An unbiased test reveals no enrichment of sexually antagonistic polymorphisms on the human X chromosome

Filip Ruzicka and Tim Connallon

School of Biological Sciences, Monash University, Clayton, Victoria, Australia

Mutations with beneficial effects in one sex can have deleterious effects in the other. Such ‘sexually antagonistic’ (SA) variants contribute to variation in life-history traits and overall fitness, yet their genomic distribution is poorly resolved. Theory predicts that SA variants could be enriched on the X chromosome or autosomes, yet current empirical tests face two formidable challenges: (i) identifying SA selection in genomic data is difficult; and (ii) metrics of SA variation show persistent biases towards the X, even when SA variants are randomly distributed across the genome. Here, we present an unbiased test of the theory that SA variants are enriched on the X. We first develop models for reproductive FST—a metric for quantifying sex-differential (including SA) effects of genetic variants on lifetime reproductive success—that control for X-linked biases. Comparing data from approximately 250,000 UK Biobank individuals to our models, we find FST elevations consistent with both X-linked and autosomal SA polymorphisms affecting reproductive success in humans. However, the extent of FST elevations does not differ from a model in which SA polymorphisms are randomly distributed across the genome. We argue that the polygenic nature of SA variation, along with sex asymmetries in SA effects, might render X-linked enrichment of SA polymorphisms unlikely.

1. Introduction

Adaptation requires genetic variation for fitness, yet we know surprisingly little about its genetic basis [1,2]. Is fitness variation mostly attributable to rare variants maintained by recurrent mutation, or to common variants maintained by various forms of balancing selection? To what extent are the frequencies of fitness-affecting variants influenced by genetic drift, departures from equilibrium or gene flow from neighbouring populations? To what extent do loci with large versus small fitness effects contribute to genome-wide fitness variation? Also, how are these genetic variants distributed across the genome? These are some of the most pressing questions in evolutionary biology, yet they are also the most challenging to answer [3].

Sexually antagonistic (SA) polymorphisms, wherein the alleles of a locus have opposing fitness effects on each sex, are an important class of fitness-affecting genetic variant. The basic conditions giving rise to SA fitness effects—including sex differences in selection on traits expressed by both sexes, and sex-differential phenotypic effects of mutations—are permissive [4] and have been widely documented in natural and experimental populations [5–8]. Accordingly, direct evidence for SA genetic variation has been reported in numerous quantitative genetic studies (e.g. [9–11]; see also [12]). Yet despite this, our understanding of the genomic distribution, fitness effects and evolutionary dynamics of SA variants remains poor [13–16].

An intriguing and oft-debated question is whether SA polymorphisms are more likely to be found on autosomes or the X chromosome. This question has

© 2022 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.
inspired an extensive body of theoretical research (e.g. [17–23]) that has identified conditions where the X chromosome should be enriched, or deficient, for SA polymorphisms. Rice [18], for example, highlighted conditions under which the X is a more permissive genomic location for balancing selection of SA alleles, while Fry [20] later emphasized the sensitivity of Rice’s predictions to the dominance relations between SA alleles, about which little is known [24]. Moreover, when SA polymorphisms are maintained by recurrent mutation rather than balancing selection, the X chromosome should typically harbour fewer SA polymorphisms, owing to stronger purifying selection and elevated rates of genetic drift at X-linked compared to autosomal loci [25–28]. In short, empirical studies—rather than theory alone—are needed to resolve whether the X, autosomes, or neither chromosome type, are enriched for SA polymorphisms.

So far, two formidable challenges have hampered empirical research on this question. The first is the logistical difficulty of identifying SA polymorphisms using current genomic approaches (e.g. [29–33]), which typically lack the statistical power needed to confidently identify SA loci [16,29,34,35]. The second is that X-linkage inflates empirical signals of SA polymorphism, which can lead to biases in detection and erroneous inferences of elevated SA polymorphism on the X [23]. These biases arise, on the one hand, from less constrained conditions for allele frequency divergence between the sexes at X-linked compared to autosomal loci [23,29,30], and, on the other hand, from the inherently stronger effects of X-linked compared to autosomal polymorphisms on male fitness variances and cross-sex fitness covariances [23,28,36–38]. These twin challenges of estimation and inference render previous tests for an enrichment of SA polymorphisms on the X chromosome ambiguous [30,39–47].

Here, we develop and implement an unbiased test for the enrichment of SA polymorphisms on the X chromosome. Building on recently developed methods for characterizing polygenic signals of SA polymorphism in population genomic data [48], we first test for polygenic signals of SA polymorphism on the X chromosome and autosomes, using genotypic and lifetime reproductive success (LRS) data from approximately 250 000 males and females from the UK Biobank [49]. Our metric of SA polymorphism, ‘reproductive \( F_{ST} \)’, quantifies sex differences in the genetic basis of LRS among UK Biobank adults, and thus controls for X-autosome biases in between-sex allele frequency differences that may arise between fertilization and reproductive maturity [23,29,30]. We then develop an idealized model of reproductive \( F_{ST} \) in which the prevalence and attributes of SA polymorphisms do not systematically differ between the X and autosomes. By comparing polygenic signals of SA polymorphism in the UK Biobank to this theoretical baseline, we control for both the elevated sampling variance and stronger effects of X-linked polymorphisms on the reproductive \( F_{ST} \) metric [23,28,36–38], allowing us to assess whether SA polymorphisms are disproportionately autosomal or X-linked.

2. Theoretical background

Our empirical analyses address two questions. First, is there a polygenic signal of SA polymorphism on the human X chromosome in the UK Biobank dataset (an autosomal signal was previously reported by Ruzicka et al. [48])? Second, do the magnitudes of these signals imply an elevation of SA polymorphisms on the X relative to autosomes? To formally test each question, we developed mathematical models for ‘reproductive \( F_{ST} \)’—a metric that potentially captures sex-differential (including SA) effects of alleles on LRS (see below and [48])—and compared empirical data to each model. For the first question, we developed null models for the distribution of reproductive \( F_{ST} \) in the absence of sex differences in selection. For the second question, we developed models of reproductive \( F_{ST} \) in which SA polymorphisms are present and inflate reproductive \( F_{ST} \) on both the X and autosomes, yet neither chromosome is enriched for them (i.e. there are no systematic differences between chromosomes in the abundance or attributes of SA polymorphisms). This second model accounts for other factors that are likely to differ between the X and autosomes, including the effects of diploidy versus haploidy on the expression of fitness variation. We elaborate on each model of reproductive \( F_{ST} \) below.

(a) Definition of reproductive \( F_{ST} \)

Our models and empirical analyses focus on bi-allelic loci (i.e. the vast majority of polymorphic sites in population genomic datasets). For an X-linked locus with alleles \( A_1 \) and \( A_2 \), females have three genotypes (\( A_1A_1, A_1A_2 \) and \( A_2A_2 \)) and males have two (\( A_1 \) and \( A_2 \)). Let \( n_{11, f} \), \( n_{12, f} \) and \( n_{22, f} \) represent the number of females of each genotype recorded in the UK Biobank dataset, respectively (\( N_f = n_{11, f} + n_{12, f} + n_{22, f} \)). The frequency of the \( A_1 \) allele in the female sample is \( p_f = (n_{11, f} + 1/2)n_{12, f}N_f^{-1} \). Likewise, \( n_{11, m} \) and \( n_{22, m} \) represent the number of males from the UK Biobank that carry alleles \( A_1 \) and \( A_2 \) (\( N_m = n_{11, m} + n_{22, m} \)), and \( p_m = n_{11, m}N_m^{-1} \) is the \( A_1 \) allele frequency in the male sample.

Letting \( f_i \) represent the total number of offspring produced by females carrying genotype \( ij \) (\( ij = \{11, 12, 22\} \)), and \( M_i \) represent the total number of offspring produced by males carrying genotype \( i \) (\( i = \{1, 2\} \)), the expected frequencies with which females and males transmit the \( A_1 \) allele to their offspring (respectively) are

\[
\begin{align*}
\tilde{p}_f' &= \frac{n_{11, f} + 1/2n_{12, f}}{N_f} \\
\tilde{p}_m' &= M_f \frac{n_{11, m} + (1/2)n_{12, m}}{M_i}
\end{align*}
\] (2.1)

Reproductive \( F_{ST} \) is a standardized measure of sex differences in allele transmission, owing to sex-specific genetic variation for LRS, and controlling for allele frequency differences between adults [48]. It is defined as

\[
F_{ST} = \frac{(\tilde{p}_f' - \tilde{p}_m')^2}{4p(1-p)},
\] (2.2)

where \( \tilde{p} = (p_f + p_m)/2 \). The same expression applies to autosomal loci, once male allele frequencies are adjusted to include diploidy in males, i.e.: \( \tilde{p}_m = (n_{11, m} + 1/2n_{12, m})N_m^{-1} \) and \( \tilde{p}_m' = (M_f + 1/2M_i)(M_{11, m} + M_{12, m} + M_{22, m})^{-1} \).

(b) Reproductive \( F_{ST} \) in the absence of sex differences in selection

With large sample sizes (as in the UK Biobank), no sex differences in selection and excluding polymorphic loci in which
the minor allele frequency (MAF) is very low (i.e. less than 1% in the UK Biobank; see below), reproductive \( F_{ST} \) estimates are well-approximated by a \( \chi^2 \) distribution, with estimates for an X-linked and autosomal locus, respectively, as

\[
\tilde{F}_{ST} \approx \frac{\hat{p}_1(1-\hat{p}_1)(1-\hat{F}_{ST})\sigma^2_f/(2N_f\mu_f^2)}{4\hat{p}_1(1-p)} \quad \text{and} \quad \tilde{F}_{ST} \approx \frac{\hat{p}_m(1-\hat{p}_m)(1-\hat{F}_{ST})\sigma^2_m/(2N_m\mu_m^2)}{4\hat{p}_m(1-p)}
\]

where \( X \) is a \( \chi^2 \) random variable with one degree of freedom, \( \mu_f, \mu_m, \sigma^2_f \) and \( \sigma^2_m \) correspond to the means (\( \mu_f, \mu_m \)) and variances (\( \sigma^2_f, \sigma^2_m \)) for female and male LRS (respectively) in the UK Biobank, \( F_{ST} \) is a measure of the deviation of female genotype frequencies from Hardy–Weinberg expectations (we define \( F_{IS} \) as in [16,34], where \( F_{IS} > 0 \) when heterozygotes are in excess of Hardy–Weinberg predictions, and \( F_{IS} < 0 \) when heterozygotes are deficient; see the electronic supplementary material); \( F_{ST} \) is the deviation in the male sample. With genotype frequencies near Hardy–Weinberg expectations and \( \hat{p}_1 \approx \hat{p}_m \), these expressions simplify—for X-linked and autosomal loci, respectively—to

\[
\tilde{F}_{ST} \approx \left( \frac{1}{8N_f\mu_f^2} + \frac{1}{4N_m\mu_m^2} \right) X \quad \text{and} \quad \tilde{F}_{ST} \approx \left( \frac{1}{8N_f\mu_f^2} + \frac{1}{8N_m\mu_m^2} \right) X.
\]

(c) Reproductive \( F_{ST} \) assuming a random genomic distribution of sexually antagonistic polymorphisms

To test whether SA polymorphisms are enriched on the X chromosome, we must first define a model for reproductive \( F_{ST} \) in which SA polymorphisms are present, randomly distributed across the genome, and where the frequencies and fitness effects of SA polymorphisms do not systematically differ between the X and autosomes. We expect that there will be many specific evolutionary genetic scenarios that can lead to equivalent patterns of SA polymorphism between autosomes and the X. For example, when SA polymorphisms are maintained at equilibrium under balancing selection, there are specific conditions of dominance between SA alleles that generate identical equilibrium frequencies of balanced SA alleles between the X and autosomes [23]. More complicated contexts of selection, including interactions between recurrent mutation, genetic drift and the distributions of sex-specific selection and dominance coefficients, may also sometimes result in similar patterns of SA polymorphism on the X and autosomes (although we expect that levels of polymorphism maintained by recurrent mutation will tend to be higher on autosomes, given the enhanced purifying selection and genetic drift that have been predicted [25] and documented (e.g. [28,50] for X-linked genes). In our baseline model of equivalence between the X and autosomes, we, therefore, focus on the case where SA polymorphisms are maintained at equilibrium under balancing selection, which allows us to set up a testable prediction regarding the signal of SA polymorphism on the X versus the autosomes. Whether this idealized scenario is a good model for SA polymorphism in the genomes of humans or other species is a point we return to in the Discussion.

For simplicity, we consider the case where adult genotype frequencies are approximately equal between the sexes and close to Hardy–Weinberg expectations, as is largely the case within the UK Biobank dataset. For a set of \( n_X \) and \( n_A \) polymorphic SA loci on the X chromosome and autosomes, respectively, the expected inflation of mean \( F_{ST} \) (relative to null expectations defined in equations (2.4a,b)) is

\[
f_{ST}^{\text{null}} = \frac{1}{n_X} \sum_{i=1}^{n_X} \left( \frac{1}{2N_{X_{ij}} d_{X_{ij}}} - \frac{1}{2N_{X_{ij}} d_{X_{ij}}} \right)^2
\]

and

\[
f_{ST}^{\text{null}} = \frac{1}{n_A} \sum_{i=1}^{n_A} \left( \frac{1}{2N_{A_{ij}} d_{A_{ij}}} - \frac{1}{2N_{A_{ij}} d_{A_{ij}}} \right)^2,
\]

where \( F_{STX} \) and \( F_{STA} \) denote mean \( F_{ST} \) among these loci (see the electronic supplementary material). The terms in square brackets capture effects of sex-differential selection on \( F_{ST} \). Note that when there are no sex differences in selection (i.e. the terms in square brackets evaluate to zero), there will be no inflation (i.e. \( f_{ST}^{\text{null}} = 0 \) and \( f_{ST}^{\text{null}} = 0 \)), and \( F_{STX} \) and \( F_{STA} \) conform to the null models in equation (2.4).

When fitness effects of SA alleles are small [19,51], single-generation allele frequency changes at X-linked and autosomal loci (respectively) are well-approximated by

\[
\Delta p_{X,i} = \frac{p_i(1-p_i)}{3} \left[ \frac{1}{d_{X_{ij}}} + \frac{1}{d_{X_{ij}}} \right]
\]

and

\[
\Delta p_{A,j} = \frac{p_j(1-p_j)}{4} \left[ \frac{1}{d_{A_{ij}}} + \frac{1}{d_{A_{ij}}} \right],
\]

(see the electronic supplementary material). SA selection is expected to maintain genetic variation (i.e. conditions for balancing selection are met) when the strength of selection for a given locus is relatively symmetric between the sexes, in which case the terms in the square brackets of equations (2.6a) and (2.6b) must sum to zero (i.e. \( w_{X_{ij}} d_{X_{ij}}/d_{X_{ij}} \approx -w_{X_{ij}} d_{X_{ij}}/d_{X_{ij}} \)). Note that this condition applies whether or not there are dominance interactions, including dominance reversals [24,52–54], between SA alleles at a given locus. Substituting this identity into equations (2.5a) and (2.5b), we obtain:

\[
f_{ST}^{X} = \frac{1}{n_X} \sum_{i=1}^{n_X} \left( \frac{1}{d_{X_{ij}}} \right)^2
\]

and

\[
f_{ST}^{A} = \frac{1}{n_A} \sum_{i=1}^{n_A} \left( \frac{1}{d_{A_{ij}}} \right)^2.
\]
frequencies and fitness effects of their alleles, then the sum-
mation terms in $f_{X}^{sel}$ and $f_{A}^{sel}$ should have the same
magnitudes, and the X-to-autosome ratio for the expected
amount of $F_{ST}$ inflation will be

$$\frac{f_{X}^{sel}}{f_{A}^{sel}} = \frac{9}{4}. \quad (2.8)$$

In other words, we expect a 2.25-fold higher inflation of
$F_{ST}$ for the X compared to the autosomes in the absence of chro-
mosomal differences for SA polymorphisms. This additional
inflation for X-linked loci arises because haplidity in males
inflates the contributions of X-linked loci to the variance for
male fitness components (relative to autosomal loci) [23].

In practice, loci contributing to an inflation of $F_{ST}$ may
either be targets of selection or in linkage disequilibrium
(LD) with such targets. While the prediction in equation (2.8)
applies to targets of SA selection, it will remain applica-
table to linked loci provided their degree of LD with target
loci does not systematically differ between the X chromosome
and autosomes (see the electronic supplementary material).
In our comparisons of the two chromosome types (see
below), we used LD pruning to minimize the differential
effects of hitchhiking between chromosomes. Because LD
tends to be higher on the human X chromosome compared
to autosomes [55], any uncontrolled chromosomal bias in
the effects of LD on $F_{ST}$ inflation should tend to increase
the likelihood of identifying a signal of enriched SA poly-
morphism on the X, which makes our eventual conclusion
(that there is no such signal) a conservative one.

Below, we empirically estimate the amount of inflation in
$F_{ST}$ for the X and the autosomes in the UK Biobank (i.e.
inflation relative to the average amount of noise in $F_{ST}$ esti-
mates, per chromosome), and compare these estimates to the
baseline prediction of 9/4.

### 3. Methods

(a) Quality-filtering of UK Biobank data

Access to UK Biobank data was granted under project number
52049. We employed the same quality-filtering settings for
X-linked loci in this study as employed for autosomal loci by
Ruzicka et al. [48]. Briefly, we excluded individuals with high
relatedness (3rd degree or closer), non-White British ancestry,
high heterozygosity and missing rates, individuals whose
reported sex and genetic sex differed, individuals whose age
was less than 45 years, and aneuploids. Across retained individ-
uals, we excluded non-diallelic sites, sites with MAF < 0.01,
males, and field 2743 ‘Number of live births’ for females (see [48]
for further details on quality-filtering). These data were used to
estimate allele frequency transmission ($F_{pm}$ and $F_{p}$) equation (2.1)
and reproductive $F_{ST}$ for each polymorphic locus (equation (2.2)).

We generated a null distribution for reproductive $F_{ST}$ by simu-
lating loci from the theoretical null model (i.e. equation (2.3a)),
with the sample sizes per locus ($N_{f}$ and $N_{pm}$ per locus) matching
those in the UK Biobank data, and estimates of the sex-specific
mean and variance for LRS based on all individuals included in
the dataset. We also generated an empirical null distribution for
$F_{ST}$ through a single permutation of LRS values (without permut-
ing sex) and re-calculating $F_{ST}$ for each locus using permuted data.
This permutation procedure ensured that neither sex differences in
the allele frequencies among adults, nor sex differences in the
relationship between genotype and LRS, contribute to $F_{ST}$ in the
permuted data, leaving only estimation error (including error aris-
ing from locus-specific heterogeneity in LRS values owing to
variation in missing rates across single nucleotide polymorphisms
(SNPs)) to contribute to $F_{ST}$ in permuted data. We chose to perform
a single permutation of LRS values for computational efficiency
and because our focus was on testing significance across the set
of loci, rather than establishing statistical significance for
individual loci. We tested whether the distribution of $F_{ST}$ in
observed data differed from both null distributions using
Wilcoxon sum-rank and Kolmogorov–Smirnov tests.

To consolidate our inference that $F_{ST}$ inflations reflect genuine
phenotypic effects (as opposed to technical artefacts), we assessed
whether sites with elevated $F_{ST}$ were more likely to be functional.
We first obtained variant effect predictions for each SNP using
annotations from the hg19 reference genome in SNPEff [57], with
sites categorized as ‘intergenic’ or ‘genic’, the latter defined broadly
to include coding and potential regulatory genomic regions. We
then performed logistic regressions, in which genic/intergenic
was the binary response variable and $F_{ST}$ the independent variable,
to assess relationships between $F_{ST}$ and genic/intergenic status. We
also assessed whether associations with genic status were greater in
observed than null (both simulated or permuted) data by re-calcul-
ating the regression coefficient of $F_{ST}$ on genic status among 1000
bootstrap replicates of the data, where one replicate consists of
the set of SNPs sampled with replacement.

(b) Quantifying polygenic signals of sexually
antagonistic polymorphism

For each X-linked locus, we estimated allele frequencies in adults
($F_{pm}$ and $F_{p}$). To quantify LRS, we used reported offspring numbers
from UK Biobank field 2405 ‘Number of children fathered’ for

### 4. Results

(a) Signals of sexually antagonistic polymorphism for
X-linked loci

The total sample size, after quality-filtering, includes 249,021
individuals ($N_{pm} = 115,531$ and $N_{f} = 133,490$), $N = 229,196
imputed polymorphic sites on the X chromosome and $N = 7,851,642$ imputed polymorphic sites on autosomes. We first compared reproductive $F_{ST}$ on the X chromosome to a distribution of $F_{ST}$ simulated from the theoretical null distribution (based on equation (2.3a)) and to an empirical null distribution obtained by permuting LRS values within each sex (see Methods). As shown for autosomal loci previously (figure 1; [48]) and as predicted under SA selection, mean $F_{ST}$ for X-linked loci was elevated relative to both theoretical and empirical null distributions (theoretical $F_{ST}^X = 1.369 \times 10^{-6}$; permuted; permuted $F_{ST} = 1.365 \times 10^{-6}$ observed $F_{ST}^X = 1.482 \times 10^{-6}$; Wilcoxon rank-sum and Kolmogorov–Smirnov tests, all $p < 0.001$; figure 1; electronic supplementary material, figure S7). We observed a 14.8% enrichment of observed sites in the top 1% quantile of the theoretical null distribution (expected number of sites = 2292; observed number of sites = 2632; $\chi^2 = 23.592$, $p < 0.001$; figure 1; electronic supplementary material, figure S7) and a 8.4% enrichment of observed sites in the top 1% quantile of the empirical null distribution (expected number of sites = 2292; observed number of sites = 2484; $\chi^2 = 7.719$, $p = 0.005$). There were no individual large-effect loci contributing to this signal: the minimum $\chi^2$ $p$-value across all sites (1.213 $\times 10^{-9}$) was well above the Bonferroni-corrected threshold (2.182 $\times 10^{-7}$) and the minimum false discovery rate $q$-value was 0.807. Thus, genomic signals of SA polymorphism for X-linked loci are polygenic.

If genic sites (broadly defined to include coding and regulatory regions) are more likely to have phenotypic effects than intergenic sites, and if $F_{ST}$ inflations reflect genuine phenotypic effects of SA polymorphisms, we should observe that sites with high $F_{ST}$ are more likely to be genic. As found for autosomes previously [48], we detected a positive association between $F_{ST}$ and genic status among X-linked polymorphisms (SNP coded as genic based on genome annotations; see Methods; binomial generalized linear model (GLM), log odds ratio (logOR) $\pm 95$ confidence interval (CI) = 16.399.81[12.189.41–20.601.39], $p < 0.001$; electronic supplementary material, figure S8). By contrast, we found no significant association between permuted $F_{ST}$ and genic status (binomial GLM, logOR $\pm 95$ CI = 3822.991–679.01–8312.16), $p = 0.096$; electronic supplementary material, figure S8), or between simulated $F_{ST}$ and genic status (binomial GLM, logOR $\pm 95$ CI = 2107.521–2408.14–6610.39, $p = 0.360$), and the association between $F_{ST}$ and genic status was stronger in the observed than permuted data (1000 bootstrap replicates of the difference between the log odds-ratio in observed and permuted data = 12.680.27[6417–18.382.84], empirical $p < 0.001$). While these associations represent suggestive evidence $F_{ST}$ inflations are attributable to genuine phenotypic effects (and align with previous enrichment of candidate SA sites in functional genomic regions; e.g. [47]), we emphasize that the evidence is not definitive, as non-genic sites may be functional as well.

(b) No evidence that sexually antagonistic polymorphisms are enriched on either chromosome type

Evidence for elevated reproductive $F_{ST}$ among X-linked loci (relative to the ‘no sex-differential selection’ null described by equation (2.3a)), together with previous evidence for elevated reproductive $F_{ST}$ on autosomes [48] (relative to the ‘no sex-differential selection’ null described by equation (2.3b)), suggest that SA polymorphisms segregate on both chromosome types. These findings provided motivation to compare signals of SA polymorphism on autosomes relative to the X.

We first compared mean $F_{ST}$ between the X chromosome and autosomes (using a set of LD-pruned sites; $N_X = 29,859$; $N_{Auto} = 1,056,003$), which should be larger for X-linked than autosomal loci (even in the absence of any SA polymorphism) because of the smaller sample sizes, and thus higher sampling variances, for X-linked loci. Accordingly, the ratio of mean X-linked to autosomal $F_{ST}$ ($\pm 95$ CI) was greater

---

**Figure 1.** Polygenic signals of SA polymorphism on the X chromosome and autosomes relative to ‘no sex-differential selection’ nulls. (a) Proportion of X-linked sites (grey, permuted; pink, observed) and autosomal sites (grey, permuted; green, observed) in each of 100 quantiles of a simulated null distribution for $F_{ST}$, with the null distributions described by equations (2.3a) and (2.3b), respectively. In the absence of sex-differential selection, approximately 1% of sites should fall into each quantile of the simulated null distribution. (b) Observed/permuted ratio of the proportion of sites in each of 100 quantiles of the simulated null distribution, for X-linked (pink) and autosomal (green) sites, respectively. Both the above panels used the full set of sites with high $F_{ST}$, with the null distributions described by equations (2.3a) and (2.3b), respectively. In the absence of sex-differential selection, approximately 1% of sites should fall into each quantile of the simulated null distribution. (i) signals of SA polymorphism relative to null distributions cannot be artificially inflated by LD between sites (when considering X-linked and autosomal sites separately; see [48]); and (ii) power to detect X versus autosome differences is reduced by LD pruning. We present equivalent figures for LD-pruned data as electronic supplementary material, figure S7. (Online version in colour.)
than one (figure 2a), whether in simulated \( F_{ST} \) values based on theoretical null distributions (\( F_{ST}^{auto} = 1.386 \times 10^{-6} \), \( F_{ST}^{X} = 8.661 \times 10^{-7} \), \( F_{ST}^{auto} / F_{ST}^{X} = 1.599 [1.575 - 1.625] \)), permuted data (\( F_{ST}^{auto} = 1.431 \times 10^{-6} \), \( F_{ST}^{X} = 8.853 \times 10^{-7} \), \( F_{ST}^{auto} / F_{ST}^{X} = 1.616 [1.589 - 1.643] \)), or observed data (\( F_{ST}^{auto} = 1.448 \times 10^{-6} \), \( F_{ST}^{X} = 9.016 \times 10^{-7} \), \( F_{ST}^{auto} / F_{ST}^{X} = 1.607 [1.581 - 1.634] \)).

To assess whether SA polymorphisms are enriched on autosomes or the X chromosome, we estimated, for each chromosome type, the degree of \( F_{ST} \) inflation relative to the null (i.e. \( f_{Sel}^{X} \) and \( f_{Sel}^{auto} \); see Theoretical Background) as the mean observed \( F_{ST} \) minus the fraction of \( F_{ST} \) that is attributable to sampling variance (i.e. mean simulated or permuted \( F_{ST} \)). We then compared the ratio of X to autosomal \( F_{ST} \) inflation to a baseline model in which the chromosome types exhibit no differences in SA polymorphism (i.e. \( f_{Sel}^{X} / f_{Sel}^{auto} = 9/4 = 2.25 \); equation (2.8)). We estimated \( f_{Sel}^{X} / f_{Sel}^{auto} \) as 1.770[0.828–3.290] based on observed relative to simulated \( F_{ST} \) values, and \( f_{Sel}^{X} / f_{Sel}^{auto} \) as 1.086[0.877–3.291] based on observed relative to permuted \( F_{ST} \) values. Whereas enrichment of SA polymorphisms on the X would predict a ratio greater than 2.25, both point estimates fell below 2.25, though neither difference was statistically significant (empirical \( p = 0.853 \) and \( p = 0.864 \) for simulated and permuted data, respectively). Overall, there is no compelling evidence for an enrichment of SA polymorphisms on either chromosome type.

### 5. Discussion

Research on the genetic basis of SA fitness variation is motivated, in part, by its relevance to broader questions about adaptive evolution, such as the distribution of dominance coefficients among adaptive mutations [1,24,58], the extent of balancing selection in genomes [3] and the evolution of sex chromosome systems [59–62]. For example, evidence for enrichment of SA variants on the X chromosome might imply that SA polymorphisms evolve under balancing selection, with male-beneficial variants typically recessive and female-beneficial variants typically dominant [18,19].

On the other hand, evidence for enrichment of SA variants on autosomes might imply that SA variants exhibit co-dominant fitness effects [17], experience beneficial reversals of dominance [20] and/or have evolutionary dynamics dominated by directional selection or drift [26,27]. Yet empirically assessing whether SA polymorphisms are typically X-linked or autosomal is extremely challenging [23,34,35].

Here, we overcome the logistical hurdle of detecting SA polymorphisms by using a study organism—humans—in which there is previous evidence for SA selection [10,63,64] and a study population—the UK Biobank—in which sample sizes are large enough to detect polygenic signals of SA polymorphism [48]. We overcome the bias towards elevated X-linked between-sex allele frequency differentiation among adults [23,29,30] by using a metric of SA polymorphism—reproductive \( F_{ST} \)—that controls for allele frequency differences among adults [48]. And we overcome the biases arising from larger effects of X-linked alleles on metrics of SA variation, owing to elevated X-linked sampling variances and X-linked haploidy in males [23,36], by developing models for reproductive \( F_{ST} \) estimates that account for these effects.

Our main finding was that X-linked sites displayed \( F_{ST} \) inflations consistent with polygenic SA selection (as shown for autosomal sites previously [48]) but the extent of \( F_{ST} \) inflation did not significantly differ from the 9/4 X-autosome ratio predicted by an idealized model in which SA polymorphisms are evenly distributed across the genome. These results accord well with re-analyses of previous empirical
research in a range of species [23], which reveal little compelling evidence for enrichment of SA polymorphisms on the X chromosome once biases towards enhanced X-linked effects are accounted for. Indeed, enrichment of SA polymorphisms on the X chromosome might be considered unlikely for additional reasons. In particular, SA variants are likely to exhibit small fitness effects owing to the polygenic nature of fitness variation [28,65,66], as also evidenced by the absence of large-effect SA loci in this dataset. Small-effect mutations, including SA mutations, are highly susceptible to genetic drift and potentially more likely to be co-dominant [66,67]—conditions that both favour autosomal enrichment of SA polymorphisms [17,26,27]. Furthermore, fitness effects of SA variants may be asymmetric between sexes owing to well-documented sex differences in strategies employed to achieve reproductive success [8,68–70]. This could render SA variants susceptible to directional (rather than balancing) selection, which also tends to favour autosomal enrichment of SA polymorphisms [26,27].

(a) Limitations of our analysis approach and future directions

Though our analyses correct for several pre-existing biases, they present some limitations. In terms of the data, the relatively small number of independently segregating polymorphisms across the human X chromosome (N ~ 30,000 LD-pruned imputed sites) reduces power to detect differences between autosomal and X-linked sites. While there is prior support for SA selection on phenotypes [64] and for polygenic signals of autosomal SA polymorphism [48] in this population, the extent to which \( F_{ST} \) inflations are driven by SA polymorphisms, as opposed to some loci showing sex differences in directional selection and neutral sites linked to selected polymorphisms, is unclear. Empirical work in systems where fitness can be measured relatively easily, sufficiently large samples of genomic sequences can be obtained, and experimental work can be carried out to test the fitness effects of candidate SA polymorphisms (e.g. the dioecious plant Silene latifolia; [9]), represent promising avenues for further research. Although there is little prospect of obtaining Biobank-scale samples of genome sequences in non-human organisms, we suspect that genetic variation for fitness may often be greater than in humans, increasing the likelihood that polygenic signals of SA selection will be detectable [16,48].

Regarding limitations of theory, previous research (including our own) has predominantly used single-locus models to make predictions about the genomic distribution of SA polymorphisms [17–23,27,71], despite the (likely) polygenic nature of fitness variation. Though we can extend single-locus predictions to polygenic scenarios when we assume that loci segregate independently and fitness effects are multiplicative, this simplifying assumption may be problematic when the extent of linked selection differs systematically between autosomal and X-linked loci [25]. Moreover, the 9/4 ratio of \( F_{ST} \) inflation is a prediction for signals of SA polymorphism assuming a random genomic distribution of balanced SA polymorphisms, yet balancing selection is only one of the multiple modes of evolution that can potentially affect SA polymorphisms [26,27,65]. Empirical data from the UK Biobank are consistent with the 9/4 prediction, yet our failure to reject this model should not be interpreted as evidence that the evolutionary scenario underlying this prediction is a reasonable description of genome-wide SA polymorphism in humans (or other species). Other evolutionary scenarios (e.g. interactions between recurrent mutation, genetic drift, and the distributions of sex-specific selection and dominance coefficients) might also lead to predictions close to 9/4. Finally, most current theoretical models of SA variation fall firmly within the classical population genetic tradition [72], in which the fitness effects of SA variants are arbitrarily assigned rather than explicitly modelled. Because fitness effects of genetic variation are properties of the distribution of mutations affecting traits under selection, models of adaptation that incorporate these features of biology (e.g. Fisher’s Geometric model [4,58] and other trait-based population genetic models [65]) may bring us closer to a robust theory for the fitness effects, dynamical properties and genomic distribution of SA polymorphisms.

Ethics. Permission to use UK Biobank data was granted under project number 52049.

Data accessibility. Code is available from GitHub (/filipluca/Polygenic_SA_selection_X_vs_A), along with metadata (within the restrictions of the UK Biobank agreement) from the Dryad Digital Repository (https://doi.org/10.5061/dryad.pvsk6k66) [73] and additional data in the electronic supplementary material [74].

Authors’ contributions. F.R.: conceptualization, data curation, formal analysis, investigation, methodology, project administration, visualization, writing—original draft, writing—review and editing; T.C.: conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by the UK Biobank (project number 52049) and by funds from the Australian Research Council.

Acknowledgements. We are grateful to two anonymous reviewers for comments that substantially improved the paper.

References

1. Orr HA. 2010 The population genetics of beneficial mutations. Phil. Trans. R. Soc. B 365, 1195–1201. (doi:10.1098/rstb.2009.0282)
2. Charlesworth B. 2015 Causes of natural variation in fitness: evidence from studies of Drosophila populations. Proc. Natl Acad. Sci. USA 112, 1662–1669. (doi:10.1073/pnas.1423275112)
3. Lewontin RC. 1974 The genetic basis of evolutionary change. New York, NY: Columbia University Press.
4. Connallon T, Clark AG. 2014 Evolutionary inevitability of sexual antagonism. Proc. R. Soc. B 281, 2013–2123. (doi:10.1098/rspb.2013.2123)
5. Mackay TFC. 2001 The genetic architecture of quantitative traits. Annu. Rev. Genet. 35, 303–339. (doi:10.1146/annurev.genet.35.102401.090633)
6. Cox RM, Caldebeck R. 2009 Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. Am. Nat. 173, 176–187. (doi:10.1086/595841)
7. De Lisle SP, Goedert D, Reedy AM, Svensson EI. 2018 Climatic factors and species range position predict sexually antagonistic selection across taxa.
