima's D and intersexual F_{ST} allows us to tease these forces apart. If sexual conflict arises from conflict over reproductive fitness, where an allele increases reproductive fitness in one sex at the same time that it exacts a reproductive cost in the other, we would not expect deviations from neutrality in intersexual F_{ST} , and only Tajima's D will be elevated (Table 1, Scenario I). However, if conflict is due to sex-differences in viability, where an allele increases viability in one sex at the same time that it exacts a viability cost in the other, we would expect both high intersexual F_{ST} and Tajima's D (Table 1, Scenario II). Figure 2A suggests that male-biased expression resolves sexual conflict and this could either be the result of conflict over reproductive fitness (Scenario I), associated with a high Tajima's D and low intersexual F_{ST} , or conflict over viability selection (Scenario II), which produces both high Tajima's D and F_{ST} . We observe a significant deficit of male-biased genes under Scenario I (Table 2, $P = 0.040$), but not Scenario II ($P = 0.790$), consistent with the notion that male-biased expression effectively resolves sexual conflict over reproductive fitness.

Male guppies show a remarkable variety of coloration patterns in the wild, and our closed, outbred population shows a similar diversity of coloration. This means that our male samples exhibit high transcriptional diversity, and this variance confounds traditional methods to determine sex-bias. We therefore used permutation testing to assess significant levels of sex-bias, a method previously implemented on guppy transcriptome analysis (Ghalambor et al. 2015), in combination with traditional fold-change thresholds (doubled expression in one sex compared to the other). Using this approach, we find a similar deficit of male-biased genes under Scenario I, however, the difference is non-significant, likely because of a limited power due to a reduction in the number of sex-biased genes (Table S7, $P = 0.100$).

We also observed higher intersexual F_{ST} for both male- and female-biased genes (Fig. 2B). High intersexual F_{ST} can arise from sexual conflict in viability (Table 1, Scenario II), however, it can also be a consequence of sex-specific viability resulting from sex-specific genetic architecture (Table 1, Scenario III). These two scenarios can be distinguished using Tajima's D , where only loci subject to ongoing sexual conflict will exhibit a signature of high Tajima's D . Using this approach, we do not observe a significant excess of sex-biased genes with both high F_{ST} and high Tajima's

D (Scenario II, Table 2, $P = 0.560$) rather a significant excess with low Tajima's D and high F_{ST} (Scenario III, Table 2, $P = 0.020$). This pattern remains significant when we impose a significant P -value threshold for defining sex-biased genes (Scenario III, Table S7, $P = 0.020$). This suggests that sex differences in viability are not due to intralocus sexual antagonism, rather loci only affecting viability in one sex due to different genetic architecture.

Discussion

The mechanisms by which sexual conflict manifests within the genome have been the focus of considerable recent debate (Innocenti and Morrow 2010; Cheng and Kirkpatrick 2016; Lonn et al. 2017). We find that the majority of sexual conflict in guppies arises from differential fitness effects related to reproduction, rather than viability. Moreover, we observe a significant deficit of male-biased genes with high Tajima's D and low intersexual F_{ST} , suggesting that male-biased gene expression largely resolves sexual conflict arising from reproductive fitness. In contrast, although intersexual F_{ST} is significantly elevated for female-biased genes, we do not find significantly reduced Tajima's D . Male-biased genes across a number of species tend to be more tissue specific than unbiased or female-biased genes (Mank et al. 2008; Meisel 2011), and although male-biased genes expressed in the gonad tend to exhibit rapid rates of evolution (Parsch and Ellegren 2013; Wright and Mank 2013), this is not the case for male-biased genes expressed in the guppy tail, which instead show lower rates of evolution than female-biased and unbiased genes (Sharma et al. 2014). The evolutionary lability and reduction in pleiotropy (Harrison et al. 2015) associated with the evolution of male-biased expression may explain in part why we observe the association between male-bias, but not female-bias, and the resolution of sexual conflict.

It is important when measuring gene expression to select a tissue related to the phenotype of interest, in our case reproduction and viability. This is because gene expression in general and sex differences in expression in particular, vary greatly across the different regions of the body (Mank 2017). We deliberately selected the guppy tail for our gene expression analysis, as this somatic tissue combines genes involved in male coloration, which are known to influence reproductive fitness (Endler 1983; Houde and Endler 1990), as well as skin and lateral line cells directly interfacing with the environment, and which therefore influence viability. Furthermore, guppies are social animals and as such, suffer from increased transmission of pathogens and parasites compared to solitary animals. This is particularly important given the trade-offs between carotenoid-based sexually selected traits and immune function (Lozano 1994; Schantz et al. 1999; Pike et al. 2007; Tomášek et al. 2016). Therefore, the gene expression patterns of our tissue sample, compared to other tissue types,

Table 2. Observed and expected numbers of genes evolving under different types of sexual conflict.

Scenario	Pattern	Sex-biased Obs/Exp <i>P</i> -value	Male-biased Obs/Exp <i>P</i> -value	Female-biased Obs/Exp <i>P</i> -value	Unbiased Obs/Exp <i>P</i> -value
I.	Sexual conflict due to differences in reproductive fitness	37/53 <i>P</i> = 0.030	22/34 <i>P</i> = 0.040	15/19 <i>P</i> = 0.400	1067/1051 <i>P</i> = 0.620
II.	Sexual conflict due to differences in viability selection	61/57 <i>P</i> = 0.560	35/37 <i>P</i> = 0.790	26/20 <i>P</i> = 0.180	1121/1125 <i>P</i> = 0.900
III.	Sex-specific viability effects	62/47 <i>P</i> = 0.020	39/30 <i>P</i> = 0.100	23/16 <i>P</i> = 0.100	907/923 <i>P</i> = 0.610

Only autosomal genes are included in this analysis. Female-biased genes are defined as genes with \log_2 fold change < -1 , male-biased genes are defined as genes with \log_2 fold change > 1 . High Tajima's *D* was defined as > 0.893 (upper tertile of empirical distribution) and low Tajima's *D* was defined as < 0.272 (lower tertile) to account for the inferred population contraction within our population (Supporting Results). High F_{ST} was defined as > 0.047 (upper tertile) and low F_{ST} was defined as < -0.008 (lower tertile) (Supporting Results). We calculated the expected number of sex-biased genes for each scenario and used chi-squared tests to identify over- or underabundance of sex-biased genes across the three scenarios.

have the unique potential to distinguish the relative importance of reproductive fitness versus viability.

It is possible that the male coloration genes in our tissue sample may increase the association between male-biased expression and reproductive fitness. However, it remains unclear why we observe a strong association between sex-biased expression and reproductive fitness in our tissue sample, and other work, based on gonad expression that should be entirely associated with reproductive fitness (Cheng and Kirkpatrick 2016), appears to reveal a pattern of intersexual F_{ST} for sex-biased genes, consistent with either sex-specific or sexually antagonistic viability. Further studies are needed to explore whether the resolution of sexual conflict via the evolution of male-biased expression is a universal feature of regulatory evolution, or is unique to tissues related to male sexually selected traits.

We also observe a significant excess of both male- and female-biased genes with elevated F_{ST} , although the pattern is much less pronounced than what we observe for Tajima's *D*. Intersexual F_{ST} has previously been interpreted as evidence for ongoing sexual antagonism (Cheng and Kirkpatrick 2016; Lucotte et al. 2016), implying that viability is a major source of sexual conflict in animals. In contrast, our results suggest that viability is less important than reproductive fitness in sexual conflict. Intersexual F_{ST} can also be influenced by sex-differences in predation (Norrdahl and Korpimäki 1998) or dispersal (Trochet et al. 2016). Although it is not known how these forces have affected estimates of intersexual F_{ST} in previous work (Cheng and Kirkpatrick 2016; Lucotte et al. 2016), our use of a closed population eliminates effects of sex-biased dispersal and predation from our estimates. Interestingly, it has been shown that bright coloration increases predation pressures in natural guppy populations (Endler 1980; Godin and McDonough 2003) and that this predation is the basis of sexual conflict. It will be interesting to assess the relative

balance of Tajima's *D* and intersexual F_{ST} , in the wild, where we might predict that male-biased genes associated with coloration exhibit elevated levels of F_{ST} due to male predation.

More importantly, recent work identifying elevated intersexual F_{ST} (Cheng and Kirkpatrick 2016; Lucotte et al. 2016), did not assess the signature of balancing selection for the same genes. Therefore, it is not clear whether the signal of intersexual F_{ST} in these studies was due to conflict or sex-specific viability effects, and intersexual F_{ST} alone cannot differentiate these latter two forces. We used the same approach to investigate patterns of intersexual F_{ST} (Cheng and Kirkpatrick 2016), but now incorporate patterns of Tajima's *D* to differentiate the type of sexual conflict. Only by incorporating patterns of Tajima's *D* with measures of F_{ST} are we able to discern that these sex differences in viability are not due to intralocus sexual conflict, rather loci only affecting viability in one sex due to different genetic architecture. This pattern is consistent with increasing evidence that many loci exhibit sex-specific phenotypic effects (Randall et al. 2013; Karp et al. 2017). Sex-specific genetic architecture, which can result from sex differences in dominance (Barson et al. 2015), is a potential mechanism of resolving sexual conflict. This together with our finding that a deficit of male-biased genes are subject to sexual conflict over reproductive fitness, indicates that sex-biased expression in general, and perhaps male-biased expression in particular, is a rapid and effective route to resolve intralocus sexual conflict (Gilks et al. 2014; Karp et al. 2017).

Measures of intersexual F_{ST} and Tajima's *D* can be influenced by population dynamics. However, we do not think they likely contributed to the patterns we observe because our analysis is based on the empirical distribution of these statistics, effectively correcting for inbreeding and population structure across our population (Supporting Information). Furthermore, we tested for changes in population size across our population, which can

influence measures of Tajima's D, and controlled for the inferred population contraction within our population (Supporting Information). Additionally, our use of a closed population also eliminates effects of sex-biased migration and sex-biased predation, which could also create patterns of intersexual F_{ST} . Finally, we would not expect the effect of population dynamics on the measurement of F_{ST} and Tajima's D to vary systematically across unbiased and sex-biased genes across the genome.

It is important to note that multiple processes can influence these population genetic measures for any particular locus, thereby hampering efforts to identify specific loci with sexually antagonistic effects. Additionally, a priori knowledge about gene function is required to scan genomes for sexual conflict given the other potential sources of balancing selection. This noise may explain the low level of variance described by our models for F_{ST} and Tajima's D and sex-biased expression. It is difficult to know whether this is significantly different from previous work using the same approach (Cheng and Kirkpatrick 2016), which did not report the amount of variance explained by the best fit model. However, our categorical analyses show consistent support for both a significant reduction in Tajima's D for male-biased genes as well as elevated F_{ST} for both male- and female-biased genes. This indicates that our findings are robust, and that these measures can be employed successfully to scan the genome to contrast the magnitude and type of sexual conflict acting across different categories of genes (Flowers et al. 2010).

It is possible that different selective regimes acting on males and females, such as recurrent selection on male-biased genes (Ellegren and Parsch 2007), have the potential to generate differences in F_{ST} and Tajima's D between classes of sex-biased genes. However, these seem unlikely to have contributed to the patterns we observe, as male-biased genes do not exhibit higher rates of functional evolution in the guppy tail (Sharma et al. 2014).

Taken together, our results suggest that majority of sexual conflict is produced through conflicting selection over reproductive fitness, and that sexual conflict has the potential to maintain substantial levels of genetic diversity through balancing selection. More importantly, our results also suggest that evolution of sex-biased gene expression and sex-specific genetic architecture are effective routes to the resolution of sexual conflict.

Materials and Methods

GENOME ASSEMBLY and TRANSCRIPTOME ANNOTATION

We previously assembled a female *Poecilia reticulata* de novo genome based on two females (Wright et al. 2017) from our outbred laboratory population originally collected from the Quare River in Trinidad, and kept in captivity since 1998 (Kotrschal et al. 2013). We annotated the transcriptome by sequencing RNA from

eleven male and four female *P. reticulata* tails (Table S1). Detailed methods for the assembly are described elsewhere (Wright et al. 2017), and in the Supporting Information, and Illumina reads have been deposited in the NCBI SRA (PRJNA353986).

RNA-SEQ ANALYSIS

We sequenced RNA from eleven male and four female *P. reticulata* tails from our population of guppies (Kotrschal et al. 2013). Male guppies show a remarkable variety of coloration patterns in the wild and our male samples exhibit high phenotypic and transcriptomic diversity. We chose to use more male samples than females to mitigate concerns over differences in transcriptional variation between the sexes and the identification of sex-biased genes. Illumina reads have been deposited in the NCBI Short Read Archive (PRJNA353986). RNA was sequenced on an Illumina HiSeq 2500 at The Wellcome Trust Centre for Human Genetics, University of Oxford, resulting in on average 32 million 100 bp paired-end reads per sample (Table S1). RNA data were quality filtered using Trimmomatic (Lohse et al. 2012).

We mapped RNA-seq reads to the de novo genome assembly using HISAT2 v2.0.4 (Kim et al. 2015), suppressing unpaired and discordant alignments for paired reads and excluding reads from the SAM output that failed to align. StringTie v1.2.3 (Pertea et al. 2015) was used to quantify gene expression. Output GTF files were merged across samples, and ncRNA and lowly expressed genes were removed (genes were removed if they were expressed < 2 FPKM in fewer than half of the individuals of either sex, a threshold that also gives us high statistical power to determine sex-bias). Expression was normalized using TMM in EdgeR (Robinson et al. 2010) and RPKM estimated for each gene. A total of 13,306 genes located on scaffolds with positional information remained after filtering (Wright et al. 2017). Further details of the sequencing procedure are described elsewhere (Wright et al. 2017). To avoid pseudo replication arising from the process of gene annotation, we identified *Poecilia formosa* reciprocal orthologs using a reciprocal BLASTn 2.3.0 (Altschul et al. 1990) with a threshold e-value of 10 e-10 and minimum percentage identify of 30%. This ensures each gene is represented only once in the analysis and removes multiple fragments of the same gene. Total of 10,079 reciprocal orthologs were used for subsequent analyses (Wright et al. 2017).

CALCULATING TAJIMA'S D

SAM files were coordinate sorted using SAMtools v1.2 (Pertea et al. 2015), converted to BAM files and filtered using ANGSD (Korneliussen et al. 2014). Reads were removed if they did not uniquely map, had a flag ≥ 256 , had a mate that was not mapped or had a mapping quality below 20. Bases were filtered if base quality fell below 13 or there was data in less than four individuals. Mapping quality scores were adjusted for excessive

mismatches and quality scores were adjusted around indels to rule out false SNPs. For subsequent analyses, only reads mapped within genic regions defined in the merged and filtered GTF file were used.

We used ANGSD (Korneliussen et al. 2014) to estimate summary statistics as it implements methods to account to sequencing uncertainty and is appropriate for uneven sequencing depth associated with transcriptome data. ANGSD was first used to calculate sample allele frequency likelihoods at each site from genotype likelihoods calculated with the SAMtools model (Korneliussen et al. 2014). Next, in the absence of ancestral state information, the overall folded site frequency spectrum (SFS) for the population was estimated using ANGSD (Nielsen et al. 2012). Finally, genetic diversity indices, including allele frequency posterior probability and Tajima's D were computed using the site frequency spectrum as prior information.

CALCULATING INTERSEXUAL F_{ST}

F_{ST} was calculated using the same procedure and filtering criteria as Tajima's D , except that RNA-seq data were instead filtered to remove bases where we had data in less than half the individuals in males and females separately. This ensures we do not exclude sex-limited genes from the analysis. Additionally, the overall unfolded SFS for the population was estimated. Hudson's F_{ST} , which is less sensitive to small sample sizes (Bhatia et al. 2013), was estimated as implemented in ANGSD (Korneliussen et al. 2014).

MODEL SELECTION FOR THE RELATIONSHIP BETWEEN SEX-BIAS, TAJIMA'S D AND F_{ST}

We followed the approach taken by Cheng and Kirkpatrick (Cheng and Kirkpatrick 2016) to fit a parametric model to describe the relationship between Tajima's D or F_{ST} and sex-biased expression for autosomal genes. First, we regressed Tajima's D or F_{ST} and sex-biased expression using polynomials. The optimal polynomial degree was determined using the Akaike Information Criterion (AIC) (Burnham and Anderson 2003) in R (Team 2016) and likelihood ratio tests to assess significance of each model using `lrtest` function in the `lmtest` package in R (Zeileis and Hothorn 2002). Models within two AIC units of the model with the lowest AIC (or $P < 0.05$) were treated as one top model set, and the model with the fewest parameters was preferred.

DISTINGUISHING TYPES OF SEXUAL CONFLICT BY CONTRASTING TAJIMA'S D AND F_{ST}

We identified three scenarios of the types of sexual conflict acting across the autosomes (Table 1). We divided the distribution of Tajima's D and F_{ST} into three quantiles and used the upper and lower tertile of each as thresholds to define the different types of sexual conflict. In doing so, we control for the inferred population contraction within our population (Supporting Results).

High Tajima's D was defined as > 0.893 and low Tajima's D was defined as < 0.272 . High F_{ST} was defined as > 0.047 and low F_{ST} was defined as < -0.008 . We calculated the observed number of sex-biased genes in each of the three scenarios. We calculated the expected number of sex-biased genes for each scenario using the proportion of all genes in each scenario and the total number of male-biased, female-biased and unbiased genes. Chi-squared tests were used to identify over- or under abundance of sex-biased genes across the three scenarios.

For discrete analysis of sex-bias and population genetic parameters, we used standard (Harrison et al. 2015; Wright et al. 2015) fold-change thresholds to define female-biased (\log_2 male/female RPKM < -1) and male-biased (\log_2 male/female RPKM > 1). Male guppies show a remarkable diversity of coloration, and this means that our samples include substantial variability in male-gene expression related to coloration. This transcriptional variability hampers traditional methods of gene expression analysis based on hard significance thresholds. We therefore used random permutation tests, previously used to assess transcriptional variation in guppies (Ghalambor et al. 2015), and shown to be ideal to detect differential expression in transcriptome data (Slonim 2002), to evaluate differential expression significance for genes showing at least twofold expression differences between males and females. Specifically, using TMM normalized counts for each gene, we performed a generalized linear model (glm) using counts as dependent variable and sex as the only fixed effect. We then generated permuted datasets using the same linear model on 1000 datasets in which the individual ID labels were randomly reassigned to sample data. This produced empirical null distributions of expression variation against which to test the hypothesis of significant expression difference for each gene. Using the computed gene-specific tests statistics from the actual data, we assessed whether these fell within the extreme tails of the permuted values for that transcript ($P < 0.05$).

AUTHOR CONTRIBUTIONS

J.E.M. and A.E.W conceived of the study and designed the experiments. N.K and S.D.B collected samples used for sequencing in this study. A.E.W., M.F., C.R.C., N.I.B and F.G.V analysed the data. All authors contributed to writing the manuscript.

ACKNOWLEDGMENTS

This work was supported by a NERC Independent Research Fellowship to A.E.W, the European Research Council (grant agreements 260233 and 680951) to J.E.M., the Swedish Research Council (grant 2012–3624) to N.K., a Human Frontier Science Program Fellowship to M.F. and Marie Skłodowska-Curie (No 654699) and NSF Postdoctoral Research Fellowships (No 1523669) to N.I.B. J.E.M. also gratefully acknowledges support from a Royal Society Wolfson Merit Award. We thank Thorfinn Korneliussen and Emma Fox for help with ANGSD and ngsTools, and Pedro Almeida, Iulia Darolti, Jake Morris and Vicencio Oostra for helpful comments. The authors acknowledge the use of the UCL Legion High

Performance Computing Facility (Legion@UCL), and associated support services, in the completion of this work.

LITERATURE CITED

- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403–410.
- Barson, N. J., T. Aykanat, K. Hindar, M. Baranski, G. H. Bolstad, P. Fiske, C. Jacq, A. J. Jensen, S. E. Johnston, S. Karlsson, *et al.* 2015. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature* 528:405–408.
- Bhatia, G., N. Patterson, S. Sankararaman, and A. L. Price. 2013. Estimating and interpreting FST: the impact of rare variants. *Genome Res.* 23:1514–1521.
- Bonduriansky, R., and S. F. Chenoweth. 2009. Intralocus sexual conflict. *Trends Ecol. Evol.* 24:280–288.
- Burnham, K. P., and D. R. Anderson. 2003. Model selection and multimodel inference: a practical information-theoretic approach. Springer Science & Business Media, Berlin, Germany.
- Charlesworth, D. 2006. Balancing selection and its effects on sequences in nearby genome regions. *PLoS Genet.* 2:e64.
- Cheng, C., and M. Kirkpatrick. 2016. Sex-specific selection and sex-biased gene expression in humans and flies. *PLoS Genet.* 12:e1006170.
- Chippindale, A. K., J. R. Gibson, and W. R. Rice. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. *Proc. Natl. Acad. Sci.* 98:1671–1675.
- Corral-López, A., N. Bloch, A. Kotschal, W. van der Bijl, S. D. Buechel, J. E. Mank, and N. Kolm. 2017. Female brain size affects the assessment of male attractiveness during mate choice. *Science Advances* 3: e1601990.
- Dapper, A. L., and M. J. Wade. 2016. The evolution of sperm competition genes: the effect of mating system on levels of genetic variation within and between species. *Evolution* 70:502–511.
- Ellegren, H., and J. Parsch. 2007. The evolution of sex-biased genes and sex-biased gene expression. *Nat. Rev. Genet.* 8:689–698.
- Endler, J. A. 1980. Natural selection on color patterns in *Poecilia reticulata*. *Evolution* 34:76–91.
- Endler, J. A. 1983. Natural and sexual selection on color patterns in poeciliid fishes. *Environ. Biol. Fishes* 9:173–190.
- Flowers, J. M., S. I. Li, A. Stathos, G. Saxer, E. A. Ostrowski, D. C. Queller, J. E. Strassmann, and M. D. Purugganan. 2010. Variation, sex, and social cooperation: molecular population genetics of the social amoeba *Dictyostelium discoideum*. *PLoS Genet.* 6:e1001013.
- Ghalambor, C. K., K. L. Hoke, E. W. Ruell, E. K. Fischer, D. N. Reznick, and K. A. Hughes. 2015. Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. *Nature* 525:372–375.
- Ghosh, R., E. C. Andersen, J. A. Shapiro, J. P. Gerke, and L. Kruglyak. 2012. Natural variation in a chloride channel subunit confers avermectin resistance in *C. elegans*. *Science* 335:574–578.
- Gilks, W. P., J. K. Abbott, and E. H. Morrow. 2014. Sex differences in disease genetics: evidence, evolution, and detection. *Trends Genet.* 30:453–463.
- Godin, J.-G. J., and H. E. McDonough. 2003. Predator preference for brightly colored males in the guppy: a viability cost for a sexually selected trait. *Behav. Ecol.* 14:194–200.
- Guirao-Rico, S., A. Sánchez-Gracia, and D. Charlesworth. 2017. Sequence diversity patterns suggesting balancing selection in partially sex-linked genes of the plant *Silene latifolia* are not generated by demographic history or gene flow. *Mol. Ecol.* 26:1357–1370.
- Harrison, P. W., A. E. Wright, F. Zimmer, R. Dean, S. H. Montgomery, M. A. Pointer, and J. E. Mank. 2015. Sexual selection drives evolution and rapid turnover of male gene expression. *Proc. Natl. Acad. Sci.* 112:4393–4398.
- Hartl, D. L., and A. G. Clark. 2007. Principles of population genetics. *Écoscience* 14: 544–545.
- Hedrick, P. W. 2011. Population genetics of malaria resistance in humans. *Heredity* 107:283–304.
- Hollis, B., D. Houle, Z. Yan, T. J. Kawecki, and L. Keller. 2014. Evolution under monogamy feminizes gene expression in *Drosophila melanogaster*. *Nat. Commun.* 5:3482.
- Houde, A. E., and J. A. Endler. 1990. Correlated evolution of female mating preferences and male color patterns in the guppy *Poecilia reticulata*. *Science* 248:1405–1408.
- Houde, A. E., and J. A. Endler. 1990. Correlate evolution of female mating preferences and male color patterns in the guppy *Poecilia reticulata*. *Science* 248:1405–1408.
- Huang, Y., S. I. Wright, and A. F. Agrawal. 2014. Genome-wide patterns of genetic variation within and among alternative selective regimes. *PLoS Genet.* 10:e1004527.
- Immonen, E., R. R. Snook, and M. G. Ritchie. 2014. Mating system variation drives rapid evolution of the female transcriptome in *Drosophila pseudoobscura*. *Ecol. Evol.* 4:2186–2201.
- Innocenti, P., and E. H. Morrow. 2010. The sexually antagonistic genes of *Drosophila melanogaster*. *PLoS Biol.* 8:e1000335.
- Jordan, C. Y., and D. Charlesworth. 2012. The potential for sexually antagonistic polymorphism in different genome regions. *Evolution* 66:505–516.
- Karp, N. A., I. M. P. Consortium, and J. K. White. 2017. Prevalence of sexual dimorphism in mammalian phenotypic traits. *Nat. Commun.* 8:15475.
- Kim, D., B. Landmead, and S. L. Salzberg. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12:357–360.
- Kirkpatrick, M., and R. F. Guerrero. 2014. Signatures of sex-antagonistic selection on recombining sex chromosomes. *Genetics* 197:531–541.
- Korneliusson, T. S., A. Albrechtsen, and R. S. D. Nielsen. 2014. Analysis of next generation sequencing data. *BMC Bioinformatics* 15:356.
- Kotschal, A., B. Rogell, A. Bundsen, B. Svensson, S. Zajitschek, I. Brännström, S. Immler, A. A. Maklakov, and N. Kolm. 2013. Artificial selection on relative brain size in the guppy reveals costs and benefits of evolving a larger brain. *Curr. Biol.* 23:168–171.
- Lewontin, R. C., and J. Krakauer. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* 74:175–195.
- Lohse, M., A. M. Bolger, A. Nagel, A. R. Fernie, J. E. Lunn, M. Stitt, and B. Usadel. 2012. RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Res.* 40:W622–W627.
- Lonn, E., E. Koskela, T. Mappes, M. Mokkonen, A. M. Sims, and P. C. Watts. 2017. Balancing selection maintains polymorphisms at neurogenetic loci in field experiments. *Proc. Natl. Acad. Sci.* 114:3690–3695.
- Lozano, G. A. 1994. Carotenoids, parasites, and sexual selection. *Oikos* 70:309–311.
- Lucotte, E. A., R. Laurent, E. Heyer, L. Ségurel, and B. Toupance. 2016. Detection of allelic frequency differences between the sexes in humans: a signature of sexually antagonistic selection. *Genome Biol. Evol.* 8:1489–1500.
- Mank, J. E. 2017. The transcriptional architecture of phenotypic dimorphism. *Nat. Ecol. Evol.* 1:0006.
- Mank, J. E., L. Hultin-Rosenberg, M. Zwahlen, and H. Ellegren. 2008. Pleiotropic constraint hampers the resolution of sexual antagonism in vertebrate gene expression. *Am. Nat.* 171:35–43.
- Meisel, R. P. 2011. Towards a more nuanced understanding of the relationship between sex-biased gene expression and rates of protein-coding sequence evolution. *Mol. Biol. Evol.* 28:1893–1900.

- Mokkonen, M., H. Kokko, E. Koskela, J. Lehtonen, T. Mappes, H. Martiskainen, and S. C. Mills. 2011. Negative frequency-dependent selection of sexually antagonistic alleles in *Myodes glareolus*. *Science* 334:972–974.
- Nielsen, R., T. Korneliussen, A. Albrechtsen, Y. Li, and J. Wang. 2012. SNP calling, genotype calling, and sample allele frequency estimation from new-generation sequencing data. *PLoS One* 7, e37558.
- Norrdahl, K., and E. Korpimäki. 1998. Does mobility or sex of voles affect risk of predation by mammalian predators? *Ecology* 79:226–232.
- Otto, S. P., J. R. Pannell, C. L. Peichel, T. L. Ashman, D. Charlesworth, A. K. Chippindale, L. F. Delph, R. F. Guerrero, S. V. Scarpino, and B. F. McAllister. 2011. About PAR: The distinct evolutionary dynamics of the pseudoautosomal region. *Trends Genet.* 27:358–367.
- Parsch, J., and H. Ellegren. 2013. The evolutionary causes and consequences of sex-biased gene expression. *Nat. Rev. Genet.* 14:83–87.
- Pertea, M., G. M. Pertea, C. M. Antonescu, T. C. Chang, J. T. Mendell, and S. L. Salzberg. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33:290–295.
- Pike, T. W., J. D. Blount, J. Lindström, and N. B. Metcalfe. 2007. Availability of non-carotenoid antioxidants affects the expression of a carotenoid-based sexual ornament. *Biol. Lett.* 3:353–356.
- Pointer, M. A., P. W. Harrison, A. E. Wright, and J. E. Mank. 2013. Masculinization of gene expression is associated with exaggeration of male sexual dimorphism. *PLoS Genet.* 9:e1003697.
- Qiu, S., R. Bergero, and D. Charlesworth. 2013. Testing for the footprint of sexually antagonistic polymorphisms in the pseudoautosomal region of a plant sex chromosome pair. *Genetics* 194:663–672.
- Randall, J. C., T. W. Winkler, Z. Kutalik, S. I. Berndt, A. U. Jackson, K. L. Monda, T. O. Kilpeläinen, T. Esko, R. Mägi, S. Li, et al. 2013. Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. *PLoS Genet.* 9:e1003500.
- Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140.
- Schantz, T., S. Bensch, M. Grahn, D. Hasselquist, and H. Wittzell. 1999. Good genes, oxidative stress and condition—dependent sexual signals. *Proc. R. Soc. Lond. B Biol. Sci.* 266:1–12.
- Sharma, E., A. Künstner, B. A. Fraser, G. Zipprich, V. A. Kottler, S. R. Henz, D. Weigel, and C. Dreyer. 2014. Transcriptome assemblies for studying sex-biased gene expression in the guppy, *Poecilia reticulata*. *BMC Genomics* 15:400.
- Slate, J. 2005. Molecular evolution of the sheep prion protein gene. *Proc. R. Soc. B* 272:2371–2377.
- Slonim, D. K. 2002. From patterns to pathways: gene expression data analysis comes of age. *Nat. Genet.* 32:502–508.
- Stahl, E. A., G. Dwyer, R. Mauricio, M. Kreitman, and J. Bergelson. 1999. Dynamics of disease resistance polymorphism at the Rpm1 locus of *Arabidopsis*. *Nature* 400:667–671.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.
- Team, R. C. 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Tomášek, O., B. Gabrielová, P. Kačer, P. Maršík, J. Svobodová, K. Syslová, M. Vinkler, and T. Albrecht. 2016. Opposing effects of oxidative challenge and carotenoids on antioxidant status and condition-dependent sexual signalling. *Sci. Rep.* 6:23546.
- Trochet, A., E. A. Courtois, V. M. Stevens, M. Bague, A. Chaine, D. S. Schmeller, and J. Clobert. 2016. Evolution of sex-biased dispersal. *Q. Rev. Biol.* 91:297–320.
- Wright, A. E., I. Darolti, N. I. Bloch, V. Oostra, B. Sandkam, S. D. Buechel, N. Kolm, F. Breden, B. Vicoso, and J. E. Mank. 2017. Convergent recombination suppression suggests role of sexual selection in guppy sex chromosome formation. *Nat. Commun.* 8: 14251.
- Wright, A. E., P. W. Harrison, F. Zimmer, S. H. Montgomery, M. A. Pointer, and J. E. Mank. 2015. Variation in promiscuity and sexual selection drives avian rate of Faster-Z evolution. *Mol. Ecol.* 24:1218–1235.
- Wright, A. E., and J. E. Mank. 2013. The scope and strength of sex-specific selection in genome evolution. *J. Evol. Biol.* 26:1841–1853.
- Zeileis, A., and T. Hothorn. 2002. Diagnostic checking in regression relationships. *R News* 2.

Associate Editor: A. Gardner

Supporting Information

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

Table S1. RNA-seq information for each sample

Table S2. Model selection for relationship between Tajima's D and sex-bias for autosomal genes.

Table S3. Model selection for relationship between Tajima's D and sex-bias for autosomal genes excluding immunity genes.

Table S4. Model selection for relationship between FST and sex-bias for autosomal genes.

Table S5. Model selection for relationship between FST and sex-bias for autosomal genes excluding immunity genes.

Table S6. Median values of FST and Tajima's D for autosomal genes excluding immunity genes.

Table S7. Distinguishing scenarios of the types of sexual conflict acting across the autosomes.

Figure S1. Principal component analysis.

Figure S2. Histogram of Tajima's D and FST across autosomal genes.

Figure S3. Test for FST bias due to unequal coverage between sexes.

Figure S4. Test for FST bias due to unequal coverage between sexes.

Figure S5. Test for FST bias due to unequal coverage between sexes.

Figure S6. Test for Tajima's D bias due to unequal coverage between sexes.

Figure S7. Test for Tajima's D bias due to unequal coverage between sexes.

Figure S8. Test for Tajima's D bias due to unequal coverage between sexes.