# Polygenic signals of sex differences in selection in contemporary humans

Filip Ruzicka[1*], Luke Holman[2,3], Tim Connallon[1]

[1] School of Biological Sciences, Monash University, Clayton, Victoria, Australia

[2] School of BioSciences, University of Melbourne, Parkville, Victoria, Australia

[3] School of Applied Sciences, Edinburgh Napier University, Edinburgh, UK

[*]**Corresponding author e-mail**: Filip.Ruzicka@monash.edu

## Abstract

Sex differences in the fitness effects of genetic variants can influence the rate of adaptation and the maintenance of genetic variation. For example, "sexually antagonistic" (SA) variants, which are beneficial for one sex and harmful for the other, can both constrain adaptation and increase genetic variability for fitness components such as survival, fertility, and disease susceptibility. However, detecting variants with sex-differential fitness effects is difficult, requiring genome sequences and fitness measurements from large numbers of individuals. Here, we develop new theory for studying sex-differential selection across a complete life cycle, and test our models with genotypic and reproductive success data from ~250,000 UK Biobank individuals. We uncover polygenic signals of sex-differential selection affecting survival, reproductive success and overall fitness, with signals of sex-differential reproductive selection reflecting a combination of SA polymorphisms and sexually concordant polymorphisms in which the strength of selection differs between the sexes. Moreover, these signals hold up to rigorous controls that minimise the contributions of potential confounders, including sequence mapping errors, population structure, and ascertainment bias. Functional analyses reveal that sex-differentiated sites are enriched in phenotype-altering genomic regions, including coding regions and loci affecting a range of quantitative traits. Population genetic analyses show that sex-differentiated sites exhibit evolutionary histories dominated by genetic drift and/or transient balancing selection, but not long-term balancing selection, which is consistent with theoretical predictions of effectively weak SA balancing selection in historically small populations. Overall, our results are consistent with polygenic sex-differential—including SA—selection in humans. Evidence for sex-differential selection is particularly strong for variants affecting reproductive success, in which the potential contributions of non-random sampling to statistical signals of sex-differentiation can be excluded.

## Introduction

Adaptation of a population to its environment requires heritable genetic variation for fitness [1]. Although many populations show substantial genetic variation for fitness components [2]—including life-history traits such as maturation rate, lifespan, mating success, and fertility [2,3]—genetic trade-offs between components, or between different types of individuals in a population, limit adaptive potential [4]. For example, a mutation that increases the probability of survival to adulthood might simultaneously decrease adult reproductive success (*e.g.*, [5]), weakening the mutation's net fitness effect [4]. In addition to slowing adaptation [6–8], genetic trade-offs can increase standing genetic variation [2,9], give rise to balancing selection [10,11], and favour evolutionary transitions between mating systems [12,13], modes of sex determination [14] and genome structures [15–18].

Sexually antagonistic (SA) genetic polymorphisms—in which the alleles that benefit one sex are harmful to the other—are a type of genetic trade-off that may be common in sexually reproducing species [19]. Theory shows that SA polymorphisms are likely to arise when mutations differentially affect trait expression in each sex, or when mutations similarly affect traits under divergent directional selection between the sexes [20]. Empirical quantitative genetic studies imply that both conditions are frequently met in nature [21–24] and, accordingly, that SA polymorphisms contribute to phenotypic variation in a range of plant and animal populations (*e.g.* [25–27]), including humans [28–31].

Although there is now abundant evidence that SA polymorphisms contribute to phenotypic variation, efforts to identify and characterise SA alleles in genomic data face two formidable challenges [32]. First, methods using explicit fitness measurements to identify SA polymorphisms (*e.g.*, genome-wide association studies (GWAS) of fitness [33]) are rarely feasible, because it is challenging to obtain fitness measurements for large numbers of genotyped individuals under natural conditions [2]. Second, methods using allele frequency differences between adult females and males as genomic signals of SA viability selection (*e.g.*,

62 between-sex $F_{ST}$ estimates [32,34–41]) are limited in several ways: they have low power to
63 detect SA loci, they cannot distinguish SA selection from sex differences in the strength of
64 selection, they are susceptible to artefacts generated by population structure and mis-mapping
65 of sequence reads to sex chromosomes [32,39,40,42], and they neglect fitness components
66 other than viability, such as reproductive success [32,43]. Previous studies of human genomic
67 data [34,35,42,44] have been affected by one or more of these issues [32,40,41,43], such that
68 we currently lack robust evidence of SA genomic variation in humans. More generally, these
69 impediments help to explain the limited catalogue of SA polymorphisms across species [45–
70 47], which currently comprises a handful of loci with exceptionally large phenotypic effects
71 (*e.g.* [48–52]).

72 Despite these challenges, new datasets and analytical approaches provide opportunities
73 to identify robust genomic signals of SA selection. First, massive 'biobank' datasets, which are
74 widely used in human genomics, sometimes include both genotype and offspring number data
75 [29,53] that can be used to detect loci with SA effects on reproductive components of fitness
76 [32]. Second, estimates of allele frequency differences between sexes—though ill-suited for
77 confidently identifying individual SA loci affecting viability—may nevertheless be amenable
78 to genome-wide tests for *polygenic* SA viability selection [32,34]. Third, population genomic
79 metrics of sex-differential selection (*e.g.*, between-sex $F_{ST}$) may include an appreciable
80 proportion of genuine SA loci in the upper tails of their distributions, providing a set of
81 candidate loci that can collectively yield insights into the general properties of SA
82 polymorphisms (*e.g.*, their functional characteristics and evolutionary dynamics), despite
83 uncertainty about individual candidates.

84 Here, we extend [32,34] and develop new statistical tests based on $F_{ST}$ metrics of
85 between-sex allele frequency differentiation to detect polygenic signals of sex-differential
86 selection affecting viability, reproduction, and total fitness during a full generational cycle.
87 Applying these tests to the UK Biobank [53]—a dataset comprising quality-filtered genotype

and offspring number data for ~250,000 men and women—reveals polygenic signals of sex-differential—including SA—polymorphism. We corroborate these results by using mixed-model statistics that explicitly control for systematic differences in the genetic ancestry of female and male individuals and potential sequencing artefacts. We further show that sex-differentiated polymorphisms are preferentially situated in functional, phenotype-altering genomic sequences. Finally, we use genetic diversity data to examine modes of evolution affecting sex-differentiated sites.

## Results

### Genomic signals of sex differences in selection: theoretical predictions

Previous studies have examined sex-differential effects of genetic variation during the zygote-to-adult stage by comparing allele frequencies between adult females and males [32,34,36–39,42]. By contrast, our analytical approach combines allele frequency with offspring number data to estimate sex-differential effects during a full generational life cycle (Fig. 1). To illustrate the approach, consider a large, well-mixed population containing many polymorphic, bi-allelic, autosomal loci. At fertilisation, Mendelian inheritance equalises allele frequencies between the sexes (Fig. 1, left box). In the zygote-to-adult stage, loci with sex-differential effects on survival accumulate allele frequency differences between the adults of each sex (*e.g.*, the black allele becomes enriched in adult males and deficient in adult females because it improves zygote-to-adult survival in males but reduces it in females; Fig. 1, middle box). Among the adults, alleles with sex-differential effects on reproductive success have different transmission rates to the next generation from surviving females *vs.* surviving males (*e.g.*, the black allele is enriched among the male gametes contributing to fertilization but deficient among female gametes, thus increasing its transmission to offspring of males but decreasing transmission to offspring of females; Fig. 1, right box).

113 Adult allele frequencies, coupled with offspring number data per individual, thus
114 provide an opportunity to estimate sex-differential effects of genetic variation during a
115 complete life cycle, even though zygotic and gametic allele frequencies are inferred and not
116 directly observed. Below, we apply our approach to the UK Biobank, a dataset that includes
117 genotypes and reported offspring numbers (hereafter "lifetime reproductive success" or LRS,
118 following standard terminology [29]) among putatively post-reproductive adults (ages 45-69
119 after filtering; see Materials and Methods). For a biallelic autosomal locus with alleles $A_1$ and
120 $A_2$, we denote $\hat{p}_m$ and $\hat{p}_f$ the respective estimated frequencies of the $A_1$ allele in adult males
121 and females of the UK Biobank. The projected frequencies of $A_1$ in paternal and maternal
122 gametes contributing to fertilisation are:

123 $$\hat{p}'_m = \frac{M_{11} + \frac{1}{2}M_{12}}{M_{11} + M_{12} + M_{22}} \quad [1a]$$

124 $$\hat{p}'_f = \frac{F_{11} + \frac{1}{2}F_{12}}{F_{11} + F_{12} + F_{22}} \quad [1b]$$

125 where $M_{ij}$ and $F_{ij}$ represent the cumulative LRS of males and females, respectively, with
126 genotype *ij* (*e.g.*, $M_{11}$, $M_{12}$ and $M_{22}$ correspond to genotypes $A_1A_1$, $A_1A_2$, and $A_2A_2$).

127 Using $F_{ST}$ [54], we partition between-sex allele frequency differentiation over one
128 generation into three components: 1) differentiation among adults, which includes effects of
129 sex-differential survival (hereafter "adult $F_{ST}$"; see [32,34,43]); 2) sex-differential variation in
130 adult LRS (hereafter "reproductive $F_{ST}$"); and 3) sex-differential variation in overall fitness
131 (hereafter "gametic $F_{ST}$"). Single-locus estimates of adult, reproductive and gametic $F_{ST}$ are
132 defined, respectively, as:

133 $$\hat{F}_{ST(Adult)} = \frac{\left(\hat{p}_f - \hat{p}_m\right)^2}{4\bar{p}(1-\bar{p})} \quad [2a]$$

134 $$\hat{F}_{ST(Reproductive)} = \frac{\left(\left(\hat{p}'_f - \hat{p}_f\right) - \left(\hat{p}'_m - \hat{p}_m\right)\right)^2}{4\bar{p}(1-\bar{p})} \quad [2b]$$

$$\hat{F}_{ST(Gametic)} = \frac{\left(\hat{p}'_f - \hat{p}'_m\right)^2}{4\bar{p}'(1-\bar{p}')} \qquad [2c]$$

where $\bar{p} = \left(\hat{p}_f + \hat{p}_m\right)/2$ and $\bar{p}' = \left(\hat{p}'_f + \hat{p}'_m\right)/2$.

*$F_{ST}$ distributions in the absence of sex-differential selection*

In the absence of sex differences in selection (*e.g.*, under neutrality, or under sexually concordant (SC) selection of equal magnitude and direction in each sex), with large sample sizes, negligible Hardy-Weinberg deviations at birth, and excluding single-nucleotide polymorphisms (SNPs) with very low minor allele frequencies, we show that the adult, reproductive and gametic $\hat{F}_{ST}$ metrics converge, respectively, to the following distributions:

$$\hat{F}_{ST(Adult)} \approx \left(\frac{1}{8N_f} + \frac{1}{8N_m}\right)X \qquad [3a]$$

$$\hat{F}_{ST(Reproductive)} \approx \frac{\frac{\hat{p}_f(1-\hat{p}_f)}{2N_f}\frac{\sigma_f^2}{\mu_f^2}\left(1-\hat{F}_{IS}^f\right) + \frac{\hat{p}_m(1-\hat{p}_m)}{2N_m}\frac{\sigma_m^2}{\mu_m^2}\left(1-\hat{F}_{IS}^m\right)}{4\hat{p}(1-\hat{p})}X \qquad [3b]$$

$$\hat{F}_{ST(Gametic)} \approx \left[\frac{1}{8N_f}\left(1+\frac{\sigma_f^2}{\mu_f^2}\right) + \frac{1}{8N_m}\left(1+\frac{\sigma_m^2}{\mu_m^2}\right)\right]X \qquad [3c]$$

where each $X$ is an independent chi-square random variable with one degree of freedom, $N_f$ and $N_m$ denote adult sample sizes, $\mu_f$ and $\mu_m$ denote mean LRS, $\sigma_f^2$ and $\sigma_m^2$ denote variances in LRS, and $\hat{F}_{IS}^f$ and $\hat{F}_{IS}^m$ quantify sex-specific departures from Hardy-Weinberg equilibrium in the sample of adults (*SI Appendix*, Section A). In datasets such as the UK Biobank, there is also between-site variation in the number of genotyped individuals and the extent of Hardy-Weinberg deviations in the adult sample. The null distributions described by eqs. [3a-c] are easily adjusted to account for this between-site variation (see Materials and Methods).

Relative to the null distributions in eqs. [3a-c], sex differences in selection inflate each $\hat{F}_{ST}$ metric (*SI Appendix*, Section A). These inflations may arise due to polymorphisms under sex-differential selection *and* neutral polymorphisms that hitchhike with selected

157 polymorphisms. However, linkage disequilibrium (LD) alone cannot inflate genome-wide $\hat{F}_{ST}$
158 in the absence of genuine selected polymorphisms (*SI Appendix*, Section B). As such, $\hat{F}_{ST}$
159 inflations represent reliable signals of sex-differentially selected polymorphism [32], provided:
160 (*i*) technical artefacts are controlled (as shown below); (*ii*) sex-specific population structure is
161 controlled; and (*iii*) males and females are sampled at random (though (*iii*) is not a requirement
162 for reproductive $\hat{F}_{ST}$; see Discussion). To simplify the presentation, we first present analyses
163 using $F_{ST}$ metrics, but we return to non-$F_{ST}$ metrics in the section "Controlling for sex-specific
164 population structure".
165
166 **Genomic signals of sex differences in selection: empirical data**
167 *UK Biobank SNP data*
168 The sample size in the UK Biobank, after removing individuals that were closely related, had
169 a recorded ancestry other than 'White British', or had missing LRS data, was $N$ = 249,021 ($N_m$
170 = 115,531 males and $N_f$ = 133,490 females). We removed rare polymorphic sites (MAF < 1%),
171 sites with low genotype or imputation quality, and sites with high potential for artefactual
172 between-sex differentiation based on criteria identified by Kasimatis et al. [42] (*i.e.*, between-
173 sex differences in missing rates, deficits of minor allele homozygotes, and heterozygosity
174 levels exceeding what can be plausibly be explained by sex differences in selection; see *SI*
175 *Appendix*, Section C). Reassuringly, none of the 8 sites that Kasimatis et al. [42] identified as
176 false-positives for sex-differential viability selection appear among the quality-filtered, LD-
177 pruned, imputed SNPs (N = 1,051,949) that are the focus of our analyses.
178
179 *Observed $F_{ST}$ distributions relative to null distributions*
180 We tested for sex differences in selection by calculating adult, reproductive and gametic $\hat{F}_{ST}$
181 (eqs. [2a-c]) in the UK Biobank, and contrasting these estimates against: (*i*) their respective
182 theoretical null distributions (eqs. [3a-c]), and (*ii*) empirical null distributions (generated by a

183 single random permutation of male and female labels among individuals or, in the case of
184 reproductive $\hat{F}_{ST}$, a single permutation of LRS among individuals of each sex; see Materials
185 and Methods).

186 All three $\hat{F}_{ST}$ metrics showed greater between-sex differentiation than predicted by their
187 theoretical and empirical null distributions, consistent with sex differences in selection
188 contributing to genome-wide allele frequency differentiation between the sexes. Mean adult
189 $\hat{F}_{ST}$ in the observed data was larger than predicted by both null distributions (theoretical null:
190 $2.039\times10^{-6}$; permuted null: $2.043\times10^{-6}$; observed: $2.104\times10^{-6}$; Wilcoxon and Kolmogorov-
191 Smirnov tests, $p<0.001$; Fig. 2A,D), with a 14.1% and 13.7% excess of SNPs in the top
192 percentile of the theoretical and empirical nulls, respectively ($\chi^2$ tests, $p<0.001$). Mean
193 reproductive $\hat{F}_{ST}$ was also larger than predicted by both nulls (theoretical null: $8.731\times10^{-7}$;
194 permuted null: $8.749\times10^{-7}$; observed: $8.900\times10^{-7}$; Wilcoxon and Kolmogorov-Smirnov tests,
195 $p<0.001$; Fig. 2B,E), with a 7.4% and 5.0% excess of SNPs in the top percentile of the
196 theoretical and empirical nulls ($\chi^2$ tests, $p<0.001$). Moreover, mean gametic $\hat{F}_{ST}$ was larger
197 than predicted by both nulls (theoretical null: $2.908\times10^{-6}$; permuted null: $2.907\times10^{-6}$;
198 observed: $2.974\times10^{-6}$; Wilcoxon and Kolmogorov-Smirnov tests, $p<0.001$; Fig. 2C,F), with a
199 9.0% and 7.8% excess of SNPs in the top percentile of the theoretical and empirical nulls ($\chi^2$
200 tests, $p<0.001$).

201 Signals of sex differences in selection in adult, reproductive, and gametic $\hat{F}_{ST}$ were
202 polygenic. For example, genetic variants situated in genomic regions with high LD tended to
203 explain more SNP-heritability of each $\hat{F}_{ST}$ metric than variants situated in low-LD regions, as
204 predicted if each sex-differential fitness component has a polygenic basis (*SI Appendix*, Section
205 D). Moreover, no individual loci had p-values below the Bonferroni-corrected threshold of
206 $4.753\times10^{-8}$, implying that the significant overall inflations were not driven by a small number
207 of strongly sex-differentiated polymorphisms (adult $\hat{F}_{ST}$: minimum p- and q-values=$2.237\times10^{-}$

$^{7}$ and 0.176; reproductive $\hat{F}_{ST}$: minimum p- and q-values =3.925×10$^{-7}$ and 0.413; gametic $\hat{F}_{ST}$: minimum p- and q-values=4.152×10$^{-6}$ and 0.821).

**Forms of sex-differential selection: theoretical predictions**

The $\hat{F}_{ST}$ elevations reported above indicate the presence of polygenic sex-differential selection in the UK Biobank. However, the signals could have arisen because of SA selection, because of sex differences in the strength but not the direction of selection (*i.e.*, sex-differential SC selection), or a combination of both scenarios. To partition signals affecting LRS into SA and SC components, we examined the effects of a given allele on LRS in each sex relative to the other. Specifically, estimates of the product $(\hat{p}'_m - \hat{p}_m)(\hat{p}'_f - \hat{p}_f)$ should tend to be negative when alleles have SA effects, and positive when alleles have SC effects (Fig. 3A). A new metric, termed "unfolded reproductive $\hat{F}_{ST}$", provides a standardized measure of the product of sex-specific effects on LRS:

$$\hat{F}_{ST(Unfolded)} = \frac{(\hat{p}'_m - \hat{p}_m)(\hat{p}'_f - \hat{p}_f)}{\sqrt{\frac{\hat{p}_m(1-\hat{p}_m)}{2n_m}\frac{\sigma_m^2}{\mu_m^2}\left(1-\hat{F}_{IS}^m\right)\frac{\hat{p}_f(1-\hat{p}_f)}{2n_f}\frac{\sigma_f^2}{\mu_f^2}\left(1-\hat{F}_{IS}^f\right)}} \quad [4].$$

In the absence of any selection on LRS, unfolded reproductive $\hat{F}_{ST}$ is distributed as the product of two independent, standard normal distributions (*i.e.*, symmetrically distributed with a mean of zero; see *SI Appendix*, Section E). SA selection generates an excess of loci in the lower quantiles of this null model, while SC selection generates an excess of loci in the upper quantiles of the null. Note that sex differences in SC selection are not required to generate an excess of positive values for unfolded reproductive $\hat{F}_{ST}$ (SC selection of equal magnitude in the sexes can generate it as well), but SA selection is required to generate an excess of negative values.

**Forms of sex-differential selection: empirical data**

232 As with previous $\hat{F}_{ST}$ metrics, we calculated unfolded reproductive $\hat{F}_{ST}$ (eq. [4]) and contrasted
233 it against its theoretical and empirical null distributions—the latter generated by a single
234 random permutation of LRS among the individuals of each sex. Doing so revealed that both
235 SC *and* SA sites contribute to the polygenic signal of sex-differential selection affecting LRS.
236 As predicted under SC selection, we observed an enrichment of sites in the upper quantiles of
237 the null distributions of unfolded reproductive $\hat{F}_{ST}$ (mean $\hat{F}_{ST}$ among sites with $\hat{F}_{ST}$>0;
238 theoretical null: 0.637; permuted null: 0.640; observed: 0.694; Wilcoxon and Kolmogorov-
239 Smirnov tests, p<0.001; Fig. 3B,C). As predicted under SA selection, we observed a smaller
240 but significant enrichment of sites in the lower quantiles of the null (mean $\hat{F}_{ST}$ among sites with
241 $\hat{F}_{ST}$<0; theoretical null: –0.635; permuted null: –0.638; observed: –0.651; Wilcoxon and
242 Kolmogorov-Smirnov tests, p<0.001; Fig. 3B,C).
243
244 **Controlling for sex-specific population structure**
245 In principle, polygenic $\hat{F}_{ST}$ elevations can arise entirely in the absence of genuine sex
246 differences in selection [32,43] if there are systematic differences in ancestry (population
247 structure) between sexes in the sampled population. We therefore replicated our analyses using
248 mixed-model association tests that are analogous to $\hat{F}_{ST}$ but which explicitly correct for sex-
249 specific population structure (see also *SI Appendix*, Section F).
250 We first re-evaluated signals of sex differences in viability selection present in adult
251 $\hat{F}_{ST}$ by performing a GWAS of sex [32,42,55] using standardised estimates of the log-odds
252 ratio ($\hat{\mathcal{L}}_{ST}$; see Materials and Methods). Like adult $\hat{F}_{ST}$, $\hat{\mathcal{L}}_{ST}$ quantifies between-sex allele
253 frequency differences among adults; moreover, it controls for population structure by including
254 a kinship matrix of genome-wide relatedness between individuals and principal components
255 which capture structure-induced axes of genetic variation (see Materials and Methods). As
256 expected, $\hat{\mathcal{L}}_{ST}$ was highly correlated with adult $\hat{F}_{ST}$ ($r_g$±SE=1.046±0.020; p<0.001) and mean

257 $\hat{\mathcal{L}}_{ST}$ was elevated relative to its empirical null distribution (null $\hat{\mathcal{L}}_{ST}$: $5.236\times10^{-7}$; observed:
258 $5.323\times10^{-7}$; Wilcoxon and Kolmogorov-Smirnov tests, p<0.001; Fig. 4A,D), with 8.9% excess
259 of SNPs in the top 1% quantile of the empirical null ($\chi^2$ test, p<0.001).

260 We then re-evaluated signals of sex-differential selection through reproductive success
261 by performing separate GWAS for LRS in females and males, each corrected for population
262 structure, and quantifying the difference between female and male effect sizes using a *t*-statistic
263 (|*t*|; see Materials and Methods). As expected, |*t*| was highly correlated with reproductive $\hat{F}_{ST}$
264 ($r_g\pm$SE=1.025$\pm$0.059, p<0.001) and mean |*t*| was elevated relative to its empirical null
265 (null=0.796, observed=0.811, Wilcoxon and Kolmogorov-Smirnov tests, p<0.001; Fig. 4B,E),
266 with an 11.9% excess of SNPs in the top 1% quantile of the empirical null ($\chi^2$ test, p<0.001).

267 We also developed an analogue of unfolded reproductive $\hat{F}_{ST}$, termed unfolded *t* (see
268 Materials and Methods), to partition signals of sex-differential reproductive selection into SA
269 and SC components. Like unfolded reproductive $\hat{F}_{ST}$, SC selection should generate an
270 enrichment of values in the upper quantiles of its null, while SA selection should generate an
271 enrichment of values in its lower quantiles; unlike unfolded reproductive $\hat{F}_{ST}$, this metric also
272 controls for population structure. Corroborating previous results, we observed an excess of
273 high values of unfolded *t* (mean *t* among sites with *t*>0; permuted null=0.639, observed=0.692,
274 Wilcoxon and Kolmogorov-Smirnov tests, p<0.001; Fig. 4C,F) and an excess of low values of
275 unfolded *t* (mean *t* among sites with *t*<0; permuted null=–0.639, observed=–0.649, Wilcoxon
276 and Kolmogorov-Smirnov tests, p<0.001), signalling the presence of SC and SA
277 polymorphisms, respectively.

278 Finally, we examined genetic correlations between metrics. These analyses showed that
279 metrics of sex-differential LRS selection were not significantly correlated with metrics of sex-
280 differential mortality selection across loci (Fig. 5A). For example, the genetic correlation

281 (estimated via LD score regression) between adult and reproductive $\hat{F}_{ST}$ was –0.24 (SE = 0.16,
282 p = 0.13) and the genetic correlation between $\hat{\mathcal{L}}_{ST}$ and |t| was –0.16 (SE = 0.16, p=0.31).
283
284 **Functional and phenotypic effects of sex-differentiated loci**
285 If sex-differentiated loci reflect genuine sex-differential selection—rather than random chance,
286 genotyping errors, or population structure—such polymorphisms should be preferentially
287 found in functionally important regions in the genome. We therefore conducted enrichment
288 tests, both to support our inference that sex-differential selection is occurring and to explore
289 functional effects of sex-differentiated loci.
290 We first used LD-score regression to test whether sites with high sex-differentiation
291 tend to be found in major functional categories in the genome (coding, 3'UTR and 5'UTR
292 regions). If a given category is enriched for genuine selected SNPs, the expected heritability
293 tagged by these SNPs (*i.e.*, what LD-score regression measures) should exceed the fraction of
294 SNPs present in that functional category. While functional enrichment estimates were noisy
295 and thus not statistically distinguishable from 1 (no enrichment) after multiple-testing
296 correction (Fig. 5B), each estimate consistently exceeded 1 across functional categories and
297 metrics, suggesting that sex-differentiated loci are more likely to have phenotype-altering
298 effects than expected by chance.
299 Further evidence for the phenotype-altering effects of sex-differentiated loci was
300 sought through direct comparisons between metrics of sex-differential selection and the Neale
301 laboratory database of UK Biobank GWAS. Specifically, we used cross-trait LD-score
302 regression to estimate genetic correlations between metrics of sex-differential selection and 30
303 phenotypes, chosen for their medical relevance and/or relationship to phenotypic sex
304 differences. Though many significant associations did not survive multiple testing correction
305 (Fig. 5C), several disease-relevant and quantitative traits (age at menarche, body fat percentage,
306 diseases of the eye and adnexa, fluid intelligence, injury, neuroticism score, SHBG [sex

307 hormone binding globulin], standing height) represent candidates for sex-differential viability
308 and LRS selection, while other traits (testosterone, high blood pressure) represent candidate
309 for sex-differential viability selection.
310
311 **Modes of evolution of sex-differentiated loci: theoretical predictions**
312 To gain insight into the modes of evolution affecting sex-differentiated sites, we investigated
313 the association between metrics of sex-differential selection and MAF in the UK Biobank. In
314 the absence of any contemporary sex differences in selection, all between-sex $\hat{F}_{ST}$ metrics
315 should be independent of MAF (*SI Appendix*, Section G). In the presence of sex-differential
316 selection, the association between each metric and MAF can potentially be positive or negative,
317 depending on the patterns of contemporary and historical selection affecting loci throughout
318 the genome. A positive covariance between $\hat{F}_{ST}$ and MAF should arise when alleles subject to
319 sex-differential selection often segregate at intermediate frequencies, as may occur under a
320 history of balancing selection or drift (*SI Appendix*, Section G), or non-equilibrium scenarios
321 such as incomplete selective sweeps. In contrast, a negative association between MAF and
322 between-sex $\hat{F}_{ST}$ is expected for loci that have evolved under sex-differential purifying
323 selection (*SI Appendix*, Section G). This negative covariance arises because purifying selection
324 disproportionately lowers the frequency of large-effect alleles (those generating larger $\hat{F}_{ST}$
325 values) relative to small-effect alleles [56]. In short, positive associations with MAF indicate
326 that purifying selection is not the dominant mode of evolution affecting loci under sex-
327 differential selection and instead signal a recent history of balancing selection, positive
328 selection, or drift.
329 While associations between metrics of sex-differential selection and MAF provide
330 insights into relatively recent and contemporary patterns of selection affecting sex-
331 differentiated sites, they do not provide insights into their deeper evolutionary histories. To
332 examine this, we tested the specific hypothesis that sex-differentiated sites are subject to *long-*

333 *term* balancing selection, as predicted for SA polymorphisms under certain scenarios of
334 selection and dominance [10]. Under long-term balancing selection, we would expect sex-
335 differentiated (and linked) loci to be old, to exhibit low between-population $\hat{F}_{ST}$, to exhibit high
336 genetic diversity, and to disproportionately co-localise with previous candidates for long-term
337 balancing selection, compared to less sex-differentiated sites with similar allele frequencies in
338 the UK Biobank.
339
340 **Modes of evolution of sex-differentiated loci: empirical data**
341 Examining the relationship between MAF and metrics of sex-differential selection in the UK
342 Biobank data revealed consistently positive correlations (adult $\hat{F}_{ST}$, $\rho$=0.009, p<0.001; $\hat{\mathcal{L}}_{ST}$:
343 $\rho$=0.006, p=0.216; reproductive $\hat{F}_{ST}$, $\rho$=0.006, p<0.001; $|t|$: $\rho$=0.005, p<0.001; gametic $\hat{F}_{ST}$,
344 $\rho$=0.007, p<0.001; Fig. 6A-D), with all correlations stronger in observed than null data (*SI*
345 *Appendix*, Section H). Given the absence of negative correlations between MAF and each
346 metric, we can reject purifying selection as the dominant mode of evolution affecting sex-
347 differentiated sites. The positive correlations instead suggest that balancing selection, drift, or
348 incomplete selective sweeps characterise the evolution of sex-differentiated loci.
349 We then tested the hypothesis that long-term balancing selection has shaped the
350 evolutionary histories of sex-differentiated loci. We focused our analyses on four measures of
351 balancing selection: allele age estimates from the Atlas of Variant Age database [57], between-
352 population $\hat{F}_{ST}$ and Tajima's D estimates from two non-European populations from the 1000
353 Genomes Project [58], and three sets of candidate loci for long-term balancing selection [59–
354 61]. In each case, we looked for associations between metrics of sex-differential selection and
355 balancing selection, while controlling for ascertainment bias of intermediate-frequency alleles
356 (which are, on average, older and thus more likely to be under long-term balancing selection
357 irrespective of the strength of sex-differential selection) among highly sex-differentiated sites
358 (see Materials and Methods). Overall, we found little support for the hypothesis of long-term

359 balancing selection affecting sex-differentiated loci. After corrections for multiple testing
360 across metrics of sex-differential selection (see *SI Appendix*, Section I, for full statistical
361 results), we found weak or absent associations with allele age (Fig. 6E-H), between-population
362 $\hat{F}_{ST}$ (*SI Appendix*, Section I), genetic diversity (*SI Appendix*, Section I), or previous candidates
363 for balancing selection (*SI Appendix*, Section I). We found some indications that candidate SA
364 alleles (*i.e.*, loci with negative values of unfolded reproductive $\hat{F}_{ST}$ and unfolded *t*) were older
365 than the genome-wide average (Fig. 6H), and loci experiencing strong SC selection (*i.e.*,
366 positive values of unfolded reproductive $\hat{F}_{ST}$ and unfolded *t*) were younger (Fig. 6H).
367

368 ## Discussion

369 Sex differences in directional selection on phenotypes have been reported in a wide range of
370 animal taxa [21–23,62], including post-industrial human populations [29,30], yet population
371 genomic signals of sex-differential selection—let alone SA selection—have been extremely
372 difficult to establish. The reason is simple: sexual reproduction equalises autosomal allele
373 frequencies between the sexes every generation, restricting genetic divergence and, in effect,
374 preventing the use of common tests (*e.g.*, McDonald-Kreitman tests for positive selection and
375 $F_{ST}$ outlier tests for spatially-varying selection [63–65]) to infer sex differences in selection.
376 Published studies using human genomic data illustrate the challenges of studying
377 polymorphisms with sex-differential fitness effects [32,40,41,43,44], including sample sizes
378 that may be insufficient for detecting polygenic signals of sex-differential selection, lack of
379 controls for population structure or technical artefacts, and/or absence of data concerning
380 reproductive fitness components.
381

382 **Signals of sex-differential selection in the UK Biobank**

383 We developed a theoretical framework for studying genomic variation with sex-differential
384 effects across a complete life cycle. Our approach extends current work based on between-sex

allele frequency differentiation among adults—a potential signal of sex-differential viability selection among juveniles [32,34,35,43]—to further include reproductive success components and total fitness. Applying this approach to data from a quarter-million UK adults, we present evidence for polygenic signals of sex-differential selection in humans. Specifically, UK Biobank individuals showed sex differences in allele frequencies—both among adults and their (projected) offspring—that consistently exceeded expectations defined by our theoretical null models for viability, reproductive and total fitness, and persisted after controlling for potential artefacts arising from mis-mapping of reads to sex chromosomes [55].

Although we focussed on $F_{ST}$ as our metric of differentiation for a variety of reasons (its simplicity, amenability to theoretical modelling, and rich history in population genetic studies of adaptation [63–65]), an important drawback of $F_{ST}$ is its inability to control for systematic sex differences in the genetic ancestry of sampled individuals. We therefore used $F_{ST}$ analogues based on mixed-model association tests to control for sex-specific population structure. These $F_{ST}$ analogues corroborated $F_{ST}$-derived signals of sex-differential selection on each component, with clear enrichments in the upper tails of each null distribution. Additional support for genuine sex-differential selection came from functional enrichment analyses, which, despite noisy individual estimates, consistently indicated that sex-differentiated sites were situated in functional genomic regions and contributed to variation for many phenotypes.

An important limitation of metrics of sex-differential selection affecting non-LRS fitness components (*i.e.*, adult $F_{ST}$, gametic $F_{ST}$ and their mixed-model analogues) applied to the UK Biobank is that UK Biobank individuals are sampled through active participation. Consequently, as noted by Pirastu et al. [66], sex differences in the genetic basis of individuals' predisposition to take part in the UK Biobank may generate sex differences in adult allele frequencies. To support this argument, Pirastu et al. reported significantly greater SNP-heritability of sex (a polygenic measure of sex differences in allele frequencies) in biobanks relying on active participation than in biobanks using passive participation. However, their

411 analysis is inconclusive because the passive participation studies they analysed were smaller
412 ($N_{\text{Biobank Japan}}$ = 178,242, $N_{\text{FinnGen}}$ = 150,831, $N_{\text{iPsych}}$ = 65,891) than active participation studies
413 ($N_{\text{UK Biobank}}$ = 452,302, $N_{\text{23andme}}$ = 2,462,132). Thus, differences in statistical power between
414 studies (and/or differences in the extent of sex-differential viability selection between
415 populations) could account for their results. Moreover, the positive point estimates of SNP-
416 heritability for passive participation studies suggest that substantial allele frequency
417 differences between the sexes are possible. For example, mortality after fertilisation, but before
418 birth, is very high in humans (on the order of 50% [67]), giving ample opportunity for mortality
419 in early life to generate allele frequency differences between sexes. In sum, neither their study
420 nor ours can conclusively distinguish the relative contributions of sex-differential selection and
421 participation bias to allele frequency differentiation between female and male adults, though
422 both sources likely contribute.

423 Importantly, participation bias should not affect metrics of sex-differential selection
424 relating to LRS. Reproductive $\hat{F}_{ST}$ and its mixed-model analogue, $|t|$, control for allele
425 frequency differences between samples of adults of each sex and rule out factors that might
426 otherwise affect estimated adult allele frequencies in the UK Biobank (*e.g.*, mis-mapping of
427 reads to sex chromosomes, participation biases [55]). Elevations in these metrics thus provide
428 the most compelling evidence for sex-differential selection in the UK Biobank. Moreover, they
429 are consistent with previous observations in post-industrial human populations, including
430 variation in female and male LRS [68] (a necessary precondition for sex-differential selection),
431 widespread sex differences in the genetic basis of quantitative traits (*e.g.*, in the UK Biobank
432 [69]), and sex-differential selection on phenotypes (*e.g.*, height [29,30] and multivariate trait
433 combinations [68]), which should collectively lead to genome-wide polymorphisms with sex-
434 differential effects on fitness and fitness components [20].

435

436 **Distinguishing between SA and SC forms of sex-differential selection**

437 Having established signals of sex-differential selection affecting LRS, we developed a new test
438 for investigating the form of selection—SC or SA—affecting these genomic variants by
439 quantifying the product of a genetic variant's effect on LRS in each sex. Applying our test to
440 UK Biobank data showed that both types of variant contribute to signals of sex-differential
441 selection on LRS, with SC variants contributing comparatively more enrichment in the upper
442 tail of the null of unfolded reproductive $\hat{F}_{ST}$ (and its mixed-model analogue, unfolded $t$) than
443 SA variants contribute in the lower tail of the null. That signals of SC polymorphism were
444 more pronounced than SA polymorphism is perhaps unsurprising, given that most traits are
445 likely to be subject to SC rather than SA selection [29]. Moreover, alleles subject to identical
446 SC selection in each sex will contribute to the upper tail of unfolded reproductive $\hat{F}_{ST}$, but will
447 not contribute to the lower tail (or to other metrics of sex-differential selection), which might
448 also account for greater apparent signal of SC than SA selection in these analyses. Nonetheless,
449 some human traits have been shown to be under SA selection—most notably standing height,
450 which positively covaries with male LRS and negatively covaries with female LRS [28–30].
451 The enrichment of sites in the lower tails of unfolded reproductive $\hat{F}_{ST}$ and unfolded $t$ is
452 consistent with these previous observations. Our finding that variants that increase height
453 tended to have male-beneficial and female-detrimental effects (*i.e.*, as reflected by a negative
454 correlation between height and $t$) is particularly reassuring [28–30], and validates the intuition
455 that SA selection at the phenotypic level (*e.g.*, over height) gives rise to SA selection
456 throughout the genome.
457
458 **Modes of evolution affecting sex-differentiated loci**
459 Sex-differentiated sites had, on average, more intermediate frequencies than less sex-
460 differentiated sites. This finding has several implications. First, we expect no association
461 between metrics of sex-differentiation and MAF in the absence of sex-differential selection.
462 Therefore, this positive association represents an independent strand of support for the

463 argument that sex-differential selection is shaping patterns of genome-wide variation in the UK
464 Biobank dataset. Second, the positive associations imply that a model of sex-differential
465 purifying selection, in which variants are maintained at mutation-selection-drift balance, is
466 inadequate to explain enrichments of sex-differentiated sites. Sex-differential purifying
467 selection is instead expected to generate negative associations between MAF and the extent of
468 sex-differentiation (a negative association that is indeed observed for many quantitative traits
469 [70]). Finally, the positive associations between sex-differentiation and MAF are consistent
470 with recent evolutionary histories of balancing selection, genetic drift or incomplete selective
471 sweeps. Balancing selection or drift can both generate a broad spectrum of allele frequency
472 states at SA loci, in which intermediate-frequency SA variants dominate signals of sex-
473 differential selection. Alternatively, SC alleles with unequal fitness effects in each sex could
474 have recently swept to intermediate frequencies and these variants now dominate signals of
475 sex-differential selection.

476 Although positive associations between metrics of sex-differential selection and MAF
477 indicate that balancing selection may be present, our analyses did not reveal clear signals of
478 *long-term* balancing selection among sex-differentiated sites. The absence of such signals may
479 stem from several factors. First, SA polymorphisms are only predicted to experience balancing
480 selection under narrow conditions [71], so SA loci may not experience balancing selection at
481 all. Second, balancing selection could affect sex-differentiated polymorphisms but be too
482 recent to generate a clear statistical signal in our analyses [72]. Third, long-term balancing
483 selection at sex-differentiated loci may be present but *effectively* weak, owing to relatively
484 small $N_e$ in humans [73] and the high susceptibility of SA alleles to genetic drift [71]. Fourth,
485 long-term balancing selection may be present, but statistical tests for it may be too weak to
486 stand out from the background noise of false positives in our metrics and the datasets used to
487 quantify balancing selection [74].

How do we reconcile these results with previous work in *Drosophila melanogaster* indicating that candidate SA polymorphisms segregate across worldwide populations and even species [33]? A parsimonious explanation for these contrasting findings is that the effectiveness of balancing selection is lower in humans than fruit flies due to much smaller $N_e$. Indeed, given the pronounced sensitivity of SA balancing selection to genetic drift [71,75], we should expect the relationship between signals of SA and balancing selection to vary with $N_e$. Moreover, previous work in *Drosophila melanogaster* focussed on SA polymorphisms [33] to the exclusion of SC polymorphisms, whereas our metrics capture both forms of sex-differential variation, thus further weakening associations with signals of balancing selection in our study compared to theirs. Interestingly, when we partitioned signals of sex-differentiation into SA and SC components, we found indications that candidate SA sites were indeed older, which implies that SA balancing selection may be present but masked by sex-differential SC polymorphisms. Overall, evidence that sex-differentiated, including SA, polymorphisms contribute to standing genetic variation—as in our study—is at present much stronger than evidence that they are maintained by balancing selection.

**Directions for future research**

Our analyses suggest a number of fruitful directions for further research. First, given the difficulty of distinguishing participation bias from selection in signals of between-sex allele frequency differentiation among adults, conclusively establishing the presence of sex-differential viability selection in genomic data remains an important research direction. Parent-offspring trio analyses that control for participation effects [76], or replication of our analysis strategy in large datasets sampled through passive rather than active participation, may yield the evidence required. Second, the extent to which variants with positive effects on mortality in a given sex have similar or opposing effects on reproduction bears further examination. Our finding that genetic correlations between metrics of viability and reproductive selection were

514 not significantly different from zero indicates a range of possible scenarios. It may suggest that
515 variants affecting each fitness component are independent (*i.e.*, because alleles affecting each
516 component are genuinely independent), that between-sex allele frequency differentiation
517 among adults is a poor signal of sex-differential viability selection, or that a similar fraction of
518 loci have concordant and antagonistic effects, thus also generating no net correlation.

519 Finally, given the increasing availability of genotypic and LRS data, further work could
520 attempt to replicate our analysis strategy in different populations and species. Many taxa
521 exhibit greater variance for reproductive success than humans [77], generating higher potential
522 for detecting polygenic signals of sex-differential selection. In line with this, polygenic
523 inflations of adult $\hat{F}_{ST}$ have previously been documented in modest samples of pipefish and
524 flycatchers [32,37,38], suggesting that sex differences in selection might be stronger in those
525 species than in humans. Moreover, these samples are immune to the issues of ascertainment
526 bias mentioned above because sampling can be randomised with respect to sex. While we
527 expect that polygenic signals of sex-differential selection will replicate across populations
528 (provided sample sizes are adequately large [32]), we caution that there may be relatively little
529 overlap in terms of the most sex-differentiated polymorphisms. One reason is that
530 environmental differences between populations (*e.g.*, cultural differences in family planning
531 between human populations) could alter the set of causal loci under sex-differential selection.
532 Another reason is that the noisiness of polygenic signals of sex-differential selection [32,43],
533 along with the near-certainty that most polymorphic loci have small effects on fitness [78],
534 generates variation in the set of candidate sex-differential polymorphisms identified across
535 populations [79], even if causal sex-differential polymorphisms do not differ.

536

537 ## Materials and Methods

538 ### Quality control of UK Biobank data

539 Access to UK Biobank data was granted under project number 52049. We used sample-level
540 information provided by the UK Biobank (see [53] for details) to perform individual-level
541 (phenotypic) quality controls. Specifically, we excluded individuals with high relatedness (3rd
542 degree or closer), non-"White British" ancestry, high heterozygosity and high missing rates.
543 We also excluded individuals whose reported sex did not match their inferred genetic sex,
544 aneuploids, and individuals with missing or unreliable lifetime reproductive success (LRS)
545 data (as detailed below).

546 We processed LRS data as follows. LRS data were obtained from UK Biobank field
547 2405 "Number of children fathered" for males, and field 2734 "Number of live births" for
548 females. Previous observations of positive genetic correlations between offspring and grand-
549 offspring numbers across generations [80] indicate that offspring number represents a good
550 proxy for LRS in post-industrial human populations. Because some individuals were asked to
551 report offspring number at repeated assessment points, we considered the maximum offspring
552 number reported as the definitive value of LRS for that individual. Though mis-estimation of
553 LRS for each individual cannot be definitively excluded (*e.g.*, individuals may mis-report and
554 include non-biological children, individuals may reproduce after data collection), we
555 minimised this possibility by removing individuals: (*i*) younger than 45 years of age (this cut-
556 off was chosen for consistency with previous research [29] and because Office for National
557 Statistics data indicates that reproduction is very limited for UK individuals aged 45 and over);
558 (*ii*) reporting fewer offspring at a later assessment point than at an earlier assessment point;
559 (*iii*) with 20 or more reported offspring numbers (large offspring numbers often ended in
560 zero—*e.g.,* 20, 30, 50, 100—and were thus considered less reliable). Furthermore, uncounted
561 LRS data adds imprecision but should not systematically bias our analyses.

562 In addition to site-level quality-controls implemented by the UK Biobank [53], we used
563 PLINK and PLINK2 [81] to remove imputed sites that were non-diallelic, had MAF<1%,

564 missing rates>5%, p-values<$10^{-6}$ in tests of Hardy-Weinberg equilibrium, and INFO score
565 ≤0.8, denoting poor imputation quality.
566
567 **Additional artefact filtering in UK Biobank data**
568 Mis-mapping of autosomal reads to sex chromosomes can generate between-sex allele
569 frequency differences among adults in the absence of sex differences in selection [32,42]. In
570 light of scant direct evidence for SA polymorphisms in humans, and still-developing
571 bioinformatic methods for distinguishing artefacts from genuine sex-differential
572 selection[39,42,82–84], our primary concern was to reduce the chance of mapping errors. We
573 did so by excluding: (*i*) sites with heterozygosity levels that exceeded what could plausibly be
574 expected under SA selection (see below and *SI Appendix*, Section C); (*ii*) sites with a deficit of
575 minor allele homozygotes; (*iii*) sites exhibiting large differences in missing rate between sexes.
576 These three patterns have previously been shown to correlate with mis-mapping of reads to sex
577 chromosomes [42]. While these filters reduce the chance of false positives, they also potentially
578 increase chance of false negatives and therefore represent a slightly conservative test of sex-
579 differential selection. For example, the removal of sites with high heterozygosity levels is
580 expected to remove sites under strong (but not weak or moderately strong) sex-differential
581 selection; similarly, the removal of sites with large missing rate differences between sexes may
582 remove genuine polymorphisms with sex-differential effects.
583 To remove sites with artificially inflated heterozygosity, we estimated $F_{IS}$ for each SNP,
584 as:
585 $$\hat{F}_{IS} = \frac{P_{Aa}}{2\bar{p}(1-\bar{p})} - 1$$
586 where $P_{Aa}$ denotes the frequency of heterozygotes for a given locus and $\bar{p}$ the sex-averaged
587 allele frequency. For a SA locus at polymorphic equilibrium, the distribution of $\hat{F}_{IS}$ is well
588 approximated by a normal distribution with expectation and variance as follows:

589 $$\mathrm{E}[\hat{F}_{IS}] \approx \frac{1}{2n} + \frac{p(1-p)}{4}\left(\frac{s_{\max}}{1-ps_{\max}}\right)^2$$

590 $$\mathrm{var}[\hat{F}_{IS}] = \frac{1}{n}$$

591 where $n$ is total sample size of adults, $p$ the minor allele frequency, and $s_{\max} = \max(s_m, s_f)$ with
592 $s_m$ and $s_f$ representing male and female selection coefficients (*SI Appendix*, Section C). To
593 identify SNPs with excess heterozygosity, we compared $\hat{F}_{IS}$ in the observed data to expected
594 $\hat{F}_{IS}$ under strong SA selection ($s_{\max}$=0.2) by performing a one-tailed Z-test for excess
595 heterozygosity. We thus obtained p-values for each locus, corrected p-values for multiple
596 testing using Benjamini-Hochberg false discovery rates [85], and removed sites with FDR q-
597 values below 0.05.

598 To identify sites with a deficit of minor allele homozygotes, we compared the observed
599 frequency of minor allele homozygotes to the expected frequency under Hardy-Weinberg
600 equilibrium ($p^2$, where $p$ is the frequency of the minor allele) by performing a one-tailed
601 binomial test, removing sites with FDR q-values below 0.05. Tests for excess heterozygosity
602 and deficits of minor allele homozygotes were performed across all individuals (regardless of
603 sex), and also for each sex separately. Sites were removed if they exhibited q-values below
604 0.05 in any of the three tests (*i.e.*, both sexes combined, females, and males). Finally, to assess
605 differences in missing rate between the sexes, we performed a $\chi^2$ test, removing sites with
606 FDR q-values below 0.05.

607

608 **Quantifying polygenic signals of sex differences in selection**

609 *$\hat{F}_{ST}$-based metrics*

610 We used $\hat{F}_{ST}$ to quantify allele frequency differences between sexes. $\hat{F}_{ST}$ is a simple metric,
611 well-established in evolutionary biology research, amenable to theoretical modelling (as in eqs.
612 [2a-c]), and independent of MAF in the absence of sex-differential selection (unlike, say, raw

allele frequency differences [32]). We obtained allele frequencies in adults of each sex directly from sequence data (after filtering individuals and sites, as described above) and used them to calculate adult $\hat{F}_{ST}$ for each polymorphic site. We obtained allele frequencies among projected gametes using LRS data (as per eq. [1]) and used them to calculate reproductive $\hat{F}_{ST}$, gametic $\hat{F}_{ST}$ and unfolded reproductive $\hat{F}_{ST}$ (as per eqs. [2a-c] and [4]).

*Statistical comparisons of null and observed distributions*

Null distributions for $\hat{F}_{ST}$ metrics were theoretically derived (see *SI Appendix*, Section A, E). The theoretical null distributions apply to genome-wide data in which the sample of female and male sequences, mean and variance in LRS, and Hardy-Weinberg deviations, are constant across loci. In practice, there is variation in sample sizes, mean LRS, variance in LRS, and the extent of Hardy-Weinberg deviations between loci. To take these factors into account, we let the multiplier in eqs. [3a-c] vary in terms of its sample size ($N_{f_i}$ and $N_{m_i}$ per diploid locus *i*), mean and variance in LRS ($\mu_{f_i}$ and $\mu_{m_i}$, and $\sigma^2_{f_i}$ and $\sigma^2_{m_i}$, per diploid locus *i*) and the extent of Hardy-Weinberg deviations in the sample ($\hat{F}^f_{IS_i}$ and $\hat{F}^m_{IS_i}$ per diploid locus *i*). We then scaled $\hat{F}_{ST}$ by the multiplier, such that, for each locus:

$$\hat{F}_{ST(Adult,scaled)} \approx \frac{\hat{F}_{ST(Adult)}}{\left(\frac{1}{8N_{f_i}} + \frac{1}{8N_{m_i}}\right)} \approx X$$

$$\hat{F}_{ST(Reprod.,scaled)} \approx \frac{4\hat{p}(1-\hat{p})\hat{F}_{ST(Reprod.)}}{\frac{\hat{p}_f(1-\hat{p}_f)}{2N_{f_i}}\frac{\sigma^2_{f_i}}{\mu^2_{f_i}}\left(1-\hat{F}^f_{IS_i}\right) + \frac{\hat{p}_m(1-\hat{p}_m)}{2N_{m_i}}\frac{\sigma^2_{m_i}}{\mu^2_{m_i}}\left(1-\hat{F}^m_{IS_i}\right)} \approx X$$

$$\hat{F}_{ST(Gametic,scaled)} \approx \frac{\hat{F}_{ST(Gametic)}}{\left[\frac{1}{8N_{f_i}}\left(1+\frac{\sigma^2_{f_i}}{\mu^2_{f_i}}\right) + \frac{1}{8N_{m_i}}\left(1+\frac{\sigma^2_{m_i}}{\mu^2_{m_i}}\right)\right]} \approx X$$

632 These scaled $\hat{F}_{ST}$ estimates, which correct for site-specific variation, can then be compared to
633 a chi-square distribution with 1 degree of freedom. For unfolded reproductive $\hat{F}_{ST}$, no scaling
634 is required because site-specific adjustments are already taken into consideration in the
635 definition of the metric (eq. 4).

636 Null distributions were also obtained empirically, through permutation, as follows. For
637 adult and gametic $\hat{F}_{ST}$, we performed a single permutation of female and male labels, and re-
638 calculated $\hat{F}_{ST}$ (scaled by the multiplier, as above) in permuted data. For reproductive and
639 unfolded reproductive $\hat{F}_{ST}$, we performed a single permutation of LRS values within each
640 sex—without permuting sex—and re-calculated the statistic (scaled by the multiplier, as above)
641 in permuted data. Permuting LRS without permuting sex is appropriate for reproductive and
642 unfolded reproductive $\hat{F}_{ST}$ because it allows allele frequencies to differ between adult males
643 and females (as would happen if, for example, sex-differential viability selection is occurring
644 among juveniles) but randomises the effects of genotype on LRS, thus ensuring that only
645 estimation error can contribute to the empirical null. We performed a single permutation for
646 each metric because performing large numbers of permutations was computationally
647 unfeasible and because we were focussed on testing a cumulative signal of selection across
648 loci, rather than establishing significance at the single-locus level.

649 To test for elevations in observed data relative to the (theoretical or empirical) nulls,
650 we LD-pruned the dataset (settings '--indep-pairwise 50 10 0.2' in PLINK) and ran Wilcoxon
651 Rank-Sum and Kolmogorov-Smirnov tests. These tests assess differences in the median and
652 distribution of the observed and null data, respectively. As a complementary way of comparing
653 observed and null data, we quantified enrichment of observed values in the top 1% of each null
654 using a $\chi^2$ test. Finally, we estimated the difference between the mean value of the metric in
655 the observed data and the mean value of the metric in each null, obtaining 95% confidence
656 intervals and empirical p-values through bootstrapping (1,000 replicates; where each replicate
657 consists of the set of relevant SNPs, sampled with replacement).

658

659 **Controlling for sex-specific population structure**

660 *Case-control GWAS of sex*

661 To complement the test for sex-differential viability selection based on adult $\hat{F}_{ST}$, we performed
662 a GWAS of sex [32,42,55]. By analogy to adult $\hat{F}_{ST}$, loci with sex-differential effects on
663 viability in a GWAS of sex will tend to have relatively large absolute log-odds ratios
664 (corresponding to relatively large allele frequency differences between sexes). Unlike adult
665 $\hat{F}_{ST}$, the GWAS of sex approach additionally permits the inclusion of covariates that account
666 for population structure and other possible confounders [32].

667 We used BOLT-LMM to run a mixed-model GWAS [86], using a kinship matrix to
668 account for population structure. The kinship matrix was constructed from an LD-pruned set
669 of quality-filtered imputed SNPs (LD-pruning settings as above). We added individual age
670 (field 54), assessment centre (field 21003) and the top 20 principal components derived from
671 the kinship matrix, as fixed-effect covariates. To facilitate comparisons with adult $\hat{F}_{ST}$, we
672 standardised the regression coefficients (log-odds ratios) from the GWAS by allele frequency,
673 such that:

674 $$\hat{\mathcal{L}}_{ST} = (\sqrt{\bar{p}(1-\bar{p})}\hat{\mathcal{L}})^2$$

675 where $\hat{\mathcal{L}}$ is the log-odds ratio and $\bar{p}$ is the sex-averaged allele frequency among adults. To
676 obtain permuted $\hat{\mathcal{L}}_{ST}$ values, we performed a single permutation of female and male labels and
677 re-calculated the statistic in the permuted data.

678

679 *t-statistics for sex-differential effects on LRS*

680 To complement the test for sex differences in selection on LRS based on reproductive $\hat{F}_{ST}$, we
681 performed a GWAS of LRS in each sex separately, using a mixed-model GWAS, which
682 allowed us to correct for population structure in effect size estimates. Following [87], we

683 quantified differences in male and female effect size estimates by means of a *t*-statistic, defined
684 as:

685 $$|t| = \left| \frac{\beta_{FM} - \beta_{FM}}{\sqrt{SE_{MF}^2 + SE_{FM}^2 - 2\rho SE_{MF} SE_{FM}}} \right|$$

686 where $\beta_F$ and $\beta_M$ are estimated effect sizes obtained from sex-stratified GWAS (implemented
687 in BOLT-LMM as above), $SE_M$ and $SE_F$ are sex-specific standard errors, and $\rho$ is the between-
688 sex rank correlation among genome-wide LD-pruned loci. Relative to permuted data (obtained
689 using an identical procedure to that implemented for reproductive $\hat{F}_{ST}$), loci with elevated $|t|$ in
690 observed data denote candidate loci with sex-differential effects on LRS.

691 To examine the relative contributions of SA and SC components to signals of sex-
692 differential reproductive selection, while correcting for sex-specific population structure, we
693 developed an analogue of unfolded reproductive $\hat{F}_{ST}$, termed "unfolded *t*", as:

694 $$\text{Unfolded } t = \frac{\beta_F \beta_M}{\sqrt{SE_F^2 SE_M^2}}$$

695 Relative to permuted data, SA selection (*i.e.*, opposing signs of $\beta_F$ and $\beta_M$) will tend to
696 generate an excess of negative values of unfolded *t*, while SC selection (*i.e.*, same sign of
697 $\beta_M$ and $\beta_F$) will tend to generate an excess of positive values of unfolded *t*.

698

699 **Functions and phenotypic effects of sex-differentiated loci**

700 We used Stratified LD Score Regression [88] to examine whether sex-differentiated loci were
701 more likely to be situated in putatively functional genomic regions (*e.g.*, coding or regulatory
702 regions) than expected by chance. This method partitions the heritability from GWAS
703 summary statistics into different functional categories, while accounting for differences in LD
704 (and thus, increased tagging of a given causal locus) in different regions of the genome (with
705 LD quantified from European-ancestry samples from the 1000 genome project, and restricted
706 to SNPs also present in the HapMap 3 reference panel [88]). Because LD-score regression

707 requires signed summary statistics as input, we first transformed our (unsigned) metrics of sex-
708 differential selection to signed metrics (e.g., $\hat{F}_{ST}$ metrics and $\hat{\mathcal{L}}_{ST}$ were transformed to Z-scores,
709 $|t|$ was transformed to $t$) where positive and negative values denote female- and male-beneficial
710 effects of the focal allele, respectively.

711 Enrichments for three putatively functional categories (coding, 3'UTR, 5'UTR) were
712 then calculated as the fraction of total heritability explained by a given category divided by the
713 fraction of all SNPs in a given category. Note that we calculated enrichment for these categories
714 while implementing the "full baseline model", which includes 50 further categories. This
715 model has been shown to provide unbiased enrichments for focal categories [88] and for total
716 SNP-heritability [89]; estimates of total SNP-heritability were used in *SI Appendix*, Section D.

717 We used cross-trait LD Score Regression [90] to examine genetic correlations between
718 metrics of sex-differential selection and a suite of phenotypic traits, as well as between the
719 metrics of sex-differential selection. The method calculates genetic correlations between pairs
720 of traits while taking into account LD-induced differences in the extent of tagging of causal
721 loci across the genome. We computed genetic correlations between each metric of sex-
722 differential selection (transformed to a signed statistic, as above, such that higher values of the
723 signed metric are more likely to benefit females than males) and an initial list of 43 traits
724 (subsequently filtered to 30 after removing traits where an accurate genetic correlation, defined
725 as SE < 0.2, could not be estimated) [91], and used FDR correction (across metrics and traits)
726 on resulting p-values. These analyses are fully documented (including code and results) at
727 https://lukeholman.github.io/UKBB_LDSC/

728

729 **Modes of evolution affecting sex-differentiated loci**

730 *Associations with MAF in the UK Biobank*

731 To test for associations between metrics of sex differences in selection and MAF, we estimated
732 a Spearman's rank correlation between each metric and MAF. We also tested whether the

733 relationship between metrics of sex differences in selection and MAF was more pronounced in
734 the observed data than in the (theoretical or empirical) nulls by estimating the difference
735 between correlations in the observed data and correlations in the null data among 1,000
736 bootstrap replicates (as above), thereby generating 95% confidence intervals and empirical p-
737 values.
738
739 *Allele ages*
740 If sex-differentiated variants experience sufficiently strong and sustained balancing selection
741 relative to the countervailing effects of genetic drift, we expect them to be older than the
742 genome-wide average [72]. We used the Atlas of Variant Age database to obtain allele age
743 estimates for genome-wide variants [57]. Estimates of allele age in this database apply to the
744 non-reference (*i.e.*, alternative) allele and are derived from coalescent modelling of the time to
745 the most recent common ancestor using the "Genealogical Estimation of Variant Age" method
746 (see [57] for details). Estimates of allele age make use of genomic data from: (*i*) the 1000
747 Genomes Project; (*ii*) the Simons Genome Diversity Project; (*iii*) both datasets combined. For
748 each site in the UK Biobank, we obtained the median estimate of allele age from the combined
749 dataset (when available), from the 1000 Genomes Project, or the Simons Genome Diversity
750 Project (when neither alternative estimate was available).
751
752 *Between-population $F_{ST}$ and Tajima's D in non-European populations*
753 If candidate SA variants experience sufficiently strong balancing selection maintaining a fixed
754 polymorphic equilibrium, they should exhibit lower-than-average allele frequency differences
755 *between* populations [72], and larger-than-average allele frequency diversity *within*
756 populations. We used bcftools [92] to obtain allele frequency data from two non-European
757 populations from the 1000 Genomes Project: Yoruba Nigerians (YRI, N=108) and Gujarati
758 Indians (GIH, N=103). We then estimated between-population $\hat{F}_{ST}$ as:

759 $$\hat{F}_{ST} = \frac{(\hat{p}_1 - \hat{p}_2)^2}{4\bar{p}(1-\bar{p})}$$

760 where $\hat{p}_1$ and $\hat{p}_2$ are allele frequency estimates in the relevant pair of populations, and $\bar{p} =$
761 $(\hat{p}_1 + \hat{p}_2)/2$. We also used vcftools [93] to calculate Tajima's D, a metric of genetic diversity
762 which takes on elevated values under certain evolutionary and demographic scenarios,
763 including balancing selection, in 10kb windows across the genome.
764
765 *Previous candidates for balancing selection*
766 If candidate SA variants experience strong balancing selection, they should disproportionately
767 co-occur with previously identified candidates for balancing selection. We used three
768 independent sets of candidate sites for balancing selection to investigate this possibility [59–
769 61]: (*i*) the dataset of Andrés et al. [59], which consists of 64 genes exhibiting elevated
770 polymorphism (as determined using the Hudson-Kreitman-Aguadé test) and/or intermediate-
771 frequency alleles across 19 African-American or 20 European-American individuals; (*ii*) the
772 dataset of DeGiorgio et al. [61], which consists of 400 candidate genes exhibiting elevated $T_1$
773 or $T_2$ statistics among 9 European (CEU) and 9 African (YRI) individuals. $T_1$ or $T_2$ statistics
774 quantify the likelihood that a genomic region exhibits levels of neutral polymorphism that are
775 consistent with a linked balanced polymorphism; (*iii*) the dataset of Bitarello et al. [60], which
776 consists of 1859 candidate genes exhibiting elevated values of "Non-Central Deviation" (NCD)
777 statistics. NCD statistics also quantify the likelihood that given genomic regions are situated
778 nearby a balanced polymorphism, using polymorphism data from 50 random individuals from
779 two African (YRI; LWK) and European (GBR; TSI) populations and divergence data from a
780 chimpanzee outgroup.
781 We assigned each site in the UK Biobank dataset to a gene using SnpEff [94] and
782 categorised sites as candidates or non-candidates for balancing selection based on whether they

783 were annotated as belonging to a candidate or non-candidate gene in each of the three
784 aforementioned datasets.
785
786 *Statistical associations between metrics of sex-differential selection and balancing selection*
787 To test whether signals of sex differences in selection were associated with signals of balancing
788 selection, we performed Spearman's rank correlations between alternative allele age (scaled
789 by alternative allele frequency in the UK Biobank, to control for ascertainment bias) and each
790 metric of sex differences in selection. For between-population $\hat{F}_{ST}$ and Tajima's D, we
791 performed multiple linear regressions, with the relevant metric of sex differences in selection
792 as the independent variable and MAF in the UK Biobank as a fixed-effect covariate (to control
793 for ascertainment bias). For previous candidates for balancing selection, we performed multiple
794 logistic regressions, where candidate/non-candidate status was the binary response variable,
795 with the relevant metric of sex differences in selection as the independent variable and MAF
796 in the UK Biobank as a fixed-effect covariate. In the case of regressions involving between-
797 population $\hat{F}_{ST}$, $\hat{F}_{ST}$ was first log-transformed to meet normality assumptions. In the case of
798 Tajima's D analyses (which are window-based rather than SNP-based), we averaged
799 independent variables across 10kb windows before performing regressions.

800 **References**


801 1. Fisher RA. The Genetical Theory of Natural Selection. Genetics. Clarendon Press;
802 1930. doi:10.1038/158453a0
803 2. Charlesworth B. Causes of natural variation in fitness: evidence from studies of
804 *Drosophila* populations. Proc Natl Acad Sci U S A. 2015;112: 1662–1629.
805 doi:10.1073/pnas.1423275112
806 3. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 years
807 of GWAS discovery: Biology, function, and translation. Am J Hum Genet. 2017;101:
808 5–22. doi:10.1016/j.ajhg.2017.06.005
809 4. Walsh B, Blows MW. Abundant Genetic Variation + Strong Selection = Multivariate
810 Genetic Constraints: A Geometric View of Adaptation. Annu Rev Ecol Evol Syst.
811 2009;40: 41–59. doi:10.1146/annurev.ecolsys.110308.120232
812 5. Johnston SE, Gratten J, Berenos C, Pilkington JG, Clutton-Brock TH, Pemberton JM,
813 et al. Life history trade-offs at a single locus maintain sexually selected genetic
814 variation. Nature. 2013;502: 93–95. doi:10.1038/nature12489
815 6. Connallon T, Hall MD. Genetic constraints on adaptation: a theoretical primer for the
816 genomics era. Ann N Y Acad Sci. 2018;1422: 65–87. doi:10.1111/nyas.13536
817 7. Arnqvist G, Tuda M. Sexual conflict and the gender load: Correlated evolution
818 between population fitness and sexual dimorphism in seed beetles. Proc R Soc B Biol
819 Sci. 2010;277: 1345–1352. doi:10.1098/rspb.2009.2026
820 8. Whitlock MC, Agrawal AF. Purging the genome with sexual selection: Reducing
821 mutation load through selection on males. Evolution. 2009;63: 569–582.
822 doi:10.1111/j.1558-5646.2008.00558.x
823 9. Gilks WP, Abbott JK, Morrow EH. Sex differences in disease genetics: Evidence,
824 evolution, and detection. Trends Genet. 2014;30: 453–463.
825 doi:10.1016/j.tig.2014.08.006

826 10. Kidwell JF, Clegg MT, Stewart FM, Prout T. Regions of stable equilibria for models
827 of differential selection in the two sexes under random mating. Genetics. 1977;85:
828 171–183.
829 11. Curtsinger JW, Service PM, Prout T. Antagonistic pleiotropy, reversal of dominance,
830 and genetic polymorphism. Am Nat. 1994;144: 210–228. doi:10.1086/285671
831 12. Charlesworth B, Charlesworth D. A model for the evolution of dioecy and gynodioecy.
832 Am Nat. 1978;112: 975–997. doi:10.1086/283342
833 13. Olito C, Connallon T. Sexually antagonistic variation and the evolution of dimorphic
834 sexual systems. Am Nat. 2019;193: 688–701. doi:10.1086/702847
835 14. Van Doorn GS, Kirkpatrick M. Turnover of sex chromosomes induced by sexual
836 conflict. Nature. 2007;449: 909–912. doi:10.1038/nature06178
837 15. Charlesworth D, Charlesworth B. Sex differences in fitness and selection for centric
838 fusions between sex-chromosomes and autosomes. Genet Res. 1980;35: 205–214.
839 doi:10.1017/S0016672300014051
840 16. Jordan CY, Charlesworth D. The potential for sexually antagonistic polymorphism in
841 different genome regions. Evolution. 2012;66: 505–516. doi:10.1111/j.1558-
842 5646.2011.01448.x
843 17. Kirkpatrick M, Guerrero RF. Signatures of sex-antagonistic selection on recombining
844 sex chromosomes. Genetics. 2014;197: 531–541. doi:10.1534/genetics.113.156026
845 18. Patten MM, Haig D. Maintenance or loss of genetic variation under sexual and
846 parental antagonism at a sex-linked locus. Evolution. 2009;63: 2888–2895.
847 doi:10.1111/j.1558-5646.2009.00764.x
848 19. Bonduriansky R, Chenoweth SF. Intralocus sexual conflict. Trends Ecol Evol.
849 2009;24: 280–288. doi:10.1016/j.tree.2008.12.005
850 20. Connallon T, Clark AG. Evolutionary inevitability of sexual antagonism. Proc R Soc B
851 Biol Sci. 2014;281: 20132123. doi:10.1098/rspb.2013.2123

852 21. Cox RM, Calsbeek R. Sexually antagonistic selection, sexual dimorphism, and the
853 resolution of intralocus sexual conflict. Am Nat. 2009;173: 176–187.
854 doi:10.1086/595841
855 22. Singh A, Punzalan D. The strength of sex-specific selection in the wild. Evolution.
856 2018;72: 2818–2824. doi:10.1111/evo.13625
857 23. De Lisle SP, Goedert D, Reedy AM, Svensson EI. Climatic factors and species range
858 position predict sexually antagonistic selection across taxa. Philos Trans R Soc B Biol
859 Sci. 2018;373: 20170415. doi:10.1098/rstb.2017.0415
860 24. Mackay TFC. The genetic architecture of quantitative traits. Annu Rev Genet.
861 2001;35: 303–339. doi:10.1146/annurev.genet.35.102401.090633
862 25. Delph LF, Andicoechea J, Steven JC, Herlihy CR, Scarpino S V., Bell DL.
863 Environment-dependent intralocus sexual conflict in a dioecious plant. New Phytol.
864 2011;192: 542–552. doi:10.1111/j.1469-8137.2011.03811.x
865 26. Svensson EI, McAdam AG, Sinervo B. Intralocus sexual conflict over immune
866 defense, gender load, and sex-specific signaling in a natural lizard population.
867 Evolution. 2009;63: 3124–3135. doi:10.1111/j.1558-5646.2009.00782.x
868 27. Chippindale AK, Gibson JR, Rice WR. Negative genetic correlation for adult fitness
869 between sexes reveals ontogenetic conflict in *Drosophila*. Proc Natl Acad Sci U S A.
870 2001;98: 1671–1675. doi:10.1073/pnas.98.4.1671
871 28. Stulp G, Kuijper B, Buunk AP, Pollet T V, Verhulst S. Intralocus sexual conflict over
872 human height. Biol Lett. 2012;8: 976–8. doi:10.1098/rsbl.2012.0590
873 29. Sanjak JS, Sidorenko J, Robinson MR, Thornton KR, Visscher PM. Evidence of
874 directional and stabilizing selection in contemporary humans. Proc Natl Acad Sci U S
875 A. 2017;115: 151–156. doi:10.1073/pnas.1707227114
876 30. Stearns SC, Govindaraju DR, Ewbank D, Byars SG. Constraints on the coevolution of
877 contemporary human males and females. Proc R Soc B Biol Sci. 2012;279: 4836–

878 4844. doi:10.1098/rspb.2012.2024
879 31. Harper JA, Janicke T, Morrow EH. Systematic review reveals multiple sexually
880 antagonistic polymorphisms affecting human disease and complex traits. Evolution.
881 2021;75: 3087–3097. doi:10.1111/evo.14394
882 32. Ruzicka F, Dutoit L, Czuppon P, Jordan CY, Li X, Olito C, et al. The search for
883 sexually antagonistic genes: Practical insights from studies of local adaptation and
884 statistical genomics. Evol Lett. 2020;4: 398–415. doi:10.1002/evl3.192
885 33. Ruzicka F, Hill MS, Pennell TM, Flis I, Ingleby FC, Mott R, et al. Genome-wide
886 sexually antagonistic variants reveal long-standing constraints on sexual dimorphism
887 in fruit flies. PLoS Biol. 2019;17: e3000244. doi:10.1371/journal.pbio.3000244
888 34. Cheng C, Kirkpatrick M. Sex-specific selection and sex-biased gene expression in
889 humans and flies. PLoS Genet. 2016;12: e1006170. doi:10.1371/journal.pgen.1006170
890 35. Lucotte EA, Laurent R, Heyer E, Ségurel L, Toupance B. Detection of allelic
891 frequency differences between the sexes in humans: a signature of sexually
892 antagonistic selection. Genome Biol Evol. 2016;8: 1489–1500.
893 doi:10.1093/gbe/evw090
894 36. Wright AE, Fumagalli M, Cooney CR, Bloch NI, Vieira FG, Buechel SD, et al. Sex-
895 biased gene expression resolves sexual conflict through the evolution of sex-specific
896 genetic architecture. Evol Lett. 2018;2: 52–61. doi:10.1101/176990
897 37. Flanagan SP, Jones AG. Genome-wide selection components analysis in a fish with
898 male pregnancy. Evolution. 2017;71: 1096–1105. doi:10.1111/evo.13173
899 38. Dutoit L, Mugal CF, Bolívar P, Wang M, Nadachowska-Brzyska K, Smeds L, et al.
900 Sex-biased gene expression, sexual antagonism and levels of genetic diversity in the
901 collared flycatcher (*Ficedula albicollis*) genome. Mol Ecol. 2018;27: 3572–3581.
902 doi:10.1111/mec.14789
903 39. Bissegger M, Laurentino TG, Roesti M, Berner D. Widespread intersex differentiation

904 across the stickleback genome – The signature of sexually antagonistic selection? Mol
905 Ecol. 2019;29: 262–271. doi:10.1111/mec.15255
906 40. Mank JE, Shu JJ, Wright AE. Signature of sexual conflict is actually conflict resolved.
907 Mol Ecol. 2020. doi:10.1111/mec.15311
908 41. Cheng C, Kirkpatrick M. The signal of sex-specific selection in humans is not an
909 artefact: Reply to Mank et al. Mol Ecol. 2020. doi:10.1111/mec.15420
910 42. Kasimatis KR, Abraham A, Ralph PL, Kern AD, Capra JA, Phillips PC. Evaluating
911 human autosomal loci for sexually antagonistic viability selection in two large
912 biobanks. Genetics. 2021;217: 1–10. doi:10.1093/genetics/iyaa015
913 43. Kasimatis KR, Ralph PL, Phillips PC. Limits to genomic divergence under sexually
914 antagonistic selection. G3. 2019;9: 3813–3824. doi:10.1534/g3.119.400711
915 44. Gardner EJ, Neville MDC, Samocha KE, Barclay K, Kolk M, Niemi MEK, et al.
916 Reduced reproductive success is associated with selective constraint on human genes.
917 Nature. 2022;603: 858–863. doi:10.1038/s41586-022-04549-9
918 45. Mank JE. Population genetics of sexual conflict in the genomic era. Nat Rev Genet.
919 2017;18: 721–730. doi:10.1038/nrg.2017.83
920 46. Rowe L, Chenoweth SF, Agrawal AF. The genomics of sexual conflict. Am Nat.
921 2018;192: 274–286. doi:10.1086/698198
922 47. Kasimatis KR, Nelson TC, Phillips PC. Genomic signatures of sexual conflict. J
923 Hered. 2017;108: 780–790. doi:10.1093/jhered/esx080
924 48. Roberts RB, Ser JR, Kocher TD. Sexual conflict resolved by invasion of a novel sex
925 determiner in Lake Malawi cichlid fishes. Science. 2009;326: 998–1001.
926 doi:10.1126/science.1174705
927 49. Rostant WG, Kay C, Wedell N, Hosken DJ. Sexual conflict maintains variation at an
928 insecticide resistance locus. BMC Biol. 2015;13: 34. doi:10.1186/s12915-015-0143-3
929 50. Barson NJ, Aykanat T, Hindar K, Baranski M, Bolstad GH, Fiske P, et al. Sex-

930 dependent dominance at a single locus maintains variation in age at maturity in
931 salmon. Nature. 2015;528: 405–408. doi:10.1038/nature16062
932 51. VanKuren NW, Long M. Gene duplicates resolving sexual conflict rapidly evolved
933 essential gametogenesis functions. Nat Ecol Evol. 2018;2: 705–712.
934 doi:10.1038/s41559-018-0471-0
935 52. Pearse DE, Barson NJ, Nome T, Gao G, Campbell MA, Abadía-Cardoso A, et al. Sex-
936 dependent dominance maintains migration supergene in rainbow trout. Nat Ecol Evol.
937 2019;3: 1731–1742. doi:10.1038/s41559-019-1044-6
938 53. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK
939 Biobank resource with deep phenotyping and genomic data. Nature. 2018;562: 203–
940 209. doi:10.1038/s41586-018-0579-z
941 54. Wright S. The genetical structure of species. Ann Eugen. 1951;15: 323–354.
942 55. Pirastu N, Cordioli M, Nandakumar P, Mignogna G, Abdellaoui A, Hollis B, et al.
943 Genetic analyses identify widespread sex-differential participation bias. bioRxiv. 2020.
944 56. Eyre-Walker A. Genetic architecture of a complex trait and its implications for fitness
945 and genome-wide association studies. Proc Natl Acad Sci U S A. 2010;107: 1752–
946 1756. doi:10.1073/pnas.0906182107
947 57. Albers PK, McVean G. Dating genomic variants and shared ancestry in population-
948 scale sequencing data. PLoS Biol. 2020;18. doi:10.1371/journal.pbio.3000586
949 58. The 1000 Genomes Project Consortium. A global reference for human genetic
950 variation. Nature. 2015;526: 68–74. doi:10.1038/nature15393
951 59. Andrés AM, Hubisz MJ, Indap A, Torgerson DG, Degenhardt JD, Boyko AR, et al.
952 Targets of balancing selection in the human genome. Mol Biol Evol. 2009;26: 2755–
953 2764. doi:10.1093/molbev/msp190
954 60. Bitarello BD, de Filippo C, Teixeira JC, Schmidt JM, Kleinert P, Meyer D, et al.
955 Signatures of long-term balancing selection in human genomes. Genome Biol Evol.

956 2018;10: 939–955. doi:10.1093/gbe/evy054
957 61. DeGiorgio M, Lohmueller KE, Nielsen R. A model-based approach for identifying
958 signatures of ancient balancing selection in genetic data. PLoS Genet. 2014;10:
959 e1004561. doi:10.1371/journal.pgen.1004561
960 62. Morrissey MB. Meta-analysis of magnitudes, differences and variation in evolutionary
961 parameters. J Evol Biol. 2016;29: 1882–1904. doi:10.1111/jeb.12950
962 63. Hoban S, Kelley JL, Lotterhos KE, Antolin MF, Bradburd G, Lowry DB, et al. Finding
963 the genomic basis of local adaptation: Pitfalls, practical solutions, and future
964 directions. Am Nat. 2016;188: 379–397. doi:10.1086/688018
965 64. Vitti JJ, Grossman SR, Sabeti PC. Detecting natural selection in genomic data. Annu
966 Rev Genet. 2013;47: 97–120. doi:10.1146/annurev-genet-111212-133526
967 65. Booker TR, Jackson BC, Keightley PD. Detecting positive selection in the genome.
968 BMC Biol. 2017;15. doi:10.1186/s12915-017-0434-y
969 66. Pirastu N, Cordioli M, Nandakumar P, Mignogna G, Abdellaoui A, Hollis B, et al.
970 Genetic analyses identify widespread sex-differential participation bias. Nat Genet.
971 2021;53: 663–671. doi:10.1038/s41588-021-00846-7
972 67. Benagiano G, Farris M, Grudzinskas G. Fate of fertilized human oocytes. Reprod
973 Biomed Online. 2010;21: 732–741. doi:10.1016/j.rbmo.2010.08.011
974 68. Stearns SC, Byars SG, Govindaraju DR, Ewbank D. Measuring selection in
975 contemporary human populations. Nat Rev Genet. 2010;11: 611–622.
976 doi:10.1038/nrg2831
977 69. Bernabeu E, Canela-Xandri O, Rawlik K, Talenti A, Prendergast J, Tenesa A. Sex
978 differences in genetic architecture in UK Biobank. Nat Genet. 2021;53: 1283–1289.
979 doi:10.1101/2020.07.20.211813
980 70. Zeng J, De Vlaming R, Wu Y, Robinson MR, Lloyd-Jones LR, Yengo L, et al.
981 Signatures of negative selection in the genetic architecture of human complex traits.

982 Nat Genet. 2018;50: 746–753. doi:10.1038/s41588-018-0101-4
983 71. Connallon T, Clark AG. A general population genetic framework for antagonistic
984 selection that accounts for demography and recurrent mutation. Genetics. 2012;190:
985 1477–1489. doi:10.1534/genetics.111.137117
986 72. Fijarczyk A, Babik W. Detecting balancing selection in genomes: Limits and
987 prospects. Mol Ecol. 2015;24: 3529–3545. doi:10.1111/mec.13226
988 73. Li H, Durbin R. Inference of human population history from individual whole-genome
989 sequences. Nature. 2011;475: 493–496. doi:10.1038/nature10231
990 74. Connallon T, Clark AG. Antagonistic versus nonantagonistic models of balancing
991 selection: characterizing the relative timescales and hitchhiking effects of partial
992 selective sweeps. Evolution. 2013;67: 908–917. doi:10.1111/j.1558-
993 5646.2012.01800.x
994 75. Mullon C, Pomiankowski A, Reuter M. The effects of selection and genetic drift on
995 the genomic distribution of sexually antagonistic alleles. Evolution. 2012;66: 3743–
996 3753. doi:10.1111/j.1558-5646.2012.01728.x
997 76. Lucotte EA, Albiñana C, Laurent R, Bhérer C, Bataillon T, Toupance B. Detection of
998 sexually antagonistic transmission distortions in trio datasets. Evol Lett. 2022.
999 doi:10.1002/evl3.271
1000 77. Janicke T, Häderer IK, Lajeunesse MJ, Anthes N. Darwinian sex roles confirmed
1001 across the animal kingdom. Sci Adv. 2016;2: e1500983. doi:10.1126/sciadv.1500983
1002 78. Eyre-Walker A, Keightley PD. The distribution of fitness effects of new mutations.
1003 Nat Rev Genet. 2007;8: 610–618. doi:10.1038/nrg2146
1004 79. Yeaman S. Local adaptation by alleles of small effect. Am Nat. 2015;186: S74–S89.
1005 doi:10.1086/682405
1006 80. Zietsch BP, Kuja-Halkola R, Walum H, Verweij KJH. Perfect genetic correlation
1007 between number of offspring and grandoffspring in an industrialized human

population. Proc Natl Acad Sci U S A. 2014;111: 1032–1036. doi:10.1073/pnas.1310058111

81. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: Rising to the challenge of larger and richer datasets. Gigascience. 2015;4. doi:10.1186/s13742-015-0047-8

82. McKinney GJ, Waples RK, Seeb LW, Seeb JE. Paralogs are revealed by proportion of heterozygotes and deviations in read ratios in genotyping-by-sequencing data from natural populations. Mol Ecol Resour. 2017;17: 656–669. doi:10.1111/1755-0998.12613

83. Dou J, Zhao X, Fu X, Jiao W, Wang N, Zhang L, et al. Reference-free SNP calling: Improved accuracy by preventing incorrect calls from repetitive genomic regions. Biol Direct. 2012;7. doi:10.1186/1745-6150-7-17

84. Tsai KL, Evans JM, Noorai RE, Starr-Moss AN, Clark LA. Novel Y chromosome retrocopies in canids revealed through a genome-wide association study for sex. Genes (Basel). 2019;10: 320. doi:10.3390/genes10040320

85. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc. 1995;57: 289–300. doi:10.2307/2346101

86. Loh P-R, Kichaev G, Gazal S, Schoech A, Price A. Mixed model association for biobank-scale data sets. Nat Genet. 2018;50: 906–908. doi:10.1101/194944

87. Randall JC, Winkler TW, Kutalik Z, Berndt SI, Jackson AU, Monda KL, et al. Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. PLoS Genet. 2013;9: e1003500. doi:10.1371/journal.pgen.1003500

88. Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh PR, et al. Partitioning heritability by functional annotation using genome-wide association

1034 summary statistics. Nat Genet. 2015;47: 1228–1235. doi:10.1038/ng.3404
1035 89. Evans LM, Tahmasbi R, Vrieze SI, Abecasis GR, Das S, Gazal S, et al. Comparison of
1036 methods that use whole genome data to estimate the heritability and genetic
1037 architecture of complex traits. Nat Genet. 2018;50: 737–745. doi:10.1038/s41588-018-
1038 0108-x
1039 90. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An atlas
1040 of genetic correlations across human diseases and traits. Nat Genet. 2015;47: 1236–
1041 1241. doi:10.1038/ng.3406
1042 91. http://www.nealelab.is/uk-biobank/.
1043 92. Li H. A statistical framework for SNP calling, mutation discovery, association
1044 mapping and population genetical parameter estimation from sequencing data.
1045 Bioinformatics. 2011;27: 2987–2993. doi:10.1093/bioinformatics/btr509
1046 93. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The
1047 variant call format and VCFtools. Bioinformatics. 2011;27: 2156–2158.
1048 doi:10.1093/bioinformatics/btr330
1049 94. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for
1050 annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly
1051 (Austin). 2012;6: 80–92. doi:10.4161/fly.19695
1052

1053 **Acknowledgements**

1054 We thank Tom Allison, Isobel Beasley, Isobel Booksmythe, Katja Kasimatis, Max Reuter,
1055 Daniel Berner and six anonymous reviewers for helpful comments on previous versions of the
1056 manuscript. This work was supported by the Monash Bioinformatics Platform, the UK Biobank
1057 (project number 52049 to FR and TC), and funds from the Australian Research Council (to
1058 TC).
1059

## Author contributions

Conceptualization: FR, LH, TC; Methodology: FR, LH, TC; Investigation: FR, LH, TC; Visualization: FR, LH, TC; Funding acquisition: TC; Project administration: FR, TC; Supervision: FR, LH, TC; Writing – original draft: FR, TC; Writing – review & editing: FR, LH, TC.

## Competing Interests

Authors declare no competing interests.

## Data accessibility

Within the restrictions of the UK Biobank MTA agreement, metadata will be made available in the Supplementary Materials. Further metadata will be made available on public data sharing platforms, and code will be made available on public github repositories upon publication (/filruzicka/polygenic_SA_selection_in_the_UK_biobank/ and /lukeholman/ukbb_sex_fst/).

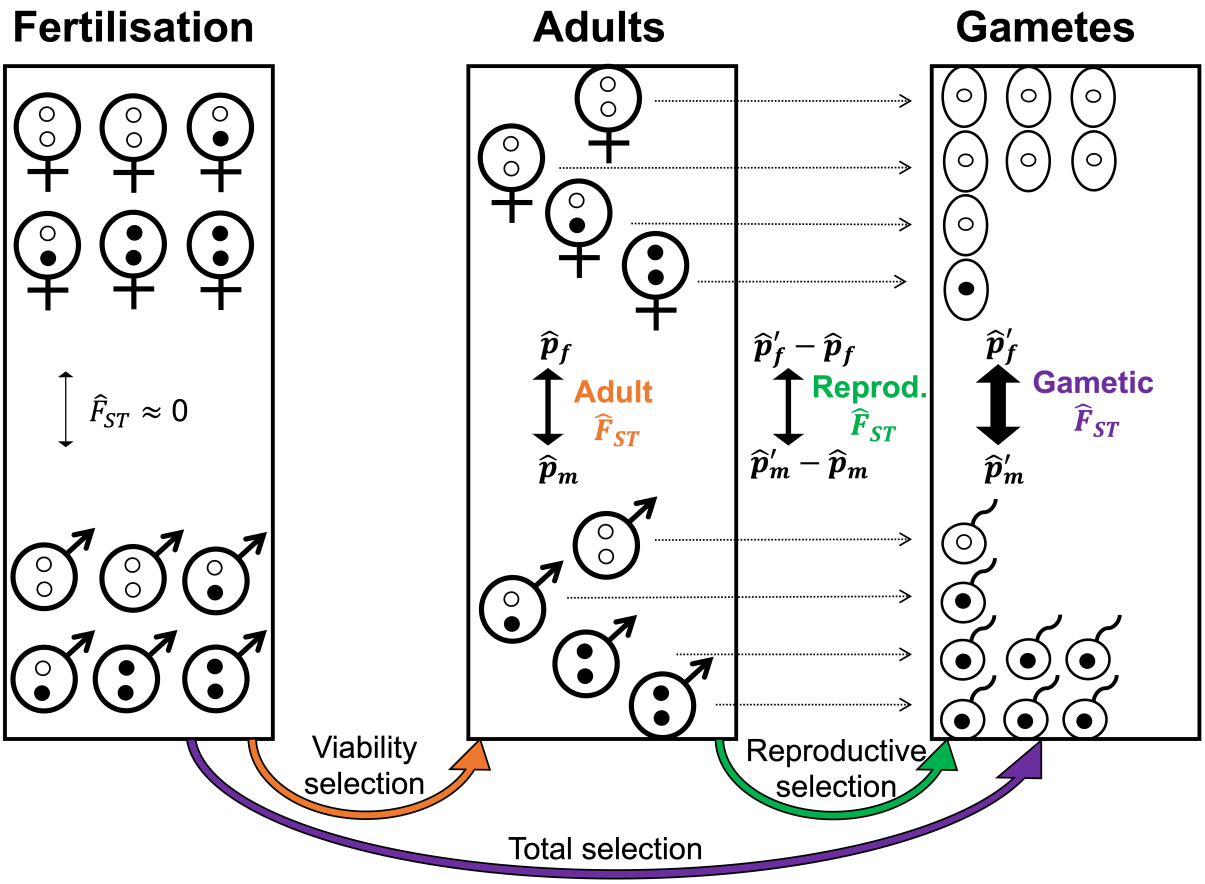



**Figure 1. Partitioning signals of sex differences in selection among fitness components.** A pair of autosomal alleles are represented by white and black dots, representing female- and male-beneficial alleles, respectively; $\hat{p}_f$, $\hat{p}_m$, $\hat{p}'_f$, and $\hat{p}'_m$ depict sex-specific frequency estimates for a given allele at different stages of the life cycle (see main text for details). Autosomal allele frequencies are equalised between sexes at fertilisation (left box; females, top; males, bottom), resulting in negligible allele frequency differentiation at this stage of the life cycle. Differentiation between sexes can arise in the sample of adults (middle box) due to sex differences in viability selection among juveniles (orange arrow), and in the projected gametes (right box) due to sex differences in lifetime reproductive success (LRS) among adults (green arrow). Data on sex-specific allele frequencies and LRS thus allow the estimation of sex-differential effects of genetic variants on each fitness component (including overall fitness; purple arrow), despite the absence of allele frequency data among zygotes (left box) and gametes (right box), which are inferred and not directly observed.

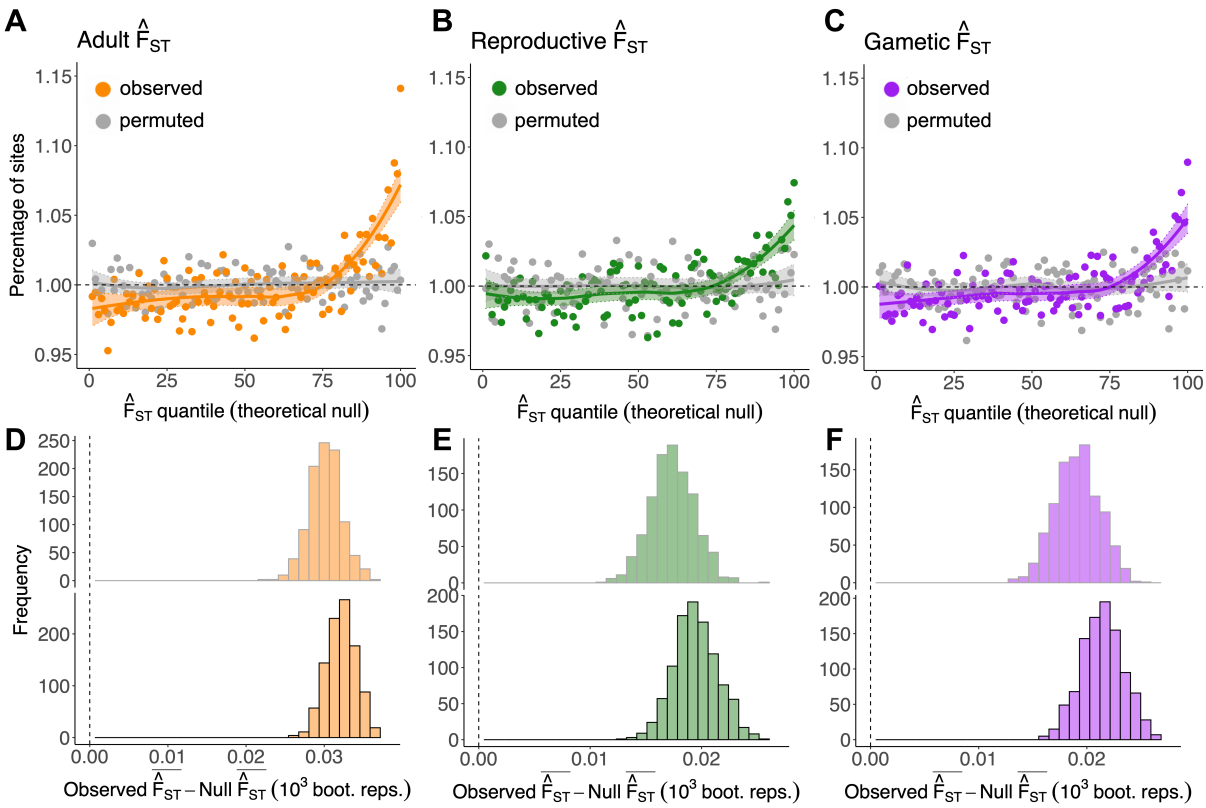



**Figure 2. Polygenic signals of sex-differential selection: inflation in $\hat{F}_{ST}$ metrics relative to their nulls. A-C**. Percentage of sites (coloured, observed; grey, permuted) falling into each of 100 quantiles of the theoretical null distributions of adult $\hat{F}_{ST}$ (A), reproductive $\hat{F}_{ST}$ (B) and gametic $\hat{F}_{ST}$ (C). Theoretical null data (*x*-axes) were generated by simulating values (nSNPs = 1,051,949) from a chi-square distribution with one degree of freedom. For each locus, observed and permuted $\hat{F}_{ST}$ values were scaled by the multiplier of the relevant theoretical null distributions (*i.e.*, the multiplier in eqs. [3a-c] for adult, reproductive and gametic $\hat{F}_{ST}$, respectively; see Materials and Methods). In the absence of sex differences in selection, ~1% of observed SNPs should fall into each quantile of the null (dashed line). LOESS curves (±SE) are presented for visual emphasis. **D-F.** Difference between the mean of observed and empirical null data for each metric (*i.e.*, adult, reproductive and gametic $\hat{F}_{ST}$, respectively) (top), and the difference between observed and theoretical null data (bottom), across 1,000 bootstrap replicates. Vertical line intersects zero (no difference between observed and null data). As in panels A-C., $\hat{F}_{ST}$ values were scaled by the relevant theoretical null distributions.

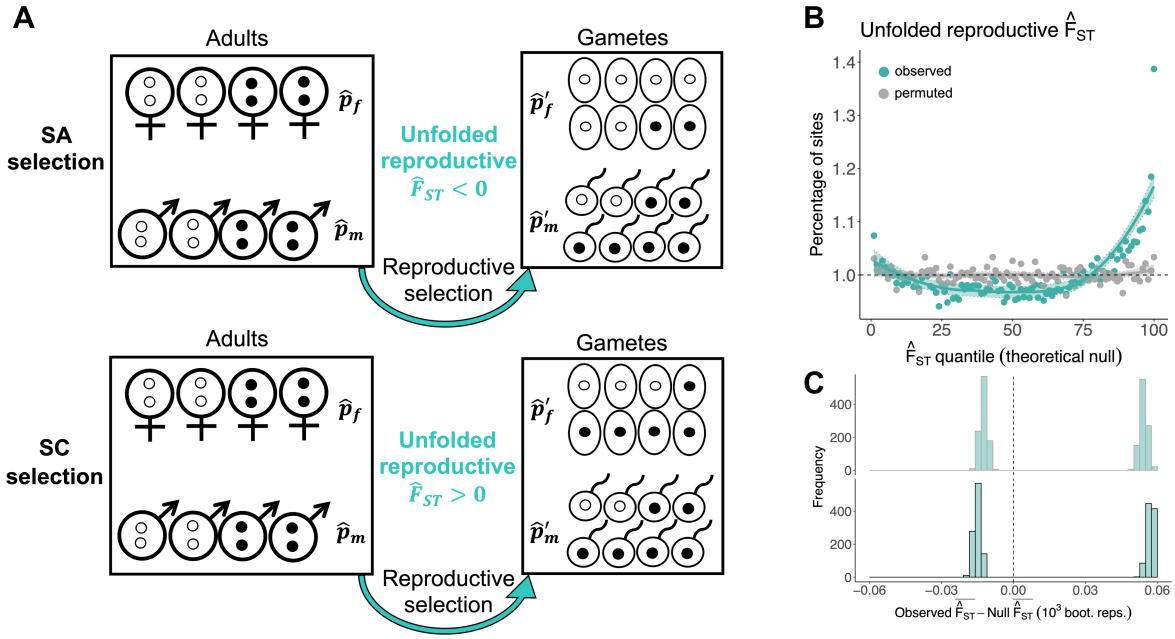
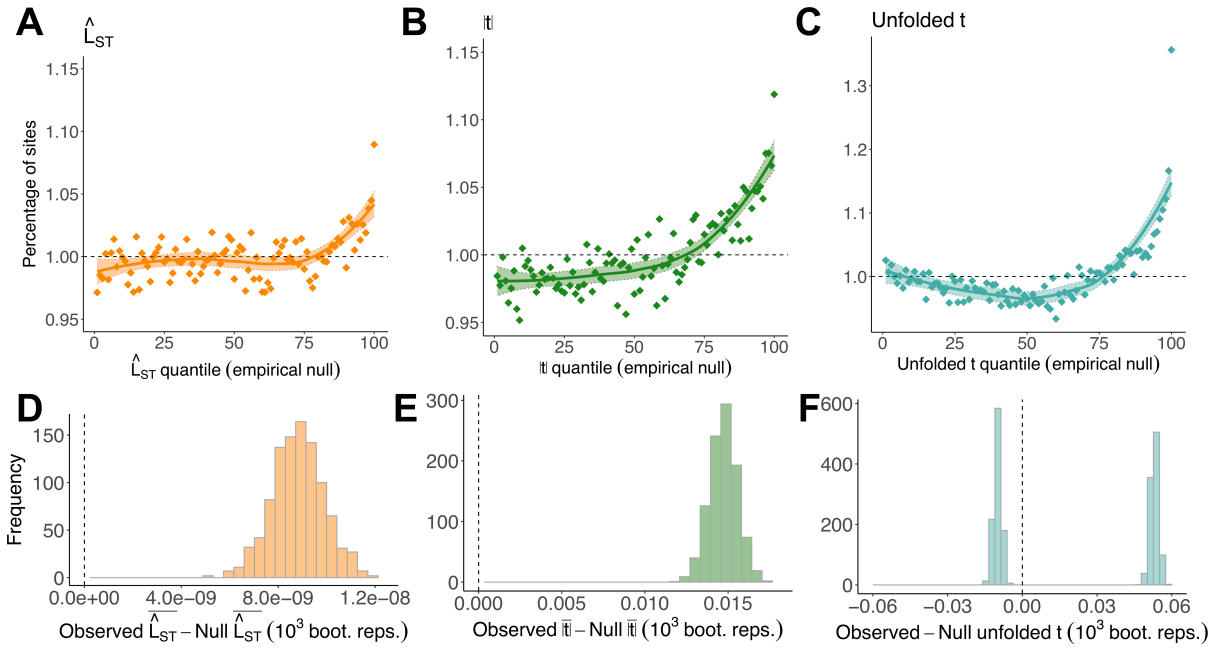



**Figure 3. Partitioning signals of sex-differential selection into SA and SC components reveals their joint contributions. A.** As in Fig. 1, $\hat{p}_f$, $\hat{p}_m$, $\hat{p}'_f$, and $\hat{p}'_m$ depict sex-specific frequency estimates for a given allele at different stages of the life cycle. Under SA selection (top), the white allele is female-beneficial and the black allele is male-beneficial, which tends to generate negative values of unfolded reproductive $\hat{F}_{ST}$. Under SC selection (bottom), the black allele is beneficial in both sexes, which tends to generate positive values of unfolded reproductive $\hat{F}_{ST}$. **B.** Percentage of sites (turquoise: observed; grey: permuted) falling into each of 100 quantiles of the theoretical null distributions of unfolded reproductive $\hat{F}_{ST}$. Theoretical null data (*x*-axes) were generated by simulating values (nSNPs = 1,051,949) from the null (*i.e.*, the product of two standard normal distributions). In the absence of sex-differential selection, ~1% of observed SNPs should fall into each quantile of the null (dashed line). LOESS curves (±SE) are presented for visual emphasis. **C.** Difference, for unfolded reproductive $\hat{F}_{ST}$, between the mean observed and empirical null data (top) and between observed and theoretical null data (bottom), across 1,000 bootstrap replicates. The vertical line intersects zero, indicating no difference between the observed and null data. Differences between observed and null data

1120 were obtained separately for negative and positive values of unfolded reproductive $\hat{F}_{ST}$. This

1121 illustrates that there is enrichment of SNPs in both tails of the null.




1122

1123 **Figure 4. Structure-corrected metrics reaffirm $\widehat{F}_{ST}$-based signals of sex-differential**
1124 **selection. A-C.** Percentage of sites falling into each of 100 quantiles of the empirical null
1125 distributions of $\hat{\mathcal{L}}_{ST}$, $|t|$ and unfolded $t$. In the absence of sex differences in selection, ~1% of
1126 observed SNPs should fall into each quantile of the null (dashed line). LOESS curves (±SE)
1127 are presented for visual emphasis. **D-F.** Difference between the mean of each metric in
1128 observed and empirical null data, across 1,000 bootstrap replicates. Vertical line intersects zero
1129 (no difference between observed and null data). For unfolded $t$, differences between observed
1130 and null data were obtained separately for negative and positive values. This illustrates that
1131 there is enrichment of SNPs in both tails of the null.

1132

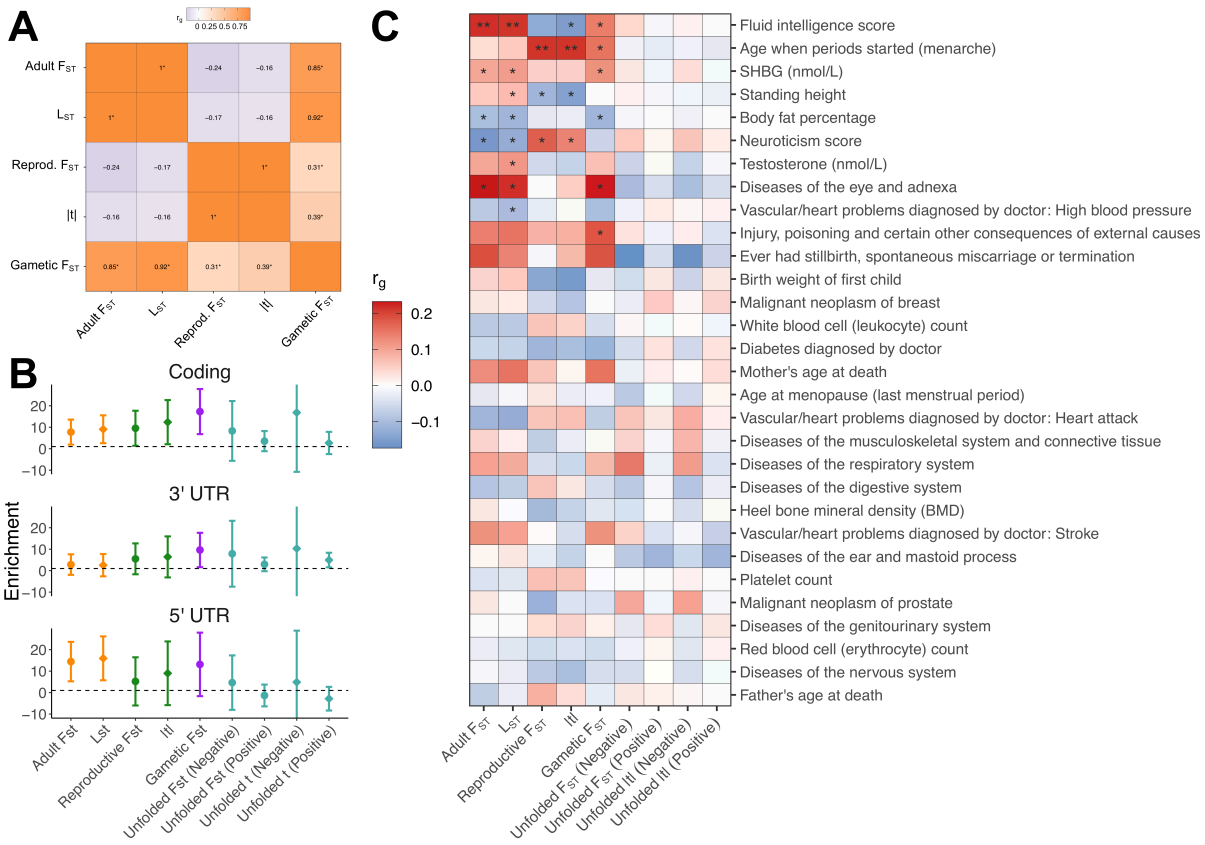



**Figure 5. Indications that sex-differentiated loci are more likely to be functional and contribute to trait variation. A.** Genetic correlations between metrics of sex-differential selection. Positive correlations (orange) imply that alleles have similar sex-specific effects on given fitness components, while negative correlations (purple) imply that alleles have opposing sex-specific effects on given fitness components. * denotes unadjusted p<0.05. **B.** Enrichments (±SE) of sex-differentiated loci in major functional categories. For each metric, enrichments were calculated as the relative SNP-heritability (as a fraction of total SNP-heritability) explained by a given functional category, divided by the relative number of SNPs (as a fraction of all SNPs) present in a given functional category. Dashed line = 1 (no enrichment). "Negative" and "Positive" refer to negative and positive values (*i.e.*, SA and SC components, respectively) of unfolded reproductive $\hat{F}_{ST}$ and unfolded t metrics. **C.** Genetic correlations between metrics of sex-differential selection and various UK Biobank phenotypes (as analysed by the Neale laboratory). Metrics of sex-differential selection have been polarised, such that positive correlations (red) imply that a trait is more beneficial to females than males (for the

1148 relevant fitness component), while negative correlations (blue) denote traits which are more
1149 beneficial to males than females. ** denotes FDR-adjusted $p<0.05$ and * denotes unadjusted
1150 $p<0.05$.
1151

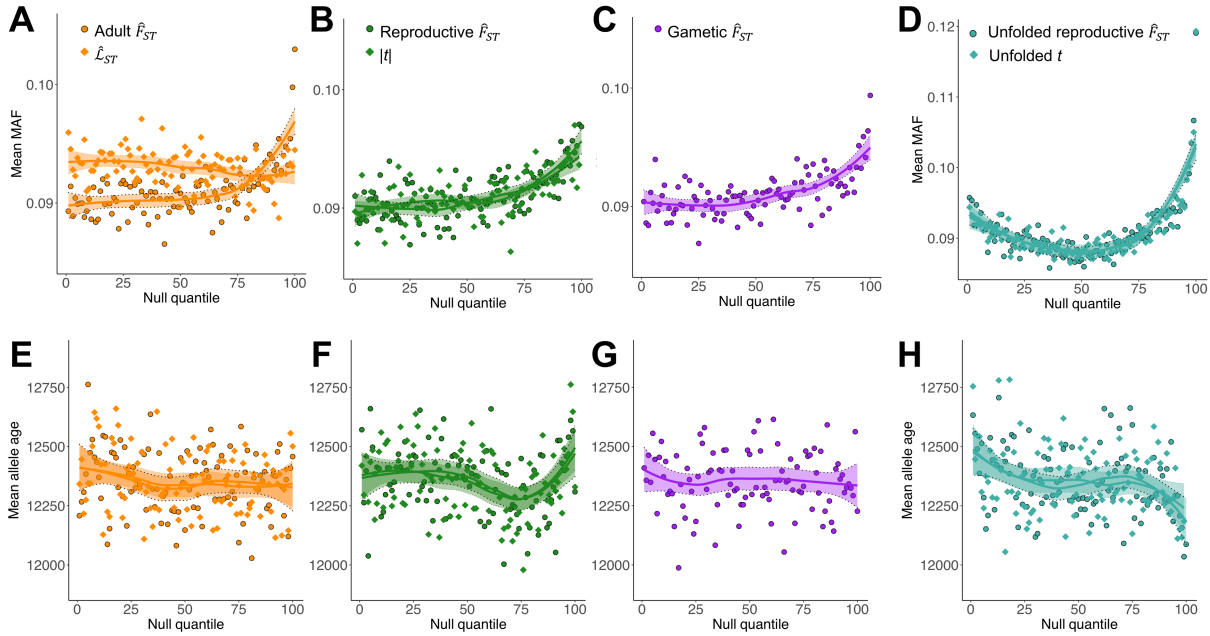



**Figure 6. Modes of evolution of sex-differentiated sites. A-D.** Mean MAF, in the UK Biobank, across 100 quantiles of the null for each metric of sex-differential selection. For $\hat{F}_{ST}$ metrics, *x*-axes correspond to Figs. 2A-C (and Fig. 3B for unfolded reproductive $\hat{F}_{ST}$). For mixed-model metrics, *x*-axes correspond to Fig. 4A-D. LOESS curves (±SE) are presented for visual emphasis. **E-H.** Mean age of the alternative (*i.e.*, non-reference) allele across 100 quantiles of the null for each metric of sex-differential selection. Each panel corrects for ascertainment bias of allele frequencies among highly sex-differentiated sites (*i.e.*, Fig. 6A-D). For visualisation purposes, this was done by averaging, in each quantile, allele age across 20 quantiles of alternative allele frequency in the UK Biobank (such that UK Biobank alternative allele frequency is approximately equal across quantiles). LOESS curves (±SE) are presented for visual emphasis.