



The search for sexually antagonistic genes: Practical insights from studies of local adaptation and statistical genomics

Filip Ruzicka,^{1,2}  Ludovic Dutoit,^{3,*} Peter Csuppon,^{4,5} Crispin Y. Jordan,⁶ Xiang-Yi Li,⁷  Colin Olito,⁸ Anna Runemark,⁸ Erik I. Svensson,⁸ Homa Papoli Yazdi,⁸ and Tim Connallon^{1,9} 

¹*School of Biological Sciences, Monash University, Clayton VIC 3800, Australia*

²*E-mail: Filip.Ruzicka@monash.edu*

³*Department of Zoology, University of Otago, Dunedin 9054, New Zealand*

⁴*Institute of Ecology and Environmental Sciences, UPEC, CNRS, IRD, INRA, Sorbonne Université, Paris 75252, France*

⁵*Center for Interdisciplinary Research in Biology, CNRS, Collège de France, PSL Research University, Paris 75231, France*

⁶*School of Biomedical Sciences, University of Edinburgh, Edinburgh EH8 9XD, United Kingdom*

⁷*Institute of Biology, University of Neuchâtel, Neuchâtel CH-2000, Switzerland*

⁸*Department of Biology, Lund University, Lund SE-22362, Sweden*

⁹*E-mail: tim.connallon@monash.edu*

Received April 28, 2020

Accepted July 28, 2020

Sexually antagonistic (SA) genetic variation—in which alleles favored in one sex are disfavored in the other—is predicted to be common and has been documented in several animal and plant populations, yet we currently know little about its pervasiveness among species or its population genetic basis. Recent applications of genomics in studies of SA genetic variation have highlighted considerable methodological challenges to the identification and characterization of SA genes, raising questions about the feasibility of genomic approaches for inferring SA selection. The related fields of local adaptation and statistical genomics have previously dealt with similar challenges, and lessons from these disciplines can therefore help overcome current difficulties in applying genomics to study SA genetic variation. Here, we integrate theoretical and analytical concepts from local adaptation and statistical genomics research—including F_{ST} and F_{IS} statistics, genome-wide association studies, pedigree analyses, reciprocal transplant studies, and evolve-and-resequence experiments—to evaluate methods for identifying SA genes and genome-wide signals of SA genetic variation. We begin by developing theoretical models for between-sex F_{ST} and F_{IS} , including explicit null distributions for each statistic, and using them to critically evaluate putative multilocus signals of sex-specific selection in previously published datasets. We then highlight new statistics that address some of the limitations of F_{ST} and F_{IS} , along with applications of more direct approaches for characterizing SA genetic variation, which incorporate explicit fitness measurements. We finish by presenting practical guidelines for the validation and evolutionary analysis of candidate SA genes and discussing promising empirical systems for future work.

KEY WORDS: Evolutionary genomics, fitness variation, F_{ST} , F_{IS} , genomic inference of natural selection, intralocus sexual conflict, local adaptation.

*Ludovic Dutoit is a co-first author, along with Filip Ruzicka

Impact Summary

Genome sequences carry a record of the evolutionary and demographic histories of natural populations. Research over the last two decades has dramatically improved our ability to detect genomic signals of adaptation by natural selection, including several widely-used methods for identifying genes underlying local adaptation and quantitative trait variation. Yet the application of these methods to identify sexually antagonistic (SA) genes—wherein variants that are adaptive for one sex are maladaptive for the other—remains underexplored, despite the potential importance of SA selection as a mechanism for maintaining genetic variation. Indeed, several lines of evidence suggest that SA genetic variation is common within animal and plant populations, underscoring the need for analytical methods that can reliably identify SA genes and genomic signals of SA genetic variation. Here, we integrate statistics and experimental designs that were originally developed within the fields of local adaptation and statistical genomics and apply them to contexts of sex-specific adaptation and SA genetic variation. First, we evaluate and extend statistical methods for identifying signals of SA variation from genome sequence data alone. We then apply these methods to reanalyze previously published data on allele frequency differences between sexes—a putative signal of SA selection. Second, we highlight more direct approaches for identifying SA genetic variation, which use experimental evolution and statistical associations between individual genetic variants and fitness. Third, we provide guidelines for the biological validation, evolutionary analysis, and interpretation of candidate SA polymorphisms. By building upon the strong methodological foundations of local adaptation and statistical genomics research, we provide a roadmap for rigorous analyses of genetic data in the context of sex-specific adaptation, thereby facilitating insights into the role and pervasiveness of SA variation in adaptive evolution.

A population's evolutionary capacity for adaptation hinges upon the nature and extent of the genetic variation it harbors (Fisher 1930). In simple environments where selection is uniform over time, across space, and among different classes of individuals within the population, adaptation may proceed by fixing unconditionally beneficial mutations and eliminating deleterious ones. Yet species exist in complex environments, where opportunities for adaptation can be limited by genetic trade-offs among traits and fitness components (Otto 2004; Gomulkiewicz and Houle 2009; Chevin 2013; Connallon and Hall 2018) or by gene flow and conflicting directional selection among habitats within a species' range (Kirkpatrick and Barton 1997; Lenormand 2002; Duputié et al. 2012). Such contexts

allow maladaptation to persist in spite of abundant genetic variation within the population (Walsh and Blows 2009).

“Sexually antagonistic” (SA) genetic variation—wherein alleles that are beneficial when expressed in one sex are harmful when expressed in the other—represents a particularly common form of genetic trade-off (Rice and Chippindale 2001; Bonduriansky and Chenoweth 2009; Van Doorn 2009). SA genetic variation arises from sex differences in selection (a.k.a. sex-specific selection) on traits that are genetically correlated between the sexes (Connallon and Clark 2014b), and may contribute substantially to fitness variation (Kidwell et al. 1977; Abbott 2011; Connallon and Clark 2014a; Olito et al. 2018) and maladaptation (Lande 1980; Matthews et al. 2019). Estimates of phenotypic selection suggest that sex differences in directional selection are common (Cox and Calsbeek 2009; Lewis et al. 2011; Gosden et al. 2012; Stearns et al. 2012; Morrissey 2016; De Lisle et al. 2018; Singh and Punzalan 2018), implying that many genetic variants affecting quantitative traits have SA effects on fitness. Likewise, estimates of genetic variation for fitness suggest that the genetic basis of female and male fitness components is partially discordant, with some multilocus genotypes conferring high fitness in one sex and low fitness in the other (Chippindale et al. 2001; reviewed in Connallon and Matthews 2019).

Although studies of sex-specific selection indicate that SA alleles contribute to fitness variation in several animal and plant populations (e.g., Chippindale et al. 2001; Fedorka and Mousseau 2004; Svensson et al. 2009; Delph et al. 2011; Mokkonen et al. 2011; Berger et al. 2014), the population genetic basis of this fitness variation is largely unknown, leaving several important questions unanswered. For example, what fraction of genetic variance for fitness is attributable to SA alleles versus other classes of genetic variation (e.g., deleterious mutations)? Is SA genetic variation attributable to many small-effect loci or to a few large-effect loci? Are SA polymorphisms maintained under balancing selection, or are they transient and primarily evolving via mutation, directional selection, and drift? Are SA alleles randomly distributed across the genome or are they enriched on certain chromosome types (e.g., sex chromosomes)? These questions are part of the broader debate about the genetic basis of fitness variation and the evolutionary forces that maintain it (Lewontin 1975; Charlesworth and Hughes 1999), which is one of the oldest in this field and perhaps the most difficult to resolve (e.g., Lewontin 1975, p. 23).

As in most areas of evolutionary biology, research on SA selection is increasingly drawing upon genomics. A few studies have identified candidate SA polymorphisms with large effects on traits related to fitness (Roberts et al. 2009; Barson et al. 2015; Rostant et al. 2015; VanKuren and Long 2018; Pearse et al. 2019), and this handful of SA loci almost certainly represents the tip of the iceberg (Ruzicka et al. 2019). Many other studies have

highlighted genomic patterns of expression or sequence diversity that could be indicative of sex-specific selection (e.g., Innocenti and Morrow 2010; reviewed in Kasimatis et al. 2017; Mank 2017; Rowe et al. 2018).

Although there is optimism that genomics will facilitate the study of sex-specific selection, we still face several challenges in applying genomic data to identify and characterize SA genetic variation. For example, some putative genomic signals of sex-specific selection, such as sex-biased gene expression, are ambiguous: at best, they may serve as indirect proxies of sex-specific selection, or at worst provide no information about contemporary selection on each sex (Kasimatis et al. 2017; Rowe et al. 2018). Allele frequency differences between sexes (e.g., between-sex F_{ST}) may represent more straightforward genomic signatures of sex-specific selection (e.g., Cheng and Kirkpatrick 2016; Lucotte et al. 2016; but see Bisseger et al. 2019; Kasimatis et al. 2020). Yet ambiguous null hypotheses for empirical estimates of between-sex F_{ST} , along with high statistical noise relative to biological signal in these estimates, raise questions about statistical power and the prevalence of false positives within such data (Kasimatis et al. 2019). In addition, we need to better understand the extent to which common pitfalls of genome sequence datasets—for example, mismapped reads (Tsai et al. 2019), sampling biases, hidden population structure, and effects of linkage and hitchhiking—yield artifactual signals of sex-specific processes, and thus questionable candidate SA genes.

The challenges in applying genomics to sex-specific selection have strong parallels in the fields of local adaptation and statistical genomics. Although the study of SA loci is still in its infancy, the fields of local adaptation and statistical genomics have already grappled intensively with many of the conceptual and methodological challenges that research on sex-specific selection now faces (Hoban et al. 2016; Visscher et al. 2017). For example, local adaptation research has long emphasized the importance of clear null models for distinguishing genes involved in local adaptation from false positives that simply reside in the tails of neutral null distributions (Lewontin and Krakauer 1973; Günther and Coop 2013; Whitlock and Lotterhos 2015; Lohse 2017). Similarly, statistical genomics researchers have repeatedly warned that hidden population structure in genome-wide association studies (GWAS) can lead to spurious conclusions about the genetic basis of quantitative traits, complex diseases, and the role of adaptation in population differentiation (Lander and Schork 1994; Price et al. 2010; Barton et al. 2019; Berg et al. 2019; Sohail et al. 2019). Lessons from local adaptation and statistical genomics research can therefore sharpen hypothesis framing, guide statistical methodology, and inform best practices for disentangling signal, noise, and artifacts in studies of sex-specific selection.

Here, by drawing insights from local adaptation and statistical genomics research, we present practical guidelines for population genomic analyses of sex-specific fitness variation. We first outline two statistics that can, in principle, provide *indirect* evidence of sex-specific fitness effects of genetic variation: between-sex F_{ST} , which is sensitive to sex differences in viability selection and some components of reproductive success, and F_{IS} , a measure of Hardy-Weinberg deviations in diploids, which is sensitive to sex differences in overall selection (i.e., cumulative effects of viability, fertility, fecundity, and mating competition). We develop theoretical null models for each metric, provide an overview of their sampling distributions and statistical power, and present a reanalysis of published F_{ST} data in light of our models. We also highlight complementary methods adapted from case-control GWAS to overcome some of the limitations of these metrics. Second, we evaluate several *direct* approaches for characterizing sex-specific genetic variation for fitness, which combine elements from quantitative genetics, reciprocal transplant studies, and experimental evolution. These direct approaches have been extensively employed to study the genetic basis of locally adapted phenotypes and quantitative traits, but rarely to identify SA loci. Third, we outline approaches for validating candidate genes and discuss best practices for the analysis and interpretation of their evolutionary histories.

Indirect Approaches for Identifying SA Genes

Estimating fitness under natural conditions is difficult, rendering approaches for identifying SA genes that rely on fitness measurements (i.e., direct approaches; see section “Direct Approaches for Identifying SA Genes”) unfeasible for many populations. Any widely applicable approach must therefore make use of *indirect* empirical signals of SA selection in genome sequence data, which can now be collected for virtually any species.

Two specific patterns of genome sequence variation could be indicative of contemporary SA selection, as emphasized by several recent studies (e.g., Cheng and Kirkpatrick 2016; Lucotte et al. 2016; Eyer et al. 2019; Kasimatis et al. 2019). First, because sex differences in selection during the life cycle are expected to generate allele frequency differences between breeding females and males (i.e., the members of each sex that contribute to offspring of the next generation; see Box 1), allele frequency differences between *samples* of females and males of a population could be indicative of sex differences in selection—including SA selection, sex-limited selection, or ongoing sexually concordant selection that differs in magnitude between the sexes. Ideally, inferences of sex-specific selection from allele frequency estimates should be based on samples of breeding adults that have

passed the filter of viability selection, sexual selection, and fertility/fecundity selection, although in practice genome sequences of random samples of adults are more readily obtainable and will reflect viability selection (Cheng and Kirkpatrick 2016; Kasimatis et al. 2019). Second, allele frequency differences between breeding females and males of a given generation elevate heterozygosity in offspring of the next generation relative to Hardy-Weinberg expectations (Kasimatis et al. 2019). Inflated heterozygosity in a large random sample of offspring could therefore reflect sex differences in viability selection, sexual selection, and/or fertility and fecundity selection during the previous generation.

SINGLE-LOCUS SIGNALS OF SEX-SPECIFIC SELECTION IN FIXATION INDICES

Fixation indices, which are widely applied in studies of population structure (Wright 1951), can be used to quantify allele frequency differences between sexes (F_{ST}) and elevations in heterozygosity in the offspring of a given generation (F_{IS}), each of which are predicted consequences of sex-specific selection (Boxes 1-2). Several studies have estimated F_{ST} between sexes using gene sequences sampled from adults (Cheng and Kirkpatrick 2016; Lucotte et al. 2016; Flanagan and Jones 2017; Wright et al., 2018, 2019; Bisseger et al. 2019; Vaux et al. 2019) or from breeding individuals (Dutoit et al. 2018), yet it remains unclear how much information about SA selection is contained within these estimates. A major problem, as recently emphasized by Kasimatis et al. (2019), is that the contribution of sex-specific selection to allele frequency differentiation between the sexes may often be weak compared to effects of sampling error in allele frequency estimates. Indeed, simulations presented in several studies (e.g., Cheng and Kirkpatrick 2016; Lucotte et al. 2016; Connallon and Hall 2018; Kasimatis et al. 2019) confirm that signals of SA selection in between-sex F_{ST} are swamped by sampling error in small population genomic datasets. Nevertheless, without clearly defined probability distributions for between-sex F_{ST} and related metrics of allele frequency differentiation, it remains difficult to evaluate whether sampling error, by itself, is sufficient to account for the empirical distributions of population genomic metrics that are putatively associated with sex-specific selection.

As we show in Appendix A (Supporting Information) (see Box 2), a null distribution for between-sex F_{ST} estimates at loci with no sex differences in selection conforms to a special case of Lewontin and Krakauer's (1973) classic null model for F_{ST} estimated between populations. An appealing feature of our two-sex null model is its insensitivity to some of the simplifying assumptions inherent in Lewontin and Krakauer's original model (i.e., that subpopulations are independent; see Nei and Maruyama 1975; Robertson 1975; Charlesworth 1998; Beaumont 2005; Whitlock and Lotterhos 2015), or issues arising from

genetic linkage (Charlesworth 1998), which do not affect the two-sex null distribution when F_{ST} is independently estimated per single nucleotide polymorphism (SNP), but can strongly impact the null for concatenated sequences (e.g., gene-wide F_{ST} estimates; see Booker et al. 2020; Appendix A [Supporting Information]). Another appealing feature of the null model for F_{ST} is its insensitivity to the distribution of allele frequencies in the population, provided SNPs with very low minor allele frequencies (MAFs) are excluded from the analysis (e.g., MAFs < 0.05 for small datasets; MAFs < 0.01 for large datasets; see Appendices A and D [Supporting Information]). This insensitivity to MAF under the null is unique to F_{ST} ; other metrics of allele frequency differentiation strongly covary with MAF, so that simulations based on the MAF distribution of the study population are required to generate genome-wide null predictions against which data can be compared (Appendix D [Supporting Information]; Figs. S6 and S7; see Berner 2019 and Bisseger et al. 2019 for examples).

By comparing the distribution of between-sex F_{ST} under the null (i.e., no sex differences in selection) with the corresponding distribution under SA selection (Box 2), we can formally evaluate the minimum strength of selection (s_{\min} , defined as the minimum cost, per sex, of inheriting the "wrong" SA allele) required for SA loci to reliably reside within the upper tail of the null distribution. For example, the 99th percentile for the null distribution is $\hat{F}_{ST(99\%)} \approx 3.32/n_H$, where n_H is the harmonic mean sample size of female- and male-derived gene sequences, and \hat{F}_{ST} refers to an estimate of F_{ST} (see Box 2; results are based on Nei's 1973 estimator for F_{ST} , which closely aligns with Wright's 1951 definition for F_{ST} between a pair of populations; see Appendix A [Supporting Information] for discussion of alternative F_{ST} estimators). Roughly 1% of F_{ST} estimates should fall above this threshold when there are no sex differences in selection. The probability that a SA locus resides within the tail of the null distribution depends on the allele frequencies at the locus, the strength of selection, and the sample size of individuals that are sequenced (Appendix A [Supporting Information]). In studies with large sample sizes (i.e., $n_H = 10^5$ or greater, as in some human genomic datasets; see Fig. 1), \hat{F}_{ST} for a SA locus with intermediate allele frequencies and a fitness effect of a few percent will reliably fall within the upper tail of the null distribution (Fig. 1; Appendix D [Supporting Information]). In contrast, studies where $n_H < 10^4$ require very strong selection ($s_{\min} > 0.05$) for SA loci to reliably reside within the upper tail of the null F_{ST} distribution (Fig. 1), and are unlikely to identify individual SA loci (i.e., significant F_{ST} outliers), even in cases where SA genetic polymorphism is common throughout the genome. Under scenarios of sex-limited selection, or sexually concordant selection that differs in magnitude between the sexes, sample sizes required to reliably identify true F_{ST} outliers must be even larger, as the degree of allele frequency differentiation under sex-limited

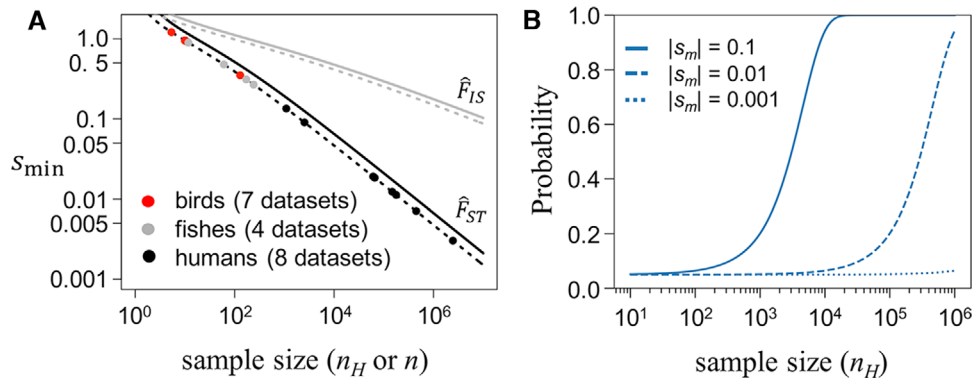


Figure 1. Signal of SA selection relative to sampling error in F_{ST} and F_{IS} estimates. (A) Lines show the minimum strength of selection (s_{\min} ; see Appendices A and B [Supporting Information]) that causes the expected values of \hat{F}_{ST} and \hat{F}_{IS} at an SA locus to reach the 5% tail (broken line) or the 1% tail (solid line) of the null distributions for \hat{F}_{ST} (black lines) and \hat{F}_{IS} (gray lines). Circles show the theoretical s_{\min} values (for the 5% tail) for published between-sex F_{ST} studies that include eight human datasets (Cheng and Kirkpatrick 2016; Lucotte et al. 2016; Kasimatis et al. 2020; Pirastu et al. 2020), four fish datasets (see Flanagan and Jones 2017; Wright et al., 2018, 2019; Bissegger et al. 2019; Vaux et al. 2019), and seven bird datasets (see Dutoit et al. 2018; Wright et al. 2019). The harmonic mean sample size, n_H , refers to the number of genes sequenced from diploid female and male samples in studies of between-sex F_{ST} ($n_H = 2(1/n_f + 1/n_m)^{-1}$, where $n_j =$ twice the number of individuals sampled from the j th sex); n refers to the sample size of offspring that are genotyped in studies of F_{IS} (see Box 2). (B) Probability that \hat{F}_{ST} for an additive SA locus (i.e., $h_f = h_m = \frac{1}{2}$) with intermediate equilibrium allele frequencies ($p, q = 1/2$) is in the top 5% tail of the null distribution for \hat{F}_{ST} . $|s_m|$ is the fitness cost of being homozygous for the “wrong” SA allele (see Appendices A and B [Supporting Information]). See Appendices C and D (Supporting Information) for further analyses of statistical power.

selection is roughly half the differentiation expected under SA selection, and differentiation is further muted under sexually concordant selection.

A second putative signal of sex-specific selection is an enrichment of heterozygotes among offspring cohorts, as inferred from high values of F_{IS} (defined in Box 2; Appendix B [Supporting Information]). Although only a single study has used estimates of F_{IS} (\hat{F}_{IS}) to test for sex-specific selection (Eyer et al. 2019; see Boxes 1-2), the potential for future applications warrants evaluation of signals of sex-specific selection using this metric. Hardy-Weinberg deviations in a sample, as captured by \hat{F}_{IS} , may arise from selection, nonrandom mating (e.g., inbreeding or population structure), or random sampling of genotypes from the population (Crow and Kimura 1970; Weir 1997; Lachance 2009). Statistical properties of Hardy-Weinberg deviations in genotype samples are well established (Weir 1997), and easily adaptable for our point of interest: the distribution of \hat{F}_{IS} in a randomly mating population in which allele frequencies may differ between the female and male parents of a given generation (Box 2). As illustrated in Box 2, the sampling variance for \hat{F}_{IS} exceeds that of \hat{F}_{ST} by a substantial margin. Consequently, \hat{F}_{IS} has far less power than \hat{F}_{ST} to distinguish signal of SA selection from noise (Fig. 1A), let alone distinguishing sex-limited selection or sexually concordant selection that differs in magnitude between the sexes. An additional issue is that signals of elevated heterozygosity are expected to be strongest

among cohorts sampled at *birth*, yet selection occurring during the life cycle can potentially decrease heterozygosity, further dampening signals of sex-specific selection in \hat{F}_{IS} estimates from adult samples.

F_{ST} DISTRIBUTIONS AND MULTI-LOCUS SIGNALS OF SEX-SPECIFIC SELECTION

Genome scans for individually significant SA loci (\hat{F}_{ST} or \hat{F}_{IS} outliers) are severely underpowered unless SA loci segregate for intermediate frequency alleles with large fitness effects and sample sizes are very large (see above). Indeed, no empirical F_{ST} study to date has yielded individually significant autosomal candidate SA SNPs that have survived corrections for multiple-testing and rigorous controls for genotyping error and read-mapping artifacts (see below).

Although F_{ST} scans for individually significant SA loci are highly conservative, the full empirical distribution of F_{ST} for autosomal SNPs may nonetheless carry a cumulative signature of sex-specific selection at many loci—even in the absence of individually significant SA genes. For example, SNPs responding to sex-specific selection (i.e., SNPs with sex-specific fitness effects or SNPs in linkage disequilibrium with them) should inflate average \hat{F}_{ST} and the proportion of observations in upper quantiles predicted by null models (Fig. 2). An excess of observations in the upper quantiles of the theoretical null may imply an enrichment of SNPs responding to sex-specific selection in the tail of the

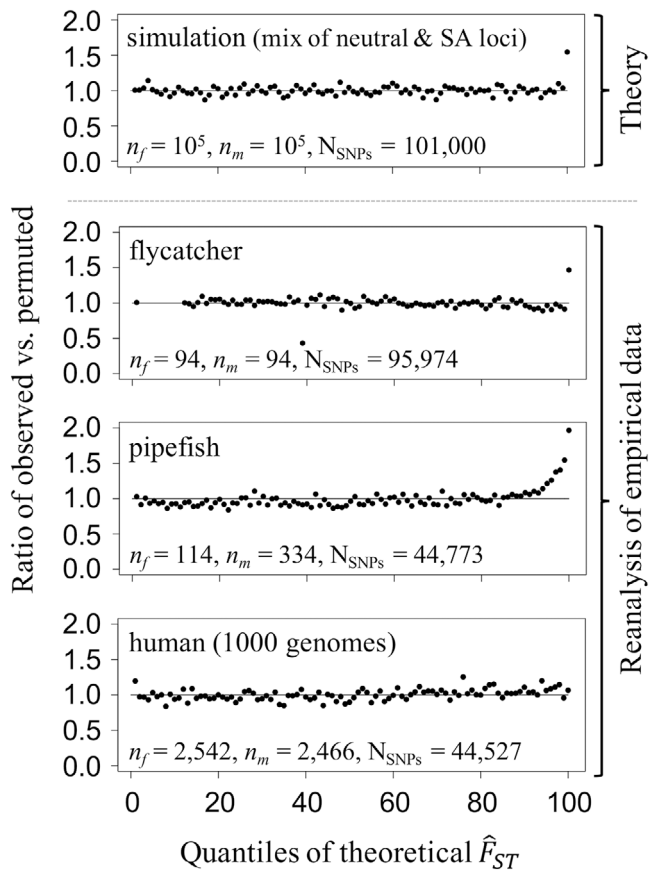


Figure 2. Reanalyses of genome-wide distributions of between-sex F_{ST} for three vertebrate species. Each panel shows the ratio of observed versus permuted \hat{F}_{ST} for SNPs within 100 quantiles of the theoretical null model of \hat{F}_{ST} . A ratio above one indicates an excess of observed \hat{F}_{ST} values in a given quantile. Theoretical data: \hat{F}_{ST} values were simulated for 10^5 neutrally evolving loci with MAF above 5% in the dataset, and for 10^3 loci responding to SA selection prior to sampling of gene sequences (each SA-responding locus had a SA fitness effect or was in perfect linkage disequilibrium with an SA locus). SA-responding loci had intermediate allele frequencies in the population ($p = \frac{1}{2}$). SA selection coefficients were drawn from an exponential distribution with mean of $s_{avg} = 0.03$, with additive fitness effects ($h = \frac{1}{2}$). The top 1% theoretical quantile is enriched by $\sim 50\%$, implying that $\sim 1/3$ of SNPs above the 99% threshold of the null model are true positives. Empirical data: population samples from flycatchers, pipefish, and humans (1000 Genomes data) (see Figs. S1-S4), excluding variants with $MAF < 0.05$; Fig. S5 illustrates the effect of including rare variants in the human reanalysis. For flycatcher data, the absence of values in some low- \hat{F}_{ST} quantiles is due to the small number of sampled sequences, which generates some discrepancy between the *discrete* distribution of \hat{F}_{ST} estimates (for both observed and permuted data) and the *continuous* theoretical null outlined in Box 2. The top 1% theoretical quantiles for flycatchers and pipefish are enriched by $\sim 50\%$ and $\sim 100\%$, respectively, implying that $\sim 1/3$ and $\sim 1/2$ of SNPs above the 99% threshold of the null are “true” positives (which still require biological validation and filtering for putative artefacts, as described in the main text). Code and data are available at https://github.com/ldutoit/male_female_fst.

empirical \hat{F}_{ST} distribution, with SNPs in the tail representing interesting candidates for follow-up analyses (see section “Validation and Follow-up Analyses of Candidate SA Genes”). Moreover, the fraction of true versus false positives among candidates can be quantified. For example, if 2% of observed SNPs fall within the top 1% quantile of the theoretical null, this implies a 1:1 ratio of true to false positives within the top 2% of observations (i.e., a false discovery rate of 50%).

To test for elevation of empirical \hat{F}_{ST} estimates relative to our theoretical null model, we reanalyzed three representative datasets from previously published studies (collared flycatcher *Ficedula albicollis*: Dutoit et al. 2018; gulf pipefish *Syngnathus scovelli*: Flanagan and Jones 2017; human: The 1000 Genomes Project Consortium 2015 data used by Cheng and Kirkpatrick 2016; Code available at https://github.com/ldutoit/male_female_fst). For human and flycatcher whole-genome resequencing datasets, we used autosomal coding variants, excluding any SNP with missing data. For the pipefish RAD-seq dataset, coding and noncoding variants with less than 50% missing data were included; sex-linked regions are unknown in this species and could not be excluded. In all datasets, polymorphic sites with MAFs below 5% were also excluded, as sites with low MAF exhibit inflated sampling variances (see Whitlock and Lotterhos 2015; Appendix A [Supporting Information] and Figs. S5-S7). Analyses were carried out in bedtools (Quinlan and Hall 2010), vcftools (Danecek et al. 2011), and R (R Core Team 2020).

For all three datasets, permuted \hat{F}_{ST} distributions (i.e., \hat{F}_{ST} calculated after randomly permuting sex labels across individuals) conform well to the theoretical null model for \hat{F}_{ST} (Figs. S1-S4). For the 1000 Genomes human dataset, \hat{F}_{ST} observations (i.e., nonpermuted estimates) are indistinguishable from both the theoretical null and permuted distributions, with no enrichment of observations within the top quantiles of the null (Fig. 2; χ^2 tests; 5% tail: $P = 0.18$; 1% tail: $P = 0.33$; comparison between the nonpermuted mean \hat{F}_{ST} and the \hat{F}_{ST} means for 1000 permutations of the data: $P > 0.05$). In contrast, flycatcher and pipefish datasets show elevated \hat{F}_{ST} values relative to the 5% and 1% tails of their null distributions (Fig. 2; χ^2 tests; $P = 4.33 \times 10^{-9}$ and $P < 2.2 \times 10^{-16}$ for flycatcher; $P < 2.2 \times 10^{-16}$ and $P < 2.2 \times 10^{-16}$ for pipefish), and exhibit inflated means for observed \hat{F}_{ST} relative to the \hat{F}_{ST} means of 1000 permutations of the data ($P < 0.05$ for both datasets). Such enrichment shows that sampling error is not sufficient to explain empirical distributions of \hat{F}_{ST} , and instead implies that many loci are responding to sex differences in selection (either directly or indirectly through hitchhiking with selected loci), or that a false signal of elevated \hat{F}_{ST} has been generated by population structure and/or data quality issues, as discussed below.

ACCOUNTING FOR SPURIOUS SIGNALS OF SEX-SPECIFIC SELECTION

The analyses presented above suggest that \hat{F}_{ST} , although severely underpowered for detecting individually significant outlier SNPs, may capture polygenic signals of sex-specific selection. However, sex differences in selection should not be invoked as the cause of such elevations in \hat{F}_{ST} without excluding artifacts that may generate similar patterns.

First, incorrect mapping of sex-linked markers to autosomes can potentially lead to artificial inflation of F_{ST} estimates. For example, Y- or W-linked sequences may be erroneously mapped to sequence paralogous on autosomes (Tsai et al. 2019), resulting in artificially high F_{ST} inferences at autosomal sites (Bissegger et al. 2019; Kasimatis et al. 2020). This problem can be mitigated in species with high-quality reference genomes, where mismapped reads can be eliminated through quality-filtering steps (i.e., removal of SNPs associated with low MAFs or extreme deviations from Hardy-Weinberg expectations), removal of candidate regions showing high sequence similarity to sex chromosome sequences (Kasimatis et al. 2020), and excluding regions with sex-biases in read coverage. However, mismapping is difficult to control for in species lacking high-quality reference genomes, including those where the sex determination system is unknown (e.g., the pipefish dataset in Fig. 2), or where sex chromosomes are young or rapidly evolving. Moreover, the effects of demographic processes, including recent admixture events or sex-biased migration, play out differently between sex chromosomes and autosomes (Hedrick 2007), so that caution is required in interpreting elevated between-sex F_{ST} on the X or Z, as has been reported in humans (e.g., Lucotte et al. 2016).

Second, sex differences in population structure—arising from the broad geographic sampling of individuals or recent migration into a single sampled population—can also generate signals of genetic differentiation between females and males in the absence of sex differences in selection (Box 1). Taxa with broad contemporary distributions (e.g., humans and *Drosophila*) often show significant genetic differentiation among populations. Uneven or unrepresentative sampling of individuals of each sex from a set of different locations can, by chance, inflate allele frequency differences between the sexes beyond expectations for a single, panmictic population. If loci showing high between-sex F_{ST} also exhibit high between-population F_{ST} , this could be indicative of population structure contributing to allele frequency divergence between the sexes in the empirical sample. Studies that sample individuals from a single population may also show artificially elevated between-sex F_{ST} if migration is sex biased (Box 1), which is common among animals (Trochet et al. 2016).

Sex-specific population structure can be accounted for by leveraging the statistical framework of case-control GWAS, in

which associations between polymorphic variants and binary phenotypic states are quantified (e.g., presence or absence of a disease). The case-control GWAS approach treats *sex* (female or male) as the binary phenotypic state and scans for loci with the strongest associations, which should exhibit elevated absolute odds ratios (see Appendix C [Supporting Information]; Kasimatis et al. 2020; Pirastu et al. 2020). Although the underlying logic is identical to between-sex F_{ST} , existing analytical methods for case-control GWAS can take population structure and relatedness in the empirical sample into account by including kinships (or the top principal components derived from kinships) as covariates (Astle and Balding 2009; Price et al. 2010). The case-control GWAS framework also permits estimation of SNP-based heritability of the phenotype (i.e., sex; Yang et al. 2011; Speed et al. 2017), which can be used to quantify a genome-wide signal of sex-specific selection.

Despite the advantages of leveraging an existing statistical framework, using case-control GWAS to test for associations between alternative alleles and sex does not sidestep all of the challenges faced by F_{ST} and F_{IS} statistics (see Appendices C and D [Supporting Information]). As with F_{ST} and F_{IS} , large sample sizes remain essential for discriminating between sampling variance and true signal of sex differences in selection (especially when selection is weak), and the methods perform poorly when MAFs are low. Additionally, association tests using odds ratios depend heavily on a normal approximation, and there is a deep and still evolving literature regarding hypothesis testing using these methods that users should be aware of (e.g., Haldane 1956; Wang and Shan 2015).

Direct Approaches for Identifying SA Genes

In exceptional study systems, candidate SA polymorphisms can be identified through explicit statistical associations between genotypes and fitness. Such *direct* inference approaches present two major advantages over indirect methods. First, the inclusion of fitness measurements can potentially increase power to detect individual SA loci, relative to indirect methods (e.g., Fig. 1). Second, association tests can be conducted across many components of fitness (e.g., viability, fecundity, and mating success), facilitating identification of the life history stages and selective contexts affected by SA loci. We outline two general approaches for direct inference of sex-specific selection—GWAS and evolve-and-resequence (E&R) studies—which have been extensively employed to identify genes associated with human trait variation and/or local adaptation (Long et al. 2015; Visscher et al. 2017), yet rarely to identify SA loci.

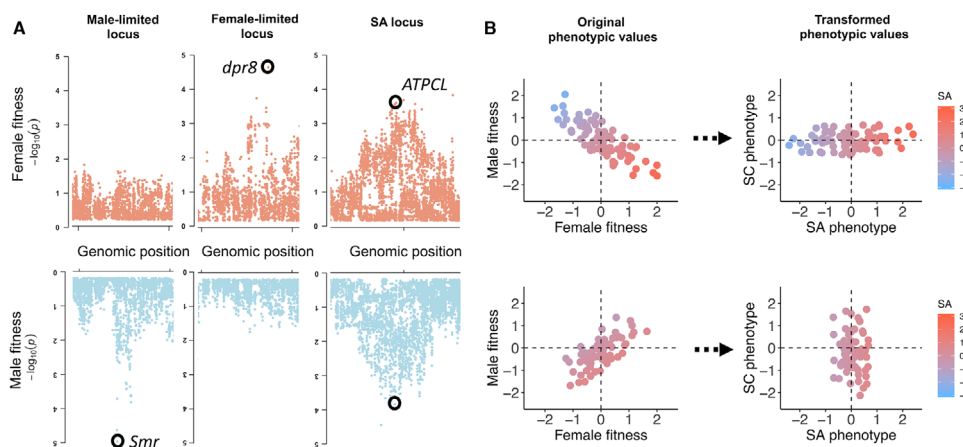


Figure 3. GWAS as a direct method for identifying SA loci. (A) Manhattan plots of $-\log_{10} P$ -values from a GWAS of female and male fitness (data from Ruzicka et al. 2019), illustrating distributions of P -values at or near candidate loci (circled) with male-limited (left), female-limited (middle), and SA (right) fitness effects. (B) Simulated phenotypic values of female and male fitness, before (left) and after (right) 45° rotation of the bivariate coordinate matrix to obtain sexually antagonistic (SA) and sexually concordant (SC) axes of phenotypic variation. Where fitness variation is predominantly SA (the top left of panel B), most variation is along the SA phenotypic axis (the top right of panel B). Where fitness variation is predominantly SC (the bottom left of panel B), most variation is along the SC phenotypic axis (the bottom right of panel B). In all panels, colors denote SA phenotypic values.

GWAS OF SEX-SPECIFIC FITNESS

GWAS quantify statistical associations between phenotypic variation and polymorphic SNPs throughout the genome. Using GWAS to identify SA loci further requires that data on fitness components and genotypes are collected from individuals of each sex. A major advantage of GWAS is the availability of statistically rigorous methods to identify candidate loci, including methods to control for covariates in analyses (Price et al. 2010), and approaches that correct for multiple testing (e.g., family-wise or false discovery rate correction; Benjamini and Hochberg 1995) or that reduce the number of tests through gene-based association analysis (Nagamine et al. 2012; Riggio et al. 2013; Bérénos et al. 2015). We discuss the application of GWAS to three dataset types: (i) datasets in which genotypes and phenotypes are each measured independently in each sex (e.g., humans); (ii) systems amenable to experimental manipulation, in which each genotype can be replicated among female and male carriers (e.g., isogenic or hemiclone fruit fly lines); and (iii) pedigreed populations, in which the genealogical relationships between all individuals are known (e.g., some sedentary vertebrate populations).

Where genotypic and phenotypic measurements are performed among independently sampled individuals of each sex, as in humans, SA loci can be identified by first performing a separate GWAS in each sex (“sex-stratified” GWAS) (Fig. 3A), and then quantifying the difference between male- and female-specific effect sizes (see also Gilks et al. 2014). Illustrating this approach, Winkler et al. (2015) performed a sex-stratified GWAS on several human anthropometric traits, then defined a t -statistic

as $t = \frac{\beta_M - \beta_F}{\sqrt{(SE_M^2 + SE_F^2 - 2\rho SE_M SE_F)}}$, where β_M and β_F are the sex-specific effect sizes, SE_M and SE_F are the sex-specific standard errors, and ρ is the between-sex rank correlation among genome-wide loci. For each polymorphic site, P -values were generated by comparing the observed t statistics to a null t -distribution with no sex-specific effects (where $E[t] = E[\beta_M - \beta_F] = 0$ under the null). This approach has been applied to nonfitness traits in humans (Randall et al. 2013; Myers et al. 2014; Winkler et al. 2015; Mitra et al. 2016; reviewed in Khramtsova et al. 2019), but has yet to be applied to fitness components (e.g., “number of children” phenotype in the UK Biobank; Sudlow et al. 2015).

In some experimental systems (e.g., fruit flies; flowering plants), the creation of isogenic or hemiclone lines (Abbott and Morrow 2011; Mackay et al. 2012; Berger et al. 2014) allows the same genotypes to be replicated and phenotypically assayed in carriers of each sex. Here, genotypes are effectively transplanted into male and female bodies or “environments,” analogous to the reciprocal transplantation of individuals sampled from different environments in local adaptation studies (Price et al. 2018). Identifying SA loci can then be achieved by transforming the bivariate coordinate system of male and female fitness values of a set of genotypes through matrix rotation (see Berger et al. 2014), which generates a univariate SA phenotype amenable to GWAS analysis (Fig. 3B). The approach is exemplified by a recent study in *D. melanogaster* (Ruzicka et al. 2019), which identified ~ 230 candidate SA polymorphic sites.

In pedigreed vertebrate populations, such as Soay sheep (*Ovis aries*) or Florida scrub jays (*Aphelocoma coerulescens*),

the genetic relationships between all individuals are known, and transmission of individual alleles across successive generations can be estimated (MacCluer et al. 1986). Because an individual's genetic contribution to future generations is a genuine representation of its Darwinian fitness, alleles transmitted more frequently by one sex relative to the other represent candidate SA variants. Analyses of pedigreed populations of Florida scrub jays have been used to identify alleles with above-average transmission rates to descendants irrespective of sex (i.e., unconditionally beneficial alleles; Chen et al. 2019), yet this type of GWAS remains to be used to identify SA loci. It should be noted, however, that many pedigreed populations are necessarily small (given the logistics of monitoring them), which may hinder detection of loci affecting fitness variation.

Although GWAS-based identification of candidate SA loci shows promise, two major drawbacks must be kept in mind. First, measurements that capture total lifetime reproductive success are difficult to obtain, and caution is required in interpreting results based on single fitness component (e.g., reproductive but not viability selection), which may correlate imperfectly with total fitness. Second, effect sizes are typically small for polygenic traits (Visscher et al. 2017), including fitness. Powerful GWAS of sex-specific fitness may therefore be logistically prohibitive, and candidate SA loci will necessarily represent the subset of loci with particularly large fitness effects.

E&R WITH SEX-LIMITED SELECTION

Experimental elimination of selection in one sex but not the other (i.e., sex-limited selection) is a powerful way to identify SA selection in action. Various sex-limited selection designs have been implemented, including (i) restricting transmission of the genome to the male line, and thereby removing selection through females (Rice, 1996, 1998; Prasad et al. 2007; Bedhomme et al. 2008; Abbott et al. 2010), or vice versa (Rice 1992), using *Drosophila* hemiclones; (ii) eliminating fitness variance in one sex (e.g., by enforcing random contributions to offspring number, or removing opportunity for mate choice) but not the other (Rundle et al. 2006; Morrow et al. 2008; Maklakov et al. 2009; Hollis et al. 2014; Immonen et al. 2014; Chenoweth et al. 2015; Veltsos et al. 2017); and (iii) applying sex-limited artificial selection on a specific fitness component, such as mating success (Dugand et al. 2019), lifespan (Berg and Maklakov 2012; Chen and Maklakov 2014; Berger et al. 2016), reproductive tactic (Bielak et al. 2014), or mating investment (Pick et al. 2017).

To identify SA loci, sex-limited selection can be combined with genotyping at multiple time points during experimental evolution within each selection regime (E&R), thereby connecting population genetic changes to the phenotypic responses accrued during experimental evolution. By tracking allele frequencies in male-limited and female-limited selection lines, alleles that show

a significant time-by-treatment interaction point to candidate SA loci, and their frequency dynamics can be characterized using current analytical tools for E&R experiments (Wiberg et al. 2017; Vlachos et al. 2019). E&R is a powerful and proven approach for identifying the genomic basis of phenotypic variation and local adaptation (Turner et al. 2011; Long et al. 2015; Barghi et al. 2019). Moreover, because it is experimental, the issues of sex-specific population structure that arise in between-sex F_{ST} studies (see section “Indirect Approaches for Identifying SA Genes”) can be minimized. Yet despite these advantages, we are not aware of any published study that has used E&R to identify SA loci (see Chenoweth et al. 2015 for the closest effort to date).

Resequencing can be performed using population samples taken at multiple time points within a single generation (Svensson et al. 2018) or across multiple generations, with the latter approach benefitting from the fact that allele frequency responses to selection are cumulative over multiple generations. E&R, like GWAS, remains best suited for detecting loci with relatively large fitness effects. Selection on complex polygenic traits typically leads to small changes in allele frequencies at large numbers of loci, resulting in genomic signals of selection that are difficult to distinguish from genetic drift (Schlötterer et al. 2015). Consequently, the study organism, the number of replicates, the effective population sizes of selection and control lines, and the duration of experiments must be carefully considered in the design of E&R experiments (Baldwin-Brown et al. 2014; Kofler and Schlötterer 2014; Kessner and Novembre 2015).

Validation and Follow-Up Analyses of Candidate SA Genes

Candidate SA genes and SNP sets enriched for SA alleles (i.e., identified using methods outlined above) provide context for addressing long-standing questions about SA variation, including the genomic distribution, biological functions, and population genetic processes shaping SA polymorphisms. We focus on two specific issues in follow-up analyses of putative SA variants. First, we outline approaches for biologically validating SA candidates—a crucial task given that candidate gene sets may include appreciable proportions of false positives and artifactual signals of sex-specific selection (see section “Indirect Approaches for Identifying SA Genes”). Second, we discuss population genetic analyses and issues of interpretation with bearing on the evolutionary histories of SA genes.

BIOLOGICAL VALIDATION OF SA GENES

Candidate SA loci can be directly validated in laboratory-amenable taxa by experimentally manipulating each allele and measuring its sex-specific fitness effect. A good example of experimental validation of naturally occurring SA genes is a

study by VanKuren and Long (2018), in which RNA interference and CRISPR-Cas9 were used to demonstrate SA effects of tandem duplicate genes *Apollo* and *Artemis* on offspring production in *D. melanogaster*. Similarly, Akagi and Charlesworth (2019) used manipulative molecular experiments to study candidate SA genes in several plant species. As a third example, several studies investigated a P450 transposable element insertion that up-regulates the *Cyp6g1* gene and increases DDT resistance in *D. melanogaster* (Smith et al. 2011; Rostant et al. 2015; Hawkes et al. 2016). Although evidence for SA effects at this particular locus is mixed, the experimental approaches used—including measurements of sex-specific fitness among isogenic lines and tracking the frequencies of each alternative allele in experimental cages—represent validation steps with potential for broad usage.

Direct experimental manipulation of candidate SA genes is not always feasible, and in instances where it is not, their biological validity can be assessed in other ways. In organisms harboring nonfunctional genomic material, it is possible to test whether candidate loci are enriched in genomic regions that are putatively functional (e.g., coding or regulatory) rather than inert (e.g., intergenic). Such “genic enrichment,” which is expected for SA polymorphisms with genuine phenotypic effects, has previously been used to strengthen validity of candidate alleles for local adaptation (Barreiro et al. 2008; Coop et al. 2009; Key et al. 2016). Another way to increase confidence in candidate loci is to look for multiple signals of SA selection. For example, candidates identified through elevated F_{ST} that are also associated with SA fitness effects in a GWAS represent the best candidates for follow-up evolutionary analyses (see below). Finally, if independent data exist on the sex-specific phenotypic effects of individual mutations (e.g., in RNAi databases), these data can be mined to support the validity of candidate SA genes.

EVOLUTIONARY DYNAMICS OF SA GENES

We do not outline the range of population genetic analyses that could be used to describe the evolutionary dynamics of candidate SA loci, as these have been comprehensively reviewed elsewhere (e.g., Vitti et al. 2013; Fijarczyk and Babik 2015). Instead we provide guidance on some common issues that are likely to arise when analyzing patterns of genetic variation at SA loci and interpreting their mode of evolution.

First, we emphasize that the evolutionary dynamics of a contemporary SA gene may have, in the past, been governed by any combination of genetic drift, net directional selection (selection favoring fixation of one SA allele), or balancing selection (selection maintaining SA polymorphism). Theory often focuses on the conditions generating balancing selection at SA loci (e.g., Kidwell et al. 1977; Patten and Haig 2009; Fry 2010), leading some empirical studies to use signals of balancing selection (e.g., elevated Tajima’s D) as indirect proxies for SA selection (e.g., Du-

toit et al. 2018; Wright et al., 2018, 2019; Sayadi et al. 2019). However, whether contemporary candidate SA alleles evolved under balancing selection hinges upon both the historical pattern of sex-specific selection and dominance at such loci (Kidwell et al. 1977; Connallon and Chenoweth 2019), which can be influenced by spatial and temporally varying selection (Connallon et al. 2019), and effective population size (Connallon and Clark, 2012, 2014a; Mullon et al. 2012). SA loci with large and symmetric selection coefficients or beneficial reversals of dominance (e.g., $h_f = 1$ and $h_m = 0$ in Box 1) are most conducive to balancing selection, whereas sufficient asymmetry between the sexes in the strength of selection (e.g., Mallet and Chippindale 2011; Mallet et al. 2011; Sharp and Agrawal 2013) should result in net directional selection that removes SA polymorphism rather than maintaining it (Kidwell et al. 1977; Kasimatis et al. 2019). Even when conditions for long-term balancing selection are met, the efficacy of balancing selection relative to drift may often be weak at SA loci, leading to genetic diversity patterns that are indistinguishable from neutrally evolving loci (Connallon and Clark, 2012, 2013; Mullon et al. 2012). In short, loci under contemporary SA selection can have a broad range of possible evolutionary histories. As such, the typical mode of evolution operating at candidate SA loci *cannot be assumed a priori* and should instead be viewed as a question that must be resolved empirically.

Second, the detection of elevated polymorphism at SA loci does not necessarily imply balancing selection. For example, SA candidate loci may exhibit significantly elevated MAFs relative to non-SA loci (Ruzicka et al. 2019), yet relaxed directional selection can account for this pattern if non-SA loci encompass a mix of neutral sites and sites evolving under sexually concordant directional selection. To establish that SA loci are evolving under balancing selection, it is necessary to show that SA genetic variation is significantly elevated compared to confirmed neutral sites (e.g., short introns, Parsch et al. 2010) and cannot be accounted for by demographic or mutational processes (Andrés et al. 2009; DeGiorgio et al. 2014; Bitarello et al. 2018). On the other hand, significant *reductions* in polymorphism at SA loci, relative to neutral sites, do not necessarily rule out balancing selection either. Counterintuitively, when the equilibrium frequency of the minor allele is low (i.e., equilibrium MAF < 0.2, approximately), balanced polymorphisms can be lost more rapidly than neutral polymorphisms, leading to reduced genetic variation relative to neutral expectations (Robertson 1962; Mullon et al. 2012).

A third and final point is that nonrandom patterns of genetic variation at SA loci can be generated by ascertainment bias alone. For example, data filtering steps that remove low-MAF SNPs (see “Indirect Approaches for Identifying SA Genes”) necessarily exclude rare SA variants from all downstream analyses. Elevated power to detect fitness effects among intermediate-frequency sites in a GWAS can also generate a spurious positive relationship

between candidate SA sites and MAF that might be mistaken for nonneutral evolution. Similarly, SA loci could be nonrandomly distributed across the genome (e.g., enriched in regions with low or high recombination), thereby generating spurious patterns in population genomic data that appear to indicate nonneutral evolution. It is therefore important to correct for such biases where possible by, for example, incorporating external data on recombination rate variation (Comeron et al. 2012; Elyashiv et al. 2016) or assessing evidence of trans-species polymorphisms among SA loci before and after removal of CpG sites (Leffler et al. 2013).

Moving Forward

We have critically assessed a broad range of methods for detecting genomic signatures of SA selection, including *indirect methods* based on genome sequence analysis (section “Indirect Approaches for Identifying SA Genes”) and *direct methods* based on associations between genome sequences and fitness measurements (section “Direct Approaches for Identifying SA Genes”). An inescapable conclusion from our indirect inference models is that very strong sex differences in selection or very large sample sizes are required to detect individual SA candidate polymorphisms with high confidence (Fig. 1; Appendices A-D [Supporting Information]), in agreement with previous simulation studies of between-sex F_{ST} (Lucotte et al. 2016; Connallon and Hall 2018; Kasimatis et al. 2019). Nevertheless, estimates of the full distribution of F_{ST} from previously published flycatcher and pipefish datasets reveal an intriguing elevation of genome-wide F_{ST} relative to our null models, which justifies future empirical studies of allele frequency differences between sexes (see below). Although an elevated signal of between-sex F_{ST} is not present in our reanalysis of human data, the absence of a signal is perhaps unsurprising given the small number of loci analyzed, following the removal of noncoding sequences and loci with rare polymorphisms. Genome-wide analyses of sex-specific selection in humans using larger datasets (Kasimatis et al. 2020), and examinations of sex-biased expression among sites with elevated F_{ST} (Cheng and Kirkpatrick 2016)—which were the focus of previous work, but not of our reanalysis—are therefore encouraged. With regard to direct methods for identifying SA genes, the substantial logistical challenge of accurately measuring fitness must be circumvented, yet the approach is powerful when feasible (see Ruzicka et al. 2019) and certain to be a key component of future work on the genetics of sex-specific fitness variation.

Although there is little doubt that identifying and characterizing SA genes is challenging, there are several reasons for optimism. First, the low power of indirect metrics to detect selection at an individual locus level does not rule out the detection of a cumulative signal of polygenic sex differences in selection (e.g.,

Fig. 2). Although such an approach implies that candidate SA genes (e.g., those in the highest F_{ST} quantiles) will include many false positives, elevated false discovery rates are not necessarily problematic if we are interested in the general properties of SA candidates relative to samples of putatively neutral (or non-SA) loci. Nevertheless, in studies with low-to-moderate sample sizes, where many candidate genes will be false positives, researchers should minimally demonstrate that (i) the empirical distribution of the metric of interest differs significantly from its appropriate null (see Kasimatis et al. 2019; our reanalyses in section “Indirect Approaches for Identifying SA Genes”), (ii) putative signals of selection are not driven by sex-specific population structure or other artefacts (see section “Indirect Approaches for Identifying SA Genes”), and (iii) candidate loci are situated in putatively functional genome regions (see section “Validation and Follow-up Analyses of Candidate SA Genes”).

Second, the power to detect SA genes using indirect metrics can often be increased in relatively simple ways. For example, pooled sequencing is a cost-effective method for estimating allele frequencies from samples of many individuals (Schlötterer et al. 2015), and well-suited for genome-wide F_{ST} scans (although not for F_{IS} scans, as estimating F_{IS} requires individual-level genotype data). Researchers could, alternatively, focus attention toward large publicly available genomic datasets that are adequately powered for detecting loci under moderately strong SA selection (see Fig. 1), or toward genomic regions predicted to have relatively high statistical power. For example, studies of pseudoautosomal regions of recombining sex chromosomes have substantially higher power to detect F_{ST} outliers driven by sex differences in selection (Qiu et al. 2013; Kirkpatrick and Guerrero 2014). Targeted sampling strategies may also amplify power to identify SA genes. For example, estimating allele frequencies among *breeding* adults—which have passed filters of viability selection and components of adult reproductive success—increases the number of episodes of selection that can contribute to allele frequency differentiation between sexes, improving the potential for detecting elevated between-sex F_{ST} .

Third, well-chosen study systems can improve prospects for accurately measuring lifetime reproductive success and identifying SA loci through direct methods (GWAS or E&R). For example, difficulties in accurately measuring fitness under field conditions can be mitigated in pedigreed populations, where the genetic contribution of each individual to successive generations is known (provided the population is well monitored), and each genotype can therefore be associated with an accurate estimate of total lifetime reproductive success in each sex. Emerging approaches to infer pedigrees from genomic data alone (Snyder-Mackler et al. 2016) may further facilitate identification of SA loci in the absence of long-term monitoring efforts. In some experimental systems, such as laboratory-adapted hemiclones of

Table 1. Sex-specific relative fitness of genotypes of a biallelic locus.

	Genotype		
	AA	Aa	aa
Female relative fitness	$1 + s_f$	$1 + s_f h_f$	1
Male relative fitness	$1 + s_m$	$1 + s_m h_m$	1

D. melanogaster (Rice et al. 2005; Abbott and Morrow 2011), relatively accurate measurements of outbred lifetime reproductive success are also possible. E&R is feasible for experimental organisms with short generation times and where large laboratory populations can be maintained (e.g., *Drosophila*, seed beetle *Callosobruchus maculatus*). Here, there is a relatively untapped opportunity to identify SA loci by combining sex-limited selection (e.g., Rice 1992; Prasad et al. 2007; Morrow et al. 2008; Abbott et al. 2010; Bonel et al. 2018) with genotyping across multiple generations (e.g., Turner et al. 2011; Long et al. 2015; Barghi et al. 2019; Abbott et al. 2020).

Finally, despite notable exceptions (e.g., the dioecious plant *Silene latifolia*; Delph et al. 2011; Muyle et al. 2012), plant systems remain underused in research on SA selection. One advantage of plants is their greater amenability to field measurements of fitness components, as widely used in studies of local adaptation and species' range limits (Hargreaves et al. 2014). Another advantage is the great diversity of reproductive systems in flowering plant species, the vast majority of which are hermaphroditic and susceptible to SA selection (Jordan and Connallon 2014; Tazzyman and Abbott 2015; Olito 2017; Olito et al. 2018), potentially leading to allele frequency differences between the pollen and ovules contributing to fertilization under haploid selection, and to elevated F_{IS} among offspring. A third advantage of plants is their greater tendency to express genetic variation during the haploid stage of their life cycle (e.g., Immler and Otto 2018). Haploid (relative to diploid) expression is expected to inflate the contribution of genetic polymorphism to fitness variance and magnify evolutionary responses to selection, including within-generation allele frequency divergence between sexes (Connallon and Jordan 2016). Exploiting plant systems may thereby increase statistical power to identify candidate SA genes or genomic signals of SA variation using direct (GWAS and E&R) or indirect inference approaches.

Box 1. Processes generating sexually divergent allele frequencies

Several evolutionary scenarios can lead to sex differences in the frequencies with which individual alleles are transmitted

to offspring (Hedrick 2007; Úbeda et al. 2011; Connallon et al. 2018). We focus on two processes—sex differences in selection and sex-biased migration—that may each commonly arise and affect estimates of allele frequency differences between sexes.

Sex differences in selection. Consider a single biallelic locus in which the focal allele (allele *A*) has a frequency of p at birth within a given generation; the alternative *a* allele has a frequency of $1 - p$. Selection during the life cycle alters the allele frequencies in the set of adults that contribute offspring to the next generation. The frequency of the *A* allele in breeding females and males (respectively) is follows:

$$p_f = p + s_f p (1 - p) (h_f + p (1 - 2h_f)) + O(s_f^2, s_m^2, s_f s_m)$$

and

$$p_m = p + s_m p (1 - p) (h_m + p (1 - 2h_m)) + O(s_f^2, s_m^2, s_f s_m),$$

where s_f and s_m are female and male selection coefficients for the *A* allele, h_f and h_m are the dominance coefficients (Table 1), and $O(s_f^2, s_m^2, s_f s_m)$ refers to second-order terms in the selection coefficients, which are negligible (and can be ignored) when s_f and s_m are small, as expected for most loci (Charlesworth and Charlesworth 2010, p. 97).

We expect allele frequency differences between breeding adults of each sex ($p_f \neq p_m$) when the fitness effects of each allele differ between sexes, that is, (1) the same allele is favored in each sex but the strength of selection differs between sexes (e.g., $p_f > p_m$ when $s_f > s_m > 0$); (2) alleles have sex-limited fitness effects (e.g., $p_f > p_m$ when $s_f > 0 = s_m$), or (3) alleles are sexually antagonistic (e.g., $p_f > p_m$ when $s_f > 0 > s_m$). Allele frequencies are expected to remain equal between sexes ($p_f = p_m$) when genetic variation is neutral ($s_f = s_m = 0$), or selection and dominance coefficients are identical between sexes ($s_f = s_m$ and $h_f = h_m$).

Sex-biased migration. Consider an island population receiving new migrants each generation, with migration occurring before reproduction during the life cycle. At birth, the frequencies of *A* and *a* alleles in the island population are p and $1 - p$, respectively. Let m_f and m_m represent the proportions of breeding females and breeding males on the island that are migrants. The expected frequency of the *A* allele in breeding females and males (respectively) will be

$$p_f = (1 - m_f) p + m_f \bar{p}$$

and

$$p_m = (1 - m_m) p + m_m \bar{p},$$

where \bar{p} is the frequency of the *A* allele in migrant individuals. The identity, $p_f - p_m = (m_f - m_m)(\bar{p} - p)$, implies that allele frequency differences between breeding females and males ($p_f \neq p_m$) would require sex-biased migration ($m_f \neq m_m$) and allele frequency differences between migrant and resident (nonmigrant) individuals ($\bar{p} \neq p$).

1. Box 2. Fixation indices (F_{ST} and F_{IS}) applied to sex differences

Allele frequency differences between breeding adults of each sex can be quantified by way of fixation indices, originally devised by Wright (1951) for characterizing genetic differentiation among populations. We again consider a biallelic autosomal locus with the focal allele (*A*) at a frequency of p_f in breeding females and p_m in breeding males.

Between-sex F_{ST} . F_{ST} is a standardized measure of the allele frequency difference between the sexes:

$$F_{ST} = \frac{(p_f - p_m)^2}{4\bar{p}(1 - \bar{p})},$$

where $\bar{p} = (p_f + p_m)/2$ (Cheng and Kirkpatrick 2016). This definition applies to the entire population, and therefore differs from empirical estimates of F_{ST} that are based upon samples of gene sequences from the population. Although there are several estimators of F_{ST} (see Appendix A [Supporting Information]; Bhatia et al. 2013; Gammerding et al. 2020), we focus on the simplest:

$$\hat{F}_{ST} = \frac{(\hat{p}_f - \hat{p}_m)^2}{4\hat{p}(1 - \hat{p})}$$

(Nei 1973), where \hat{p}_f and \hat{p}_m are the allele frequency estimates from samples of females and males, and $\hat{p} = (\hat{p}_f + \hat{p}_m)/2$. As we show in Appendix A (Supporting Information), the ratio

$$\frac{4n_f n_m \bar{p}(1 - \bar{p}) \hat{F}_{ST}}{n_m p_f (1 - p_f) + n_f p_m (1 - p_m)}$$

has an approximately noncentral chi-squared distribution with one degree of freedom and noncentrality parameter of $\lambda = (p_f - p_m)^2 \left(\frac{p_f(1-p_f)}{n_f} + \frac{p_m(1-p_m)}{n_m} \right)^{-1}$, where n_f and n_m are the numbers of sequences derived from females and males, respectively. The approximation can break down when n_f and n_m are small or the minor (rarer) allele at the locus has a frequency close to zero. Under the statistical null distribution, the true allele frequencies do not differ between the sexes ($p_f = p_m$), and therefore $\hat{F}_{ST} \approx X_0/2n_H$, where X_0 is a chi-squared

random variable with one degree of freedom, and $n_H = 2(1/n_f + 1/n_m)^{-1}$ is the harmonic mean sample size.

F_{IS} in offspring. F_{IS} can be used to quantify deviations between the observed heterozygosity in a cohort of individuals before selection (e.g., individuals sampled and sequenced at birth) and the expected heterozygosity under Hardy-Weinberg equilibrium. With random mating among breeding adults with female and male allele frequencies p_f and p_m , and ignoring effects of genetic drift or segregation distortion, the frequency of the *A* allele in offspring of the next generation will be $\bar{p} = (p_f + p_m)/2$, and the proportion that is heterozygous will be $P_{Aa} = p_f(1 - p_m) + p_m(1 - p_f)$. Under these conditions, F_{IS} will be

$$F_{IS} = \frac{P_{Aa}}{2\bar{p}(1 - \bar{p})} - 1 = \frac{(p_f - p_m)^2}{4\bar{p}(1 - \bar{p})},$$

where the final expression is equivalent to F_{ST} between breeding females and males of the prior generation (Kasimatis et al. 2019). The above expression for F_{IS} applies to the entire set of offspring in a population, whereas empirical estimates of F_{IS} will be based on the genotypes of offspring sampled from the population. As shown in Appendix B (Supporting Information), estimates of F_{IS} will be approximately normally distributed with a mean and variance of

$$E[\hat{F}_{IS}] = \frac{1}{2n} + \frac{(p_f - p_m)^2}{4\bar{p}(1 - \bar{p})}$$

and

$$\text{var}[\hat{F}_{IS}] = \frac{1}{n},$$

where n is the number of offspring genotyped for the locus. The approximation applies when the sample size is large. Under a null model, in which offspring are outbred and mating is random, estimates \hat{F}_{IS} will be normal with mean of $1/2n$ and variance of $1/n$.

ACKNOWLEDGMENTS

We thank A. Wright, H. Ellegren, and S. Flanagan for sharing data; S. Flanagan for support through pipefish data analysis; and I. Booksmythe, M. Kirkpatrick, and three anonymous reviewers for helpful comments on an earlier draft of this manuscript. The study was supported by funds from the Australian Research Council to TC and FR, the Swedish Research Council (VR: grant number 2016-03356) to EIS, and the Swiss National Science Foundation (Ambizione number 180145) to XYL. The New Zealand eScience Infrastructure (NeSI) provided support through their high-performance computing facilities; NeSI and collaborator institutions are funded through the Ministry of Business, Innovation and Employment's Research Infrastructure program (<https://www.nesi.org.nz>). We are particularly grateful to the European Society for Evolutionary Biology (ESEB) for funding the Special Topics Network workshops "Linking local adaptation with the evolution of sex differences," which motivated and enabled this study.

AUTHOR CONTRIBUTIONS

All authors contributed to the conceptualization of the manuscript. TC, CO, and PC developed mathematical models. LD conducted data reanalyses, with input from TC and FR. TC and FR coordinated manuscript writing, with initial drafts of "Direct approaches for identifying SA genes" by AR, XYL, ES, and FR, initial drafts of "Validation and follow-up analyses of candidate SA genes" by HPY and CYJ, and remaining sections drafted by TC and FR. All authors contributed to manuscript editing.

DATA ARCHIVING

All code for reproducing reanalyses presented in the manuscript is available at https://github.com/ldutoit/male_femalefst. Scripts and other data can be found on dryad: <https://doi.org/10.5061/dryad.b2rbnzsc>

LITERATURE CITED

- Abbott, J. K. 2011. Intra-locus sexual conflict and sexually antagonistic genetic variation in hermaphroditic animals. *Proc. R. Soc. B Biol. Sci.* 278:161–169.
- Abbott, J. K. and E. H. Morrow. 2011. Obtaining snapshots of genetic variation using hemiclinal analysis. *Trends Ecol. Evol.* 26:359–368.
- Abbott, J. K., S. Bedhomme, and A. K. Chippindale. 2010. Sexual conflict in wing size and shape in *Drosophila melanogaster*. *J. Evol. Biol.* 23:1989–1997.
- Abbott, J. K., A. K. Chippindale, and E. H. Morrow. 2020. The microevolutionary response to male-limited X-chromosome evolution in *Drosophila melanogaster* reflects macroevolutionary patterns. *J. Evol. Biol.* 33:738–750.
- Akagi, T. and D. Charlesworth. 2019. Pleiotropic effects of sex-determining genes in the evolution of dioecy in two plant species. *Proc. R. Soc. B Biol. Sci.* 286:20191805.
- Andrés, A. M., M. J. Hubisz, A. Indap, D. G. Torgerson, J. D. Degenhardt, et al. 2009. Targets of balancing selection in the human genome. *Mol. Biol. Evol.* 26:2755–2764.
- Astle, W. and D. J. Balding. 2009. Population structure and cryptic relatedness in genetic association studies. *Stat. Sci.* 24:451–471.
- Baldwin-Brown, J. G., A. D. Long, and K. R. Thornton. 2014. The power to detect quantitative trait loci using resequenced, experimentally evolved populations of diploid, sexual organisms. *Mol. Biol. Evol.* 31:1040–1055.
- Barghi, N., R. Tobler, V. Nolte, A. M. Jakšić, F. Mallard, K. A. Otte, et al. 2019. Genetic redundancy fuels polygenic adaptation in *Drosophila*. *PLoS Biol.* 17:e3000128.
- Barreiro, L. B., G. Laval, H. Quach, E. Patin, and L. Quintana-Murci. 2008. Natural selection has driven population differentiation in modern humans. *Nat. Genet.* 40:340–345.
- Barson, N. J., T. Aykanat, K. Hindar, M. Baranski, G. H. Bolstad, P. Fiske, et al. 2015. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature* 528:405–408.
- Barton, N., J. Hermisson, and M. Nordborg. 2019. Population genetics: why structure matters. *eLife* 8:e45380.
- Beaumont, M. A. 2005. Adaptation and speciation: what can F_{ST} tell us? *Trends Ecol. Evol.* 20:435–440.
- Bedhomme, S., N. G. Prasad, P. P. Jiang, and A. K. Chippindale. 2008. Reproductive behaviour evolves rapidly when intralocus sexual conflict is removed. *PLoS One* 3:e2187.
- Benjamini, Y. and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* 57:289–300.
- Béréanos, C., P. A. Ellis, J. G. Pilkington, S. H. Lee, J. Gratten, and J. M. Pemberton. 2015. Heterogeneity of genetic architecture of body size traits in a free-living population. *Mol. Ecol.* 24:1810–1830.
- Berg, E. C. and A. A. Maklakov. 2012. Sexes suffer from suboptimal lifespan because of genetic conflict in a seed beetle. *Proc. R. Soc. B Biol. Sci.* 279:4296–4302.
- Berg, J. J., A. Harpak, N. Sinnott-Armstrong, A. M. Joergensen, H. Mostafavi, Y. Field, et al. 2019. Reduced signal for polygenic adaptation of height in UK biobank. *eLife* 8:e39725.
- Berger, D., K. Grieshop, M. I. Lind, J. Goenaga, A. A. Maklakov, and G. Arnqvist. 2014. Intralocus sexual conflict and environmental stress. *Evolution* 68:2184–2196.
- Berger, D., T. You, M. R. Minano, K. Grieshop, M. I. Lind, G. Arnqvist, et al. 2016. Sexually antagonistic selection on genetic variation underlying both male and female same-sex sexual behavior. *BMC Evol. Biol.* 16:88.
- Berner, D. 2019. Allele frequency difference AFD—an intuitive alternative to F_{ST} for quantifying genetic population differentiation. *Genes* 10:308.
- Bhatia, G., N. Patterson, S. Sankararaman, and A. L. Price. 2013. Estimating and interpreting F_{ST} : the impact of rare variants. *Genome Res.* 23:1514–1521.
- Bielak, A. P., A. M. Skrzynecka, K. Miler, and J. Radwan. 2014. Selection for alternative male reproductive tactics alters intralocus sexual conflict. *Evolution* 68:2137–2144.
- Bissegger, M., T. G. Laurentino, M. Roesti, and D. Berner. 2019. Widespread intersex differentiation across the stickleback genome – the signature of sexually antagonistic selection? *Mol. Ecol.* 29:262–271.
- Bitarello, B. D., C. de Filippo, J. C. Teixeira, J. M. Schmidt, P. Kleinert, D. Meyer, et al. 2018. Signatures of long-term balancing selection in human genomes. *Genome Biol. Evol.* 10:939–955.
- Bonduriansky, R. and S. F. Chenoweth. 2009. Intralocus sexual conflict. *Trends Ecol. Evol.* 24:280–288.
- Bonel, N., E. Noël, T. Janicke, K. Sartori, E. Chapuis, A. Ségard, et al. 2018. Asymmetric evolutionary responses to sex-specific selection in a hermaphrodite. *Evolution* 72:2181–2201.
- Booker, T. R., S. Yeaman, and M. C. Whitlock. 2020. Variation in recombination rate affects detection of F_{ST} outliers under neutrality. *Mol. Ecol.* <https://doi.org/10.1111/mec.15501>.
- Charlesworth, B. 1998. Measures of divergence between populations and the effect of forces that reduce variability. *Mol. Biol. Evol.* 15:538–543.
- Charlesworth, B. and D. Charlesworth. 2010. Elements of evolutionary genetics. W. H. Freeman, New York.
- Charlesworth, B. and K. Hughes. 1999. The maintenance of genetic variation in life history traits. Pp. 369–391 in R. S. Singh and C. B. Krimbas, eds. *Evolutionary genetics: from molecules to morphology*. Vol. 1. Cambridge Univ. Press, Cambridge, U.K.
- Chen, H. Y. and A. A. Maklakov. 2014. Condition dependence of male mortality drives the evolution of sex differences in longevity. *Curr. Biol.* 24:2423–2427.
- Chen, N., I. Juric, E. J. Cosgrove, R. Bowman, J. W. Fitzpatrick, S. J. Schoech, et al. 2019. Allele frequency dynamics in a pedigreed natural population. *Proc. Natl. Acad. Sci. USA* 116:2158–2164.
- Cheng, C. and M. Kirkpatrick. 2016. Sex-specific selection and sex-biased gene expression in humans and flies. *PLoS Genet.* 12:e1006170.
- Chenoweth, S. F., N. C. Appleton, S. L. Allen, and H. D. Rundle. 2015. Genomic evidence that sexual selection impedes adaptation to a novel environment. *Curr. Biol.* 25:1860–1866.
- Chevin, L. M. 2013. Genetic constraints on adaptation to a changing environment. *Evolution* 67:708–721.

- Chippindale, A. K., J. R. Gibson, and W. R. Rice. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 98:1671–1675.
- Comeron, J. M., R. Ratnappan, and S. Bailin. 2012. The many landscapes of recombination in *Drosophila melanogaster*. *PLoS Genet.* 8:e1002905.
- Connallon, T. and A. G. Clark. 2012. A general population genetic framework for antagonistic selection that accounts for demography and recurrent mutation. *Genetics* 190:1477–1489.
- . 2013. Antagonistic versus nonantagonistic models of balancing selection: characterizing the relative timescales and hitchhiking effects of partial selective sweeps. *Evolution* 67:908–917.
- . 2014a. Balancing selection in species with separate sexes: insights from Fisher's geometric model. *Genetics* 197:991–1006.
- . 2014b. Evolutionary inevitability of sexual antagonism. *Proc. R. Soc. B Biol. Sci.* 281:2013–2123.
- Connallon, T. and C. Y. Jordan. 2016. Accumulation of deleterious mutations near sexually antagonistic genes. *G3* 6:2273–2284.
- Connallon, T. and G. Matthews. 2019. Cross-sex genetic correlations for fitness and fitness components: connecting theoretical predictions to empirical patterns. *Evol. Lett.* 3:254–262.
- Connallon, T. and M. D. Hall. 2018. Genetic constraints on adaptation: a theoretical primer for the genomics era. *Ann. N. Y. Acad. Sci.* 1422: 65–87.
- Connallon, T. and S. F. Chenoweth. 2019. Dominance reversals and the maintenance of genetic variation for fitness. *PLoS Biol.* 29:e3000118.
- Connallon, T., F. Débarre, and X. Y. Li. 2018. Linking local adaptation with the evolution of sex differences. *Philos. Trans. R. Soc. B Biol. Sci.* 373:20170414.
- Connallon, T., S. Sharma, and C. Olito. 2019. Evolutionary consequences of sex-specific selection in variable environments: four simple models reveal diverse evolutionary outcomes. *Am. Nat.* 193:93–105.
- Coop, G., J. K. Pickrell, J. Novembre, S. Kudaravalli, J. Li, D. Absher, et al. 2009. The role of geography in human adaptation. *PLoS Genet.* 5:e1000500.
- Cox, R. M. and R. Calsbeek. 2009. Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. *Am. Nat.* 173:176–187.
- Crow, J. F. and M. Kimura. 1970. *An introduction to population genetics theory*. Harper and Row, Manhattan, NY.
- Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, et al. 2011. The variant call format and VCFtools. *Bioinformatics* 27:2156–2158.
- De Lisle, S. P., D. Goedert, A. M. Reedy, and E. I. Svensson. 2018. Climatic factors and species range position predict sexually antagonistic selection across taxa. *Philos. Trans. R. Soc. B Biol. Sci.* 373:20170415.
- DeGiorgio, M., K. E. Lohmueller, and R. Nielsen. 2014. A model-based approach for identifying signatures of ancient balancing selection in genetic data. *PLoS Genet.* 10:e1004561.
- Delph, L. F., J. Andicoechea, J. C. Steven, C. R. Herlihy, S. V. Scarpino, and D. L. Bell. 2011. Environment-dependent intralocus sexual conflict in a dioecious plant. *New Phytol.* 192:542–552.
- Dugand, R. J., J. L. Tomkins, and W. J. Kennington. 2019. Molecular evidence supports a genic capture resolution of the lek paradox. *Nat. Commun.* 10:1359.
- Duputié, A., F. Massol, I. Chuine, M. Kirkpatrick, and O. Ronce. 2012. How do genetic correlations affect species range shifts in a changing environment? *Ecol. Lett.* 15:251–259.
- Dutoit, L., C. F. Mugal, P. Bolívar, M. Wang, K. Nadachowska-Brzyska, L. Smeds, et al. 2018. Sex-biased gene expression, sexual antagonism and levels of genetic diversity in the collared flycatcher (*Ficedula albicollis*) genome. *Mol. Ecol.* 27:3572–3581.
- Elyashiv, E., S. Sattath, T. T. Hu, A. Strutsovsky, G. McVicker, P. Andolfatto, et al. 2016. A genomic map of the effects of linked selection in *Drosophila*. *PLoS Genet.* 12:e1006130.
- Eyer, P. A., A. J. Blumenfeld, and E. L. Vargo. 2019. Sexually antagonistic selection promotes genetic divergence between males and females in an ant. *Proc. Natl. Acad. Sci. USA* 116:24157–24163.
- Fedorka, K. M. and T. A. Mousseau. 2004. Female mating bias results in conflicting sex-specific offspring fitness. *Nature* 429:65–67.
- Fijarczyk, A. and W. Babik. 2015. Detecting balancing selection in genomes: limits and prospects. *Mol. Ecol.* 24:3529–3545.
- Fisher, R. A. 1930. *The genetical theory of natural selection*. Clarendon Press, Oxford, U.K.
- Flanagan, S. P. and A. G. Jones. 2017. Genome-wide selection components analysis in a fish with male pregnancy. *Evolution* 71:1096–1105.
- Fry, J. D. 2010. The genomic location of sexually antagonistic variation: some cautionary comments. *Evolution* 64:1510–1516.
- Gammerding, W. J., M. A. Toups, and B. Vicoso. 2020. Disagreement in FST estimators: a case study from sex chromosomes. *Mol. Ecol. Resour.* <https://doi.org/10.1111/1755-0998.13210>.
- 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.
- Gomulkiewicz, R. and D. Houle. 2009. Demographic and genetic constraints on evolution. *Am. Nat.* 174:E218–E229.
- Gosden, T. P., K. L. Shastri, P. Innocenti, and S. F. Chenoweth. 2012. The B-matrix harbors significant and sex-specific constraints on the evolution of multicharacter sexual dimorphism. *Evolution* 66:2106–2116.
- Günther, T. and G. Coop. 2013. Robust identification of local adaptation from allele frequencies. *Genetics* 195:205–220.
- Haldane, J. B. S. 1956. The estimation and significance of the logarithm of a ratio of frequencies. *Ann. Hum. Genet.* 20:309–311.
- Hargreaves, A. L., K. E. Samis, and C. G. Eckert. 2014. Are species' range limits simply niche limits writ large? A review of transplant experiments beyond the range. *Am. Nat.* 183:157–173.
- Hawkes, M. F., C. E. Gamble, E. C. R. Turner, M. R. Carey, N. Wedell, and D. J. Hosken. 2016. Intralocus sexual conflict and insecticide resistance. *Proc. R. Soc. B Biol. Sci.* 283:20161429.
- Hedrick, P. W. 2007. Sex: differences in mutation, recombination, selection, gene flow, and genetic drift. *Evolution* 61:2750–2771.
- Hoban, S., J. L. Kelley, K. E. Lotterhos, M. F. Antolin, G. Bradburd, D. B. Lowry, et al. 2016. Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *Am. Nat.* 188:379–397.
- 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.
- Immler, S. and S. P. Otto. 2018. The evolutionary consequences of selection at the haploid gametic stage. *Am. Nat.* 192:241–249.
- 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 T. Connallon. 2014. Sexually antagonistic polymorphism in simultaneous hermaphrodites. *Evolution* 68:3555–3569.
- Kasimatis, K. R., T. C. Nelson, and P. C. Phillips. 2017. Genomic signatures of sexual conflict. *J. Hered.* 108:780–790.
- Kasimatis, K. R., P. L. Ralph, and P. C. Phillips. 2019. Limits to genomic divergence under sexually antagonistic selection. *G3* 9:3813–3824.
- Kasimatis, K. R., A. Abraham, P. L. Ralph, A. D. Kern, J. A. Capra, and P. C. Phillips. 2020. Sexually antagonistic selection on genetic variation is rare in humans. *bioRxiv*. <https://doi.org/10.1101/2020.03.26.009670>

- Kessner, D. and J. Novembre. 2015. Power analysis of artificial selection experiments using efficient whole genome simulation of quantitative traits. *Genetics* 199:991–1005.
- Key, F. M., Q. Fu, F. Romagne, M. Lachmann, and A. M. Andres. 2016. Human adaptation and population differentiation in the light of ancient genomes. *Nat. Commun.* 7:10775.
- Khrantsova, E. A., L. K. Davis, and B. E. Stranger. 2019. The role of sex in the genomics of human complex traits. *Nat. Rev. Genet.* 20:173–190.
- Kidwell, J. F., M. T. Clegg, F. M. Stewart, and T. Prout. 1977. Regions of stable equilibria for models of differential selection in the two sexes under random mating. *Genetics* 85:171–183.
- Kirkpatrick, M. and N. H. Barton. 1997. Evolution of a species' range. *Am. Nat.* 150:1–23.
- Kirkpatrick, M. and R. F. Guerrero. 2014. Signatures of sex-antagonistic selection on recombining sex chromosomes. *Genetics* 197:531–541.
- Kofler, R. and C. Schlötterer. 2014. A guide for the design of evolve and resequencing studies. *Mol. Biol. Evol.* 31:474–483.
- Lachance, J. 2009. Detecting selection-induced departures from Hardy-Weinberg proportions. *Genet. Sel. Evol.* 41:15.
- Lande, R. 1980. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* 34:292–305.
- Lander, E. S. and N. J. Schork. 1994. Genetic dissection of complex traits. *Science* 265:2037–2048.
- Leffler, E. M., Z. Gao, S. Pfeifer, L. Ségurel, A. Auton, O. Venn, et al. 2013. Multiple instances of ancient balancing selection shared between humans and chimpanzees. *Science* 339:1578–1582.
- Lenormand, T. 2002. Gene flow and the limits to natural selection. *Trends Ecol. Evol.* 17:183–189.
- Lewis, Z., N. Wedell, and J. Hunt. 2011. Evidence for strong intralocus sexual conflict in the Indian meal moth, *Plodia interpunctella*. *Evolution* 65:2085–2097.
- Lewontin, R. C. 1975. The genetic basis of evolutionary change. Columbia Univ. Press, New York.
- 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, K. 2017. Come on feel the noise – from metaphors to null models. *J. Evol. Biol.* 30:1506–1508.
- Long, A., G. Liti, A. Luptak, and O. Tenaillon. 2015. Elucidating the molecular architecture of adaptation via evolve and resequencing experiments. *Nat. Rev. Genet.* 16:567–582.
- 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.
- MacCluer, J. W., J. L. VandeBerg, B. Read, and O. A. Ryder. 1986. Pedigree analysis by computer simulation. *Zoo Biol.* 5:147–160.
- Mackay, T. F. C., S. Richards, E. A. Stone, A. Barbadilla, J. F. Ayroles, D. Zhu, et al. 2012. The *Drosophila melanogaster* genetic reference panel. *Nature* 482:173–178.
- Maklakov, A. A., R. Bonduriansky, and R. C. Brooks. 2009. Sex differences, sexual selection, and ageing: an experimental evolution approach. *Evolution* 63:2491–2503.
- Mallet, M. A. and A. K. Chippindale. 2011. Inbreeding reveals stronger net selection on *Drosophila melanogaster* males: implications for mutation load and the fitness of sexual females. *Heredity* 106:994–1002.
- Mallet, M. A., J. M. Bouchard, C. M. Kimber, and A. K. Chippindale. 2011. Experimental mutation-accumulation on the X chromosome of *Drosophila melanogaster* reveals stronger selection on males than females. *BMC Evol. Biol.* 11:156.
- Mank, J. E. 2017. Population genetics of sexual conflict in the genomic era. *Nat. Rev. Genet.* 18:721–730.
- Matthews, G., S. Hangartner, D. G. Chapple, and T. Connallon. 2019. Quantifying maladaptation during the evolution of sexual dimorphism. *Proc. R. Soc. B Biol. Sci.* 286:20191372.
- Mitra, I., K. Tsang, C. Ladd-Acosta, L. A. Croen, K. A. Aldinger, R. L. Hendren, et al. 2016. Pleiotropic mechanisms indicated for sex differences in autism. *PLoS Genet.* 12:e1006425.
- Mokkonen, M., H. Kokko, E. Koskela, J. Lehtonen, T. Mappes, H. Martiskainen, et al. 2011. Negative frequency-dependent selection of sexually antagonistic alleles in *Myodes glareolus*. *Science* 334:972–974.
- Morrissey, M. B. 2016. Meta-analysis of magnitudes, differences and variation in evolutionary parameters. *J. Evol. Biol.* 29:1882–1904.
- Morrow, E. H., A. D. Stewart, and W. R. Rice. 2008. Assessing the extent of genome-wide intralocus sexual conflict via experimentally enforced gender-limited selection. *J. Evol. Biol.* 21:1046–1054.
- Mullon, C., A. Pomiankowski, and M. Reuter. 2012. The effects of selection and genetic drift on the genomic distribution of sexually antagonistic alleles. *Evolution* 66:3743–3753.
- Muyle, A., N. Zemp, C. Deschamps, S. Mousset, A. Widmer, and G. A. B. Marais. 2012. Rapid de novo evolution of X chromosome dosage compensation in *Silene latifolia*, a plant with young sex chromosomes. *PLoS Biol.* 10:e1001308.
- Myers, R. A., N. M. Scott, W. J. Gauderman, W. Qiu, R. A. Mathias, I. Romieu, et al. 2014. Genome-wide interaction studies reveal sex-specific asthma risk alleles. *Hum. Mol. Genet.* 23:5251–5259.
- Nagamine, Y., R. Pong-Wong, P. Navarro, V. Vitart, C. Hayward, I. Rudan, et al. 2012. Localising loci underlying complex trait variation using Regional Genomic Relationship Mapping. *PLoS One* 7:e46501.
- Nei, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* 70:3321–3323.
- Nei, M. and T. Maruyama. 1975. Letters to the editors: Lewontin-Krakauer test for neutral genes. *Genetics* 80:395.
- Olito, C. 2017. Consequences of genetic linkage for the maintenance of sexually antagonistic polymorphism in hermaphrodites. *Evolution* 71:458–464.
- Olito, C., J. K. Abbott, and C. Y. Jordan. 2018. The interaction between sex-specific selection and local adaptation in species without separate sexes. *Philos. Trans. R. Soc. B Biol. Sci.* 373:20170426.
- Otto, S. P. 2004. Two steps forward, one step back: the pleiotropic effects of favoured alleles. *Proc. R. Soc. B Biol. Sci.* 271:705–714.
- Parsch, J., S. Novozhilov, S. S. Saminadin-Peter, K. M. Wong, and P. Andolfatto. 2010. On the utility of short intron sequences as a reference for the detection of positive and negative selection in *Drosophila*. *Mol. Biol. Evol.* 27:1226–1234.
- Patten, M. M. and D. Haig. 2009. Maintenance or loss of genetic variation under sexual and parental antagonism at a sex-linked locus. *Evolution* 63:2888–2895.
- Pearse, D. E., N. J. Barson, T. Nome, G. Gao, M. A. Campbell, A. Abadía-Cardoso, et al. 2019. Sex-dependent dominance maintains migration supergene in rainbow trout. *Nat. Ecol. Evol.* 3:1731–1742.
- Pick, J. L., P. Hutter, and B. Tschirren. 2017. Divergent artificial selection for female reproductive investment has a sexually concordant effect on male reproductive success. *Evol. Lett.* 1:222–228.
- Pirastu, N., M. Cordioli, P. Nandakumar, G. Mignogna, A. Abdellaoui, B. Hollis, et al. 2020. Genetic analyses identify widespread sex-differential participation bias. *bioRxiv*. <https://doi.org/10.1101/2020.03.22.001453>
- Prasad, N. G., S. Bedhomme, T. Day, and A. K. Chippindale. 2007. An evolutionary cost of separate genders revealed by male-limited evolution. *Am. Nat.* 169:29–37.

- Price, A. L., N. A. Zaitlen, D. Reich, and N. Patterson. 2010. New approaches to population stratification in genome-wide association studies. *Nat. Rev. Genet.* 11:459–463.
- Price, N., B. T. Moyers, L. Lopez, J. R. Lasky, J. Grey Monroe, J. L. Mullen, et al. 2018. Combining population genomics and fitness QTLs to identify the genetics of local adaptation in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 115:5028–5033.
- 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.
- Quinlan, A. R. and I. M. Hall. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842.
- R Core Team. 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.
- Randall, J. C., T. W. Winkler, Z. Kutalik, S. I. Berndt, A. U. Jackson, K. L. Monda, 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.
- Rice, W. R. 1998. Male fitness increases when females are eliminated from gene pool: implications for the Y chromosome. *Proc. Natl. Acad. Sci. USA* 95:6217–6221.
- . 1992. Sexually antagonistic genes: experimental evidence. *Science* 256:1436–1439.
- . 1996. Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. *Nature* 381:232–234.
- Rice, W. R. and A. K. Chippindale. 2001. Intersexual ontogenetic conflict. *J. Evol. Biol.* 14:685–693.
- Rice, W. R., J. E. Linder, U. Friberg, T. A. Lew, E. H. Morrow, and A. D. Stewart. 2005. Inter-locus antagonistic coevolution as an engine of speciation: assessment with hemiclinal analysis. *Proc. Natl. Acad. Sci. USA* 102:6527–6534.
- Riggio, V., O. Matika, R. Pong-Wong, M. J. Stear, and S. C. Bishop. 2013. Genome-wide association and regional heritability mapping to identify loci underlying variation in nematode resistance and body weight in Scottish Blackface lambs. *Heredity* 110:420–429.
- Roberts, R. B., J. R. Ser, and T. D. Kocher. 2009. Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes. *Science* 326:998–1001.
- Robertson, A. 1962. Selection for heterozygotes in small populations. *Genetics* 47:1291–1300.
- . 1975. Gene frequency distributions as a test of selective neutrality. *Genetics* 81:775–785.
- Rostant, W. G., C. Kay, N. Wedell, and D. J. Hosken. 2015. Sexual conflict maintains variation at an insecticide resistance locus. *BMC Biol.* 13:34.
- Rowe, L., S. F. Chenoweth, and A. F. Agrawal. 2018. The genomics of sexual conflict. *Am. Nat.* 192:274–286.
- Rundle, H. D., S. F. Chenoweth, and M. W. Blows. 2006. The roles of natural and sexual selection during adaptation to a novel environment. *Evolution* 60:2218–2225.
- Ruzicka, F., M. S. Hill, T. M. Pennell, I. Flis, F. C. Ingleby, R. Mott, et al. 2019. Genome-wide sexually antagonistic variants reveal long-standing constraints on sexual dimorphism in fruit flies. *PLoS Biol.* 17:e3000244.
- Sayadi, A., A. Martínez Barrio, E. Immonen, J. Dainat, D. Berger, C. Tellgren-Roth, et al. 2019. The genomic footprint of sexual conflict. *Nat. Ecol. Evol.* 3:1725–1730.
- Schlötterer, C., R. Kofler, E. Versace, R. Tobler, and S. U. Franssen. 2015. Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. *Heredity* 114:431–440.
- Sharp, N. P. and A. F. Agrawal. 2013. Male-biased fitness effects of spontaneous mutations in *Drosophila melanogaster*. *Evolution* 67:1189–1195.
- Singh, A. and D. Punzalan. 2018. The strength of sex-specific selection in the wild. *Evolution* 72:2818–2824.
- Smith, D. T., D. J. Hosken, W. G. Rostant, M. Yeo, R. M. Griffin, A. Bretman, et al. 2011. DDT resistance, epistasis and male fitness in flies. *J. Evol. Biol.* 24:1351–1362.
- Snyder-Mackler, N., W. H. Majoros, M. L. Yuan, A. O. Shaver, J. B. Gordon, G. H. Kopp, et al. 2016. Efficient genome-wide sequencing and low-coverage pedigree analysis from noninvasively collected samples. *Genetics* 203:699–714.
- Sohail, M., R. M. Maier, A. Ganna, A. Bloemendal, A. R. Martin, M. C. Turchin, et al. 2019. Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *eLife* 8:e39702.
- Speed, D., N. Cai, M. R. Johnson, S. Nejentsev, and D. J. Balding. 2017. Reevaluation of SNP heritability in complex human traits. *Nat. Genet.* 49:986–992.
- Stearns, S. C., D. R. Govindaraju, D. Ewbank, and S. G. Byars. 2012. Constraints on the coevolution of contemporary human males and females. *Proc. R. Soc. B Biol. Sci.* 279:4836–4844.
- Sudlow, C., J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, et al. 2015. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* 12:e1001779.
- Svensson, E. I., A. G. McAdam, and B. Sinervo. 2009. Intralocus sexual conflict over immune defense, gender load, and sex-specific signaling in a natural lizard population. *Evolution* 63:3124–3135.
- Svensson, E. I., D. Goedert, M. A. Gómez-Llano, F. Spagopoulou, A. Nava-Bolaños, and I. Booksmythe. 2018. Sex differences in local adaptation: what can we learn from reciprocal transplant experiments? *Philos. Trans. R. Soc. B Biol. Sci.* 373:20170420.
- Tazzyman, S. J. and J. K. Abbott. 2015. Self-fertilization and inbreeding limit the scope for sexually antagonistic polymorphism. *J. Evol. Biol.* 28:723–729.
- Trochet, A., E. A. Courtois, V. M. Stevens, M. Baguette, A. Chaine, D. S. Schmeller, et al. 2016. Evolution of sex-biased dispersal. *Q. Rev. Biol.* 91:297–320.
- Tsai, K. L., J. M. Evans, R. E. Noorai, A. N. Starr-Moss, and L. A. Clark. 2019. Novel Y chromosome retrocopies in canids revealed through a genome-wide association study for sex. *Genes* 10:320.
- Turner, T. L., A. D. Stewart, A. T. Fields, W. R. Rice, and A. M. Tarone. 2011. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. *PLoS Genet.* 7:e1001336.
- Úbeda, F., D. Haig, and M. M. Patten. 2011. Stable linkage disequilibrium owing to sexual antagonism. *Proc. R. Soc. B Biol. Sci.* 278:855–862.
- Van Doorn, G. S. 2009. Intralocus sexual conflict. *Ann. N. Y. Acad. Sci.* 1168:52–71.
- VanKuren, N. W. and M. Long. 2018. Gene duplicates resolving sexual conflict rapidly evolved essential gametogenesis functions. *Nat. Ecol. Evol.* 2:705–712.
- Vaux, F., L. K. Rasmuson, L. A. Kautzi, P. S. Rankin, M. T. O. Blume, K. A. Lawrence, et al. 2019. Sex matters: otolith shape and genomic variation in deacon rockfish (*Sebastes diaconus*). *Ecol. Evol.* 9:13153–13173.
- Veltsos, P., Y. Fang, A. R. Cossins, R. R. Snook, and M. G. Ritchie. 2017. Mating system manipulation and the evolution of sex-biased gene expression in *Drosophila*. *Nat. Commun.* 8:2072.
- Visscher, P. M., N. R. Wray, Q. Zhang, P. Sklar, M. I. McCarthy, M. A. Brown, et al. 2017. 10 years of GWAS discovery: biology, function, and translation. *Am. J. Hum. Genet.* 101:5–22.

- Vitti, J. J., S. R. Grossman, and P. C. Sabeti. 2013. Detecting natural selection in genomic data. *Annu. Rev. Genet.* 47:97–120.
- Vlachos, C., C. Burny, M. Pelizzola, R. Borges, A. Futschik, R. Kofler, et al. 2019. Benchmarking software tools for detecting and quantifying selection in evolve and resequencing studies. *Genome Biol.* 20:169.
- Walsh, B. and M. W. Blows. 2009. Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. *Annu. Rev. Ecol. Evol. Syst.* 40:41–59.
- Wang, W. and G. Shan. 2015. Exact confidence intervals for the relative risk and the odds ratio. *Biometrics* 71:985–995.
- Weir, B. S. 1997. *Genetic data analysis II*. Sinauer Associates, Sunderland, MA.
- Whitlock, M. C. and K. E. Lotterhos. 2015. Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of F_{ST} . *Am. Nat.* 186:S24–S36.
- Wiberg, R. A. W., O. E. Gaggiotti, M. B. Morrissey, and M. G. Ritchie. 2017. Identifying consistent allele frequency differences in studies of stratified populations. *Methods Ecol. Evol.* 8:1899–1909.
- Winkler, T. W., A. E. Justice, M. Graff, L. Barata, M. F. Feitosa, S. Chu, et al. 2015. The influence of age and sex on genetic associations with adult body size and shape: a large-scale genome-wide interaction study. *PLoS Genet.* 12:e1006166.
- Wright, A. E., M. Fumagalli, C. R. Cooney, N. I. Bloch, F. G. Vieira, S. D. Buechel, et al. 2018. Sex-biased gene expression resolves sexual conflict through the evolution of sex-specific genetic architecture. *Evol. Lett.* 2:52–61.
- Wright, A. E., T. F. Rogers, M. Fumagalli, C. R. Cooney, and J. E. Mank. 2019. Phenotypic sexual dimorphism is associated with genomic signatures of resolved sexual conflict. *Mol. Ecol.* 28:2860–2871.
- Wright, S. 1951. The genetical structure of species. *Ann. Eugen.* 15:323–354.
- Yang, J., S. H. Lee, M. E. Goddard, and P. M. Visscher. 2011. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 88:76–82.

Associate Editor: J. Mank

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Permuted versus observed \hat{F}_{ST} for simulated data.

Figure S2. Permuted versus observed \hat{F}_{ST} for the flycatcher dataset.

Figure S3. Permuted versus observed \hat{F}_{ST} for the pipefish dataset.

Figure S4. Permuted vs. observed \hat{F}_{ST} for the human 1000 Genomes data.

Figure S5. Human 1000 Genomes data with SNPs with low minor allele frequencies (MAFs) included in the analysis.

Appendix A. Distribution of F_{ST} estimates.

Table S1. Sex-specific relative fitness of genotypes of a biallelic locus with additive fitness effects in each sex.

Appendix B. Sex-specific allele frequencies and F_{IS} estimates.

Appendix C. Case-control GWAS and the Log-Odds Ratio.

Figure S6. Probability that \mathcal{L} / SE for an additive SA locus.

Figure S7. Theoretical signal of multilocus SA polymorphism: the ratio of observed versus permuted absolute log odds ratio estimates within 100 quantiles of the theoretical null of \mathcal{L} / SE .

Appendix D. Comparison of null model distributions for different metrics allele frequency differentiation between sexes.