Amplification is the Primary Mode of Gene-by-Sex Interaction in Complex Human Traits

Carrie Zhu[1,2], Matthew J. Ming[1,2], Jared M. Cole[2], Mark Kirkpatrick[2], Arbel Harpak[1,2,+]

[1] Department of Population Health, The University of Texas at Austin

[2] Department of Integrative Biology, The University of Texas at Austin

[+] To whom correspondence should be addressed: arbelharpak@utexas.edu

Abstract

Sexual dimorphism is observed in many complex traits and diseases and is suspected to be in part due to widespread gene-by-sex interactions (GxSex). To date, empirical evidence for GxSex in GWAS data has been elusive. We hypothesized that GxSex may be pervasive but largely missed by current approaches if it acts primarily through sex differences in the magnitude of many genetic effects ("amplification"), regulated by a shared cue such as a sex hormone, rather than differences in the identity of causal variants or the direction of their effect. To test this hypothesis, we inferred the genetic covariance structure between males and females across 27 physiological traits in the UK Biobank. We found amplification to be a pervasive mode of GxSex across traits. As one example, we estimate that 38% of variants have a greater effect on urate levels in females than males. For some traits, notably those related to body mass, testosterone levels are associated with the magnitude of genetic effects in both males and females, but the association is opposite in sign between the sexes. Finally, we developed a novel test of sexually-antagonistic viability selection linking GxSex signals to allele frequency divergence between adult males and females. Using independent allele frequency data, we find marginally-significant evidence for contemporary sexually-antagonistic selection on genetic variation associated with testosterone. In summary, our results suggest that the systematic amplification of genetic effects is a common mode of GxSex that may contribute to sexual dimorphism and fuel its evolution.

Introduction

Genetic effects are not exerted in a vacuum, but in environments that mediate them. Several lines of evidence suggest large context dependency of genetic effects on polygenic (complex) human traits[1–5], or "gene-by-environment interaction" (GxE). At the same time, unlike studies in study systems where the environment can be experimentally manipulated[6–13], there are but a few cases in which GxE models consistently explain data—such as Genome-Wide Association Study (GWAS) data—better than parsimonious models that assume independent contributions of genetic and environmental factors[2,4,14–16]. This discrepancy may point to a pervasive mediation of polygenic effects by the environment—alongside the inability of our current toolbox to quantify polygenic GxE[17–19].

Perhaps nowhere is this discrepancy more apparent than in the case of gene-by-sex interaction (GxSex). Sexual dimorphism is observed in many complex traits and diseases and is suspected to be in part due to widespread GxSex involving autosomal loci[20,21]. GxSex may be a result of sex differences in environmental contexts or epistatic interaction with sex chromosomes. Sex differences in polygenic effects on traits are clearly of high evolutionary[22–26] and translational[27–31] importance. Yet with the exception of testosterone levels[32–35], the genetic basis of sexual dimorphism is not well-understood[31]. To date, empirical evidence for GxSex in GWAS data—whether focused on identifying large GxSex effects at individual loci or by estimating genetic correlations between the sexes for polygenic traits—has been lacking.

Considering sex as an "environmental" variable offers a methodological advantage as well. The study of context-dependency in humans is often complicated by study participation biases, leading to genetic ancestry structure that confounds genotype-phenotype associations[3,36–38], reverse causality between the phenotype and environment variable, collider bias, gene-by-environment correlation and other problems[39–41]. Focusing on sex as a case study circumvents many of these "usual suspects" problems: Problems involving the phenotype causally affecting sex are unlikely. Study participation in large biobanks like the UK Biobank (UKB) differs by sex[42]; and work by Piratsu et al. further argued that allele frequency differences between males and females may reflect sex-specific recruitment biases[43]. However, a recent study by Benonisdottir and Kong found no evidence of sex-specific genetic associations with UKB participation[38], and another by Kasimatis et al. showed that many apparent associations of autosomal genotypes and biological sex in the UKB are instead primarily due to a bioinformatic artifact—the mis-

hybridization of autosomal genotyping probes with sex chromosomes[44]. Finally, sex is a variable that is almost always measured; so insight into sex-specific genetic effects can be incorporated straightforwardly in future studies and in clinical risk prediction.

Here, we set out to characterize GxSex across complex human traits. We begin by suggesting a mode of GxSex that may have gone largely undetected by previous studies: A systematic difference in the magnitude of effect of many variants between the sexes, which we refer to as "amplification"[45]. Amplification can happen for a large set of variants regulating a specific pathway if the pathway responds to a shared cue[46–49]. In classic hypothesis testing approaches that test for a GxE effect separately in each variant, the signal of amplification may be crushed under the multiple hypothesis burden. On the other hand, even state-of-the-art tools designed with complex traits in mind may miss amplification signals: They often treat genetic correlation (between GWAS estimates based on samples from two environments) as a litmus test for whether effects are the same in two groups[50–54], but correlations are scaleless and thus may entirely miss amplification effects.

To quantify the role of genetic amplification in sex differences, we developed a new approach for flexibly characterizing a mixture of male-female genetic covariance relationships, and applied it to 27 traits in the UK Biobank. We found that amplification is pervasive across traits, and that it is in fact the primary mode of GxSex that helps explain sex differences in phenotypic variance. Finally, we consider an implication of pervasive amplification for sexually-antagonistic selection: Trait-associated variants may be subject to sexually-antagonistic selection when male and female trait optima are very different; but, surprisingly, even if the trait optima are very similar. We developed a novel test for sexually-antagonistic polygenic selection, which connects GxSex signals—including amplification—to signals of contemporary viability selection. Using this test, we find subtle evidence of sexually-antagonistic selection on variants associated with testosterone levels.

Results

The limited scope of single-locus analysis. We conducted GWASs stratified by sex chromosome karyotype for 27 continuous physiological traits in the UKB using a sample of ~150K individuals with two X chromosomes and another sample of ~150K individuals with XY, and a

combined sample that included both the XX and XY samples. We chose to analyze traits with SNP heritabilities over 7.5% in the combined sample, to have higher statistical power. While there is not a strict one-to-one relationship between sex chromosome karyotype and biological sex, we label XX individuals as females and XY individuals as males, and view these labels as capturing group differences in distributions of environmental context, rather than as a dichotomy[55]. Throughout, we analyze GWAS on the raw measurement units as provided by UKB. (See note on the rationale behind this choice in the section **Amplification of genetic effects is the primary mode of GxSex**).

Among the 27 traits, we observed substantial discordance between males and females in associations with the trait only for testosterone and waist:hip ratio (whether or not it is adjusted for BMI; **Fig. S1**). For testosterone, as noted in previous analyses, genes neighboring sex-specific associations are often in separate pathways in males and females[32,34]. This is reflected in the small overlap of top associations in our GWAS. For example, in females, the gene CYP3A7 is involved in the hydroxylation of testosterone, resulting in its inactivation. In males, FKBP4 plays a role in the downstream signaling of testosterone on the hypothalamus. Both genes, to our knowledge, do not affect testosterone levels in the other sex.

For waist:hip ratio, we saw multiple associations in females only, such as variants near ADAMTS9, a gene involved in insulin sensitivity[56]. As previous work established[32,34,35], testosterone and waist:hip ratio are the exception, not the rule: Most traits did not display many sex differences in top associations. For instance, arm fat-free mass, a highly heritable dimorphic trait, showed near-perfect concordance in significant loci (**Fig. S1**). A previous study[35] examining the concordance in top SNPs between males and females found few uniquely-associated SNPs (<20) across the 84 continuous traits they studied; waist:hip ratio was an exception with 100 associations unique to one sex. Considering the overwhelming evidence on the vast polygenicity of additive genetic variation affecting complex traits[57–59], it stands to reason that looking beyond lead associations, through a polygenic prism, may aid in the characterization of non-additive effects (such as GxSex) as well.

The limited scope of analyzing GxSex via heritability differences and genetic correlations. We therefore turned to consider the polygenic nature of GxSex, first by employing commonly-used approaches. Specifically, we used LD Score Regression (LDSC)[54,60] to estimate sex-specific

SNP heritabilities and the genetic correlation between the sexes for each trait. In most traits (17/27), males and females had a genetic correlation greater than 0.9. Testosterone had the lowest genetic correlation of 0.01, which suggests very little sharing of signals between males and females (see similar results by Flynn et al.[34] and Sinnott-Armstrong et al.[32]).

For the majority of traits (18/27), male and female heritabilities were both greater than the heritability in a sample that included both sexes. For instance, in arm fat-free mass (right), the heritability in the both-sex sample was 0.232 (± 0.009), while the heritabilities for male and female were 0.279 (± 0.012) and 0.255 (± 0.011), respectively. In particular, all body mass-related traits, excluding BMI-adjusted waist:hip ratio, had greater sex-specific heritabilities **(Table S6)**.

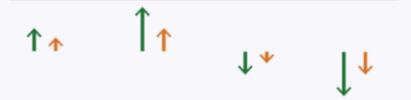
Model	Motivation	Illustration of Effect Covariance	Expectation from Heritability Analysis
No GxSex	Little previous evidence for GxSex		(a) h^2_m can only differ from h^2_f through environmental variance differences (b) $h^2_m < h^2$ or $h^2_f < h^2$
Weakly or negatively correlated genetic effects	Sexual dimorphism is pervasive and heritable contribution is expected to lie primarily in autosomes		(a) Low or negative genetic correlation (b) h^2_m , $h^2_f > h^2$, and the larger the difference, the lower the genetic correlation
Highly correlated effects, difference in magnitude ("amplification")	Response to cues such as testosterone; evidence for GxE in non-human organisms		(a) High genetic correlation (b) h^2_m or $h^2_f < h^2$
Mixture of covariance relationships	Heritability analysis often incompatible with either model or cannot distinguish between models		Compatible with all observations. motivates this work, including: (a) Direct estimation of genetic effect covariance. (b) Modelling mixture components

Table 1. Polygenic Models of GxSex. We examine different models of the nature of GxSex in complex traits that link to previous studies and motivations. Each model leads to different expectations from the analysis of heritability and genetic correlations (**Fig. 1**). The illustrations in the third column depict examples of directions and magnitudes of genetic effects, corresponding to each model. h^2_m, h^2_f and h^2 denote narrow-sense heritabilities in males, females, and a combined sample, respectively.

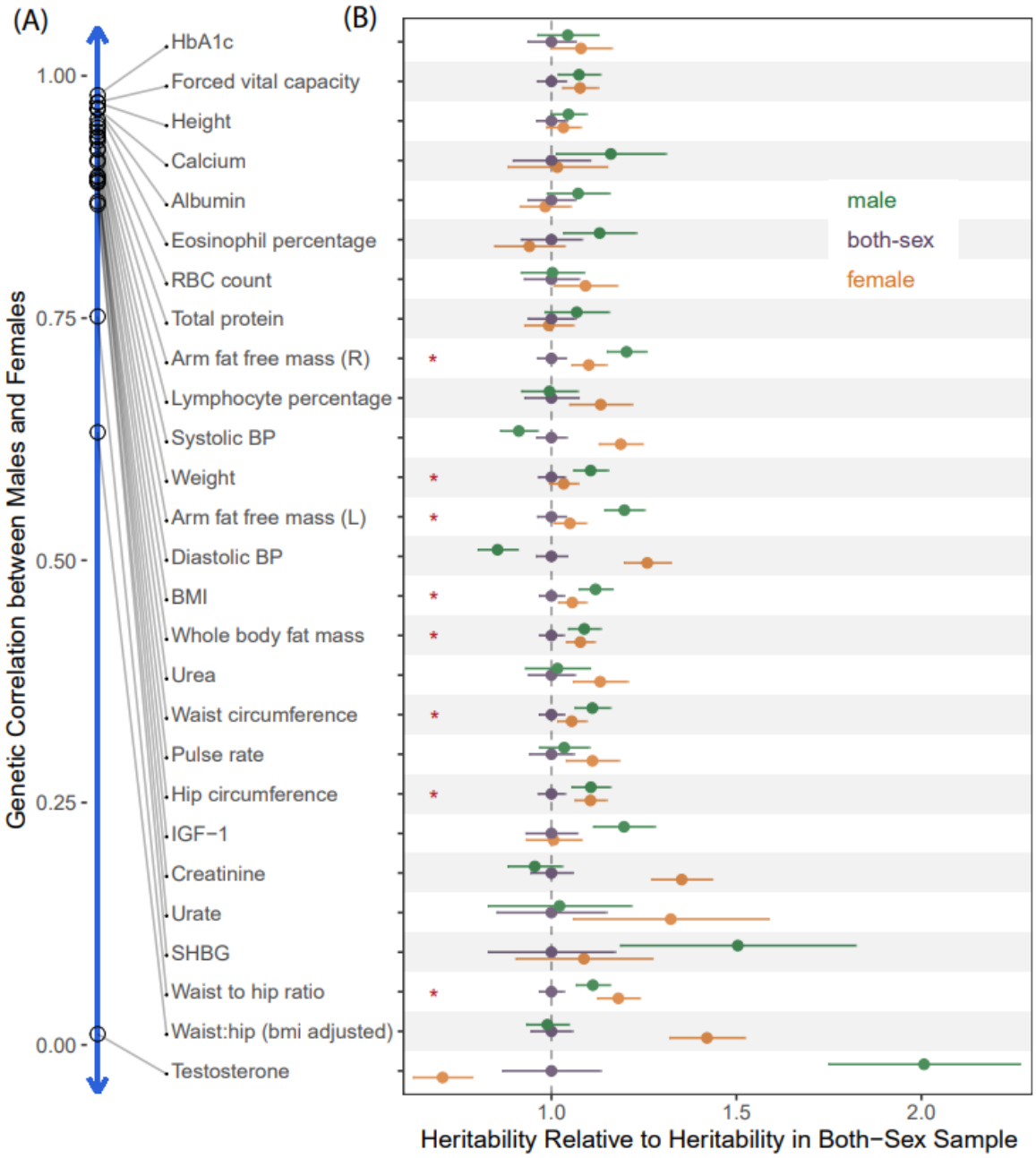


Figure 1: Heritabilities and Genetic Correlations Cannot Fully Distinguish Models of GxSex. (A) Genetic correlations between the male and females, estimated using bi-variate LD Score Regression, are shown in descending order. **(B)** The x-axis represents the relative heritability, i.e. the SNP heritability divided by the SNP heritability estimated in the sample with both sexes combined; therefore, the point estimates for the purple values are set to 1. Red asterisks show body-mass related traits with greater heritabilities in both sex-specific samples than in the sample combining both sexes.

In addition, we noticed a trend in which, as the genetic correlation decreased, the difference between the heritabilities within each sex and in the sample combining both sexes tended to become larger (Pearson r = -0.88, paired t-test p = 10^{-10}, **Fig. 1**). Nonetheless, several traits with genetic correlation above 0.9 also present relatively large sex differences in heritability: For example, diastolic blood pressure and arm fat-free mass (left) had differences of 5.2% (two-sample z-test p=$3 \cdot 10^{-6}$) and 3.4% (two-sample z-test p = 0.04), respectively. These examples are incompatible with a model of pervasive uncorrelated genetic effects driving sex-specific genetic contributions to variation in the trait (**Table 1**, second model).

We therefore considered two other alternative hypotheses under a simple additive model of variance in a trait Y within each sex,

$$Var[Y_z] = Var[G_z] + Var[E_z], \tag{1}$$

where Y_z, G_z, E_z represent the trait value, additive effect, and environmental effect (including all non-genetic context aside from sex) in sex z, respectively. Under this model, the sex-specific heritability h_z^2 is

$$h_z^2 = \frac{Var[G_z]}{Var[G_z] + Var[E_z]}, \tag{2}$$

and therefore, sex differences in heritability are either due to sex differences in genetic variance, in environmental variance, or both. If genetic effects are similar, differences in environmental variance alone could cause heritability differences (**Table 1**, first model). But as we show in the **Methods** section, under such a model, the heritability in the combined sample cannot be smaller than both sex-specific heritabilities.

Therefore, the observation of higher sex-specific heritabilities for most traits suggests that the genetic variance must differ between males and females. Given the random segregation of autosomal alleles, independent of an individual's sex chromosome karyotype, and assuming, further, that there is little-to-no interaction of sex and genotype affecting participation in the UKB[38], UKB allele frequencies in males and females are expected to be very similar. Thus, this observation suggests that causal genetic effects differ between males and females for most traits analyzed.

A less appreciated possibility for GxSex is via amplification, namely that the identity and direction of effects are shared between sexes (leading to high genetic correlation), but with

pervasive differences in the magnitude of genetic effects—e.g., larger genetic effects on blood pressure in females—which in turn lead to differences in genetic variance (**Table 1**, third model).

We can test the hypothesis that amplification acts systematically—across a large fraction of causal variants—by examining the effects of polygenic scores, genetic predictors of a complex trait. Under this hypothesis, regardless of whether the polygenic score is estimated in a sample of males, females, or a combined sample of both males and females, it should be predictive in both sexes, since the causal variants and the direction of their effects are shared (**Table 1**, third model). At the same time, in the sex for which genetic effects are larger, the effect of the polygenic score is expected to be larger. To evaluate evidence for the systematic amplification model, we estimated polygenic scores based on our sex-specific GWASs, and examined their effect in both sexes. For some traits, like albumin and lymphocyte percentage, the effects of the same polygenic score on trait value in males and females were statistically indistinguishable (**Fig. 2A,E,I,J**). In a few other traits, such as diastolic blood pressure the result was contingent on the sample in which the polygenic score was estimated (**Fig. 2C,G,I,J**). However, for most traits examined, regardless of whether the polygenic score was estimated in males or females, the effect of the polygenic score was predictive in both sexes yet significantly larger in one of the sexes (19/27 traits with t-test $p < 0.05$ using the male polygenic score. 15/27 using the female polygenic score; **Fig. 2B,D,F,H,I,J**). These observations are consistent with systematic amplification.

The results presented in **Figs. 1,2** suggested to us that both amplification and other modes of polygenic GxSex ought to be jointly evaluated. This motivated us to estimate the covariance between genetic effects in males and females directly, and further examine whether multiple, distinct GxSex patterns may exist across subsets of genetic factors affecting a trait (**Table 1**, fourth model).

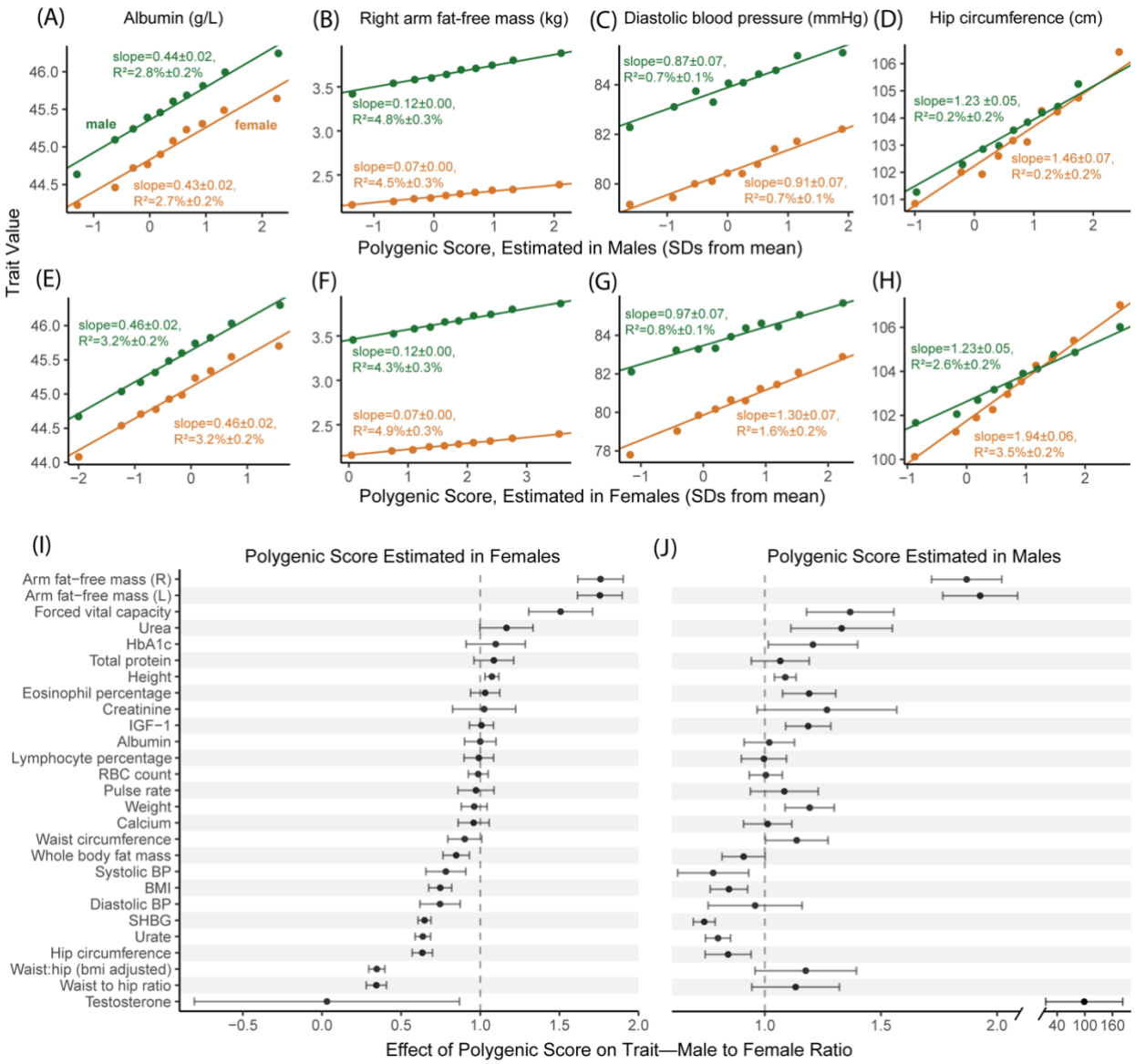


Figure 2: Evaluating evidence for systematic amplification. (A-D) We regressed trait values in males (green) and separately in females (orange) to a polygenic score estimated in an independent sample of males. Points show mean values in one decile of the polygenic score; the fitted line and associated effect estimate and R^2 correspond to regressions on the raw, non-binned data. In some traits, like Albumin (A), the polygenic score has a similar effect on the trait in both sexes. In other traits (B,D), the estimated effect of the polygenic score differs significantly, consistent with a substantial difference in the magnitude of genetic effects of sites included in the polygenic score. **(E-H)** Same analysis as A-D, but with a polygenic score pre-estimated in an independent sample of females. **(I-J)** Summary of the ratio of the effect of the polygenic score on the trait (±2 SE) in males to the effect in females across physiological traits.

Amplification of genetic effects is the primary mode of GxSex. We set to infer the mixture of covariance relationships of genetic effects among the sexes directly, based on the marginal SNP effect estimates from our sex-stratified GWAS. We analyzed all traits in their raw measurement units as provided by the UKB. In particular, we did not normalize or standardize phenotypes within each sex before performing the sex-stratified GWAS, because sex differences in trait variance may be partly due to amplification. Standardization would have therefore resulted in masking amplification signals that may exist in the data. In some cases, this is indeed the purpose of standardization[13]. More generally, while each scaling choice has it merits, we view the measurement of genetic effects in their raw units as the most biologically interpretable.

We used multivariate adaptive shrinkage (*mash*)[61], a tool that allows the inference of genome-wide frequencies of genetic covariance relationships. Namely, we model the marginal SNP effect estimates as sampled (with SNP-specific, sex-specific noise) from a mixture of zero-centered Normal distributions with various prespecified covariance relationships (2x2 Variance-Covariance matrices for male and female effects; Eq. 1 in Urbut et al.[61]). Our prespecified covariance matrices ("hypothesis matrices") span a wide array of amplification and correlation relationships, and use *mash* to estimate the mixture weights. Loosely, these weights can be interpreted as the proportion of variants that follow the pattern specified by the covariance matrix (**Fig. 3B, C**). Our covariance matrices ranged from -1 to 1 in correlation, and 11 levels of magnitude in females relative to males, including matrices corresponding to no effect in one or both sexes (**Fig. 3B**).

Inferring Polygenic Covariance Structure between Males and Females

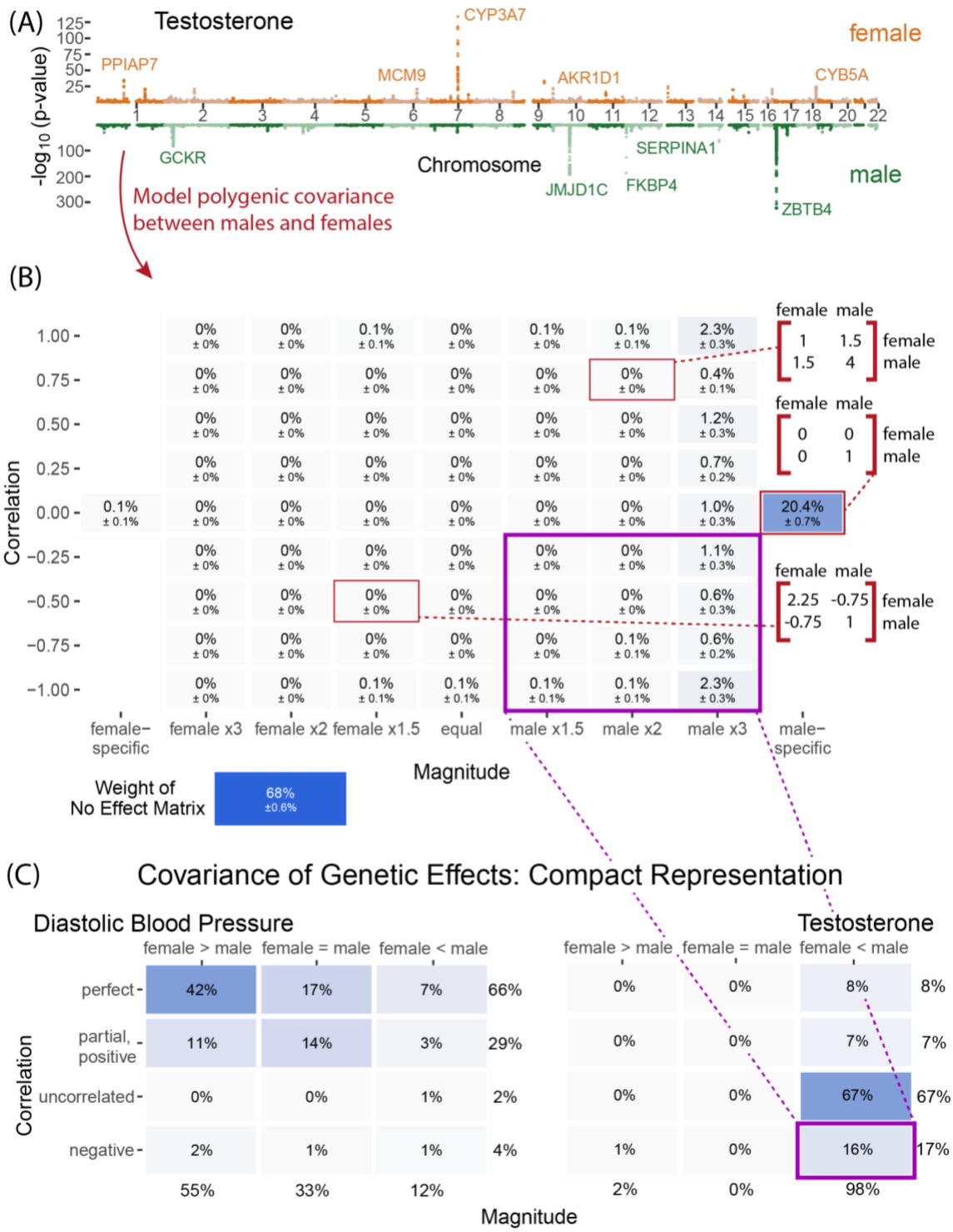


Figure 3: Inferring polygenic covariance structure between males and females. (A) Our analysis of the polygenic covariance between males and females is based on sex-stratified GWAS. Shown for illustration, is a "Miami plot" for testosterone. Highly-associated SNPs that are less than 5kbp away from a transcription start site were annotated with the gene corresponding to the closest one. **(B)** We modelled the sex-stratified GWAS estimates as sampled with error from true effects arising from a mixture of hypothesized covariance relationships between female and male genetic effects; see examples in red frames. Each box shows the weight (± SE) that we infer for one hypothesis matrix (hypothesis about covariance of causal effects in males and females). Each weight, also indicated by the shade, corresponds to the probability that a variant's effects adhere to that covariance relationship. The boxes are shown along axes indicating the relative magnitude (amplification) and the correlation between males and females, which together fully specify the covariance relationship. **(C)** The x and y axes are a condensed version of the x and y axes from (B) for testosterone, as well as the condensed version for diastolic blood pressure. The weights are the percentage of non-null weights, i.e., the weight divided by the sum of all weights except for the weight on the all-zeros matrix corresponding to no effect in either sex. For example, the square in purple sums over all 12 weights for matrices corresponding to larger effects on testosterone in males that are negatively correlated with effects in females, 5.1%, divided by the total weight on matrices with nonzero effects, 32%.

We first focus on testosterone, for which previous research sets the expectation for polygenic male-female covariance. In terms of magnitude, the vast majority of effects should have much greater effect in males. In terms of correlation, we expect a class of genetic effects acting through largely independent and uncorrelated pathways alongside a class of effects via shared pathways[32]. Independent pathways include the role of hypothalamic-pituitary-gonadal axis in male testosterone regulation and the contrasting role of the adrenal gland in female testosterone production. Shared pathways involve sex hormone-binding globulin (SHBG), which decreases the amount of bioavailable testosterone in both males and females. As expected, we found that mixture weights for testosterone concentrated on greater magnitudes in males and largely uncorrelated effects. Out of the 32% total weights on matrices with an effect in at least one sex, 98% of the weights were placed on matrices representing larger effects in males, including 20.4% (± 0.7%) having no effect in females at all (**Fig. 3B,C**).

While testosterone is unusual in the relative lack of correlation between sexes and the extent of male-lopsided amplification, about half (13/27) of the traits analyzed had the majority of weights placed on greater effects in just one of the sexes (**Fig. 4B**). For instance, BMI-adjusted waist:hip ratio showed higher percentages for greater effects in females (64%; **Fig. S4AA**) and arm fat-free mass (right) showed higher percentages of greater effects in males (92%; **Fig. S4C**). Both traits had weights concentrated on highly correlated effects. We confirmed, using a

simulation study, that amplification signals indeed suggest sex differences in the magnitude of genetic effects and are not due to differences in the extent of estimation noise (e.g., variation in environmental factors independent of allelic effects; **Fig. S6**; **Methods**).

In addition, the only trait of the 27 for which a large fraction (≥10%) of the non-zero weights were on matrices with negative correlation was testosterone (17%). Most weights were rather on perfect or highly positive correlation matrices. For example, diastolic blood pressure and eosinophil percentage had 66% (**Fig. 3C**) and 68% (**Fig. S4H**), respectively, of the non-zero weights on matrices representing a perfect correlation. Overall, the low weights on matrices representing negative correlation do not support opposite directions of effects being a major mode of GxSex (**Fig. S7**).

Across traits, the difference between male-larger and female-larger weights correlates strongly with male-to-female phenotypic variance ratio (Pearson $r = 0.873$, $p = 6 \cdot 10^{-9}$ after removing testosterone as an outlier; **Fig. 4B**). This observation is consistent with our hypothesis of amplification leading to differences in genetic variance between sexes and thereby contributing substantially to sex differences in phenotypic variance. Together, these observations point to amplification, rather than uncorrelated effects, as a primary mode of polygenic GxSex.

The pervasiveness of GxSex, alongside the mixture of covariance relationships across the genome for many traits, may mean that GxSex is important to consider in phenotypic prediction. However, we found but a slight improvement in predictive performance of GxSex-aware polygenic scores and only for 2/27 traits analyzed (**Supplementary Materials**; **Fig. S14**). We suggest a possible reason for this observation in the **Discussion** section.

Another important question about the implication of pervasive amplification is whether it is a major driver of mean phenotypic differences. The ratio between male and female mean phenotypic values is strongly correlated with the difference between male-larger and female-larger amplification (Pearson $r=0.75$; $p=2 \cdot 10^{-5}$ after removing testosterone and BMI-adjusted waist:hip ratio as outliers); but since both are strongly correlated with phenotypic variance ratios (**Fig. 4A,B**), it is hard to discern a causal hypothesis for this relationship.

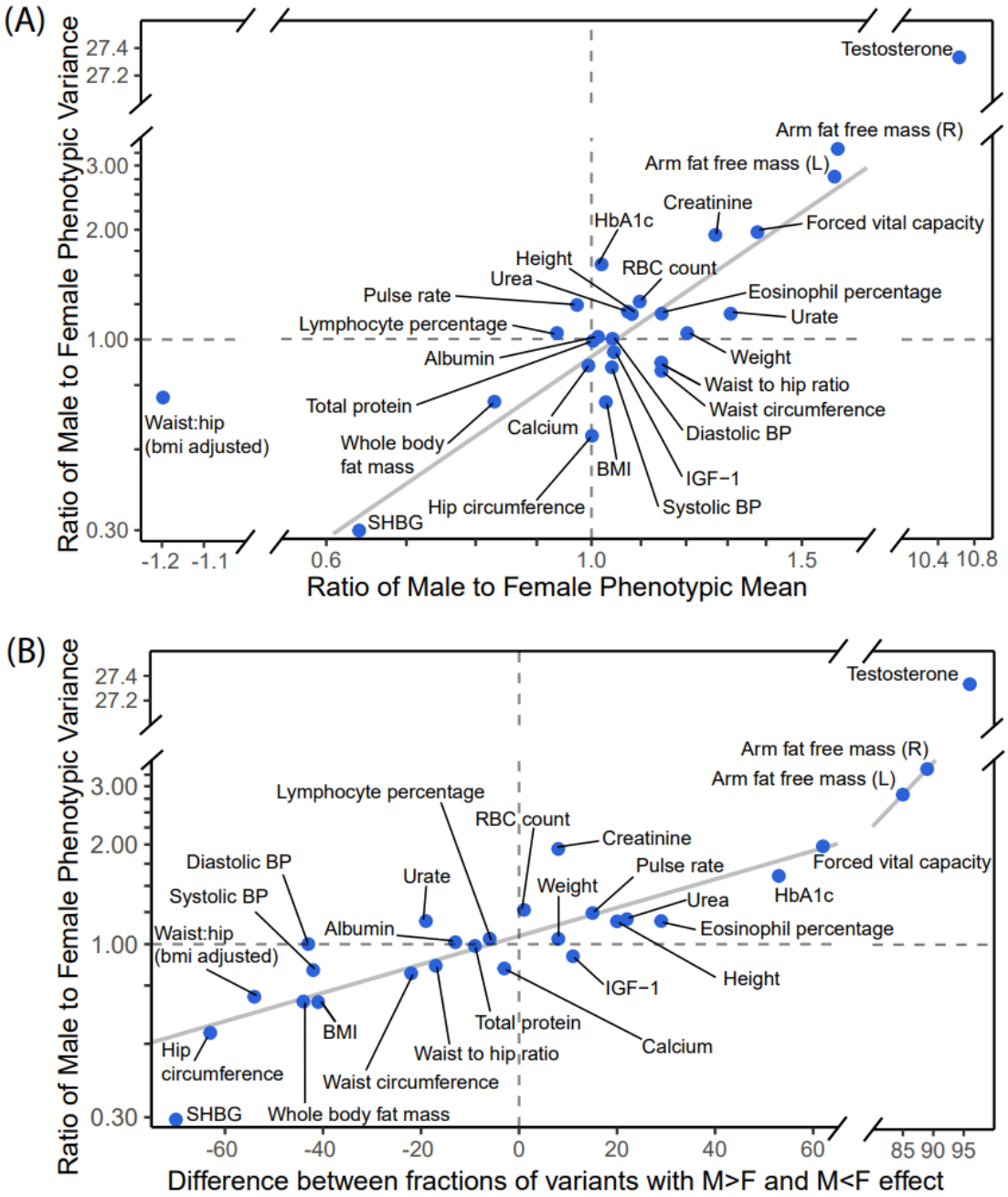


Figure 4. (A) Across traits, phenotypic variance sex ratio correlates with the phenotypic mean sex ratio. The solid gray line shows a linear fit, excluding two outliers: Testosterone and waist:hip ratio adjusted for BMI. **(B) Phenotypic variance strongly correlates with amplification.** The x-axis summarized the sex-biased amplification of polygenic effects, and is calculated by taking the difference between the sum of weights on matrices with male effects greater in magnitude than female effects (M>F) and the sum of weights of M<F matrices. The solid gray line shows a linear fit across traits, excluding testosterone as an outlier.

Testosterone as a regulating factor. Thus far, we treated the genetic interaction as discretely mediated by biological sex. One mechanism that may underlie GxSex is a cue or exposure that modulates the magnitude (and less often, the direction) of genetic effects, and varies in its distribution between the sexes. A natural candidate is testosterone. Testosterone may be a key instigator since the hormone is present in distinctive pathways and levels between the sexes and a known contributor to the development of male secondary characteristics, so therefore could modulate genetic causes on sex-differentiated traits.

For each sex, we first binned individuals by their testosterone levels. Then, for each trait, we quantified the amplification of total genetic effect as the linear regression coefficient, within the bin, of phenotypes to the additive PGS for the trait (**Methods**; see **Fig. S10** for results obtained using sex-specific PGS). For several traits—notably, traits related to body mass—testosterone (mean per bin) and the magnitude of genetic effect were correlated for both males and females. BMI and whole body fat mass (**Fig. 5A**) showed significant correlations (Pearson $p < 0.05$) for both sexes. For all the body-mass-related traits, there was a negative correlation between the magnitude of genetic effect and testosterone levels for males and a positive correlation for females (**Fig. 5B**); for example, the Pearson correlation coefficient for whole body fat mass in males was -0.86 ($p = 0.001$) and 0.76 for females ($p = 0.011$). Since the relationship with testosterone remains contingent on sex, a model of testosterone as the sole driver of the observed sex-specificity would be invalid. These observations may help explain previous reports of positive correlations between obesity and free testosterone in women, and negative correlations in men[62]. We conclude that in body-mass related traits, testosterone may be modulating genetic effects in a sexually antagonistic manner.

We performed two additional analyses designed to control for possible caveats to the association of testosterone and the magnitude of polygenic effect. An association test that controls for possible confounding with age (**Fig. S12**), and a test that mitigates confounding with other variables or reverse causality (wherein the magnitude of polygenic effect affecting the focal trait causally affecting testosterone levels; **Fig. S11**). The evidence for an effect of testosterone on the magnitude of polygenic effect did not remain significant in either of these tests. It is possible, however, that this was due to low statistical power (**Methods**).

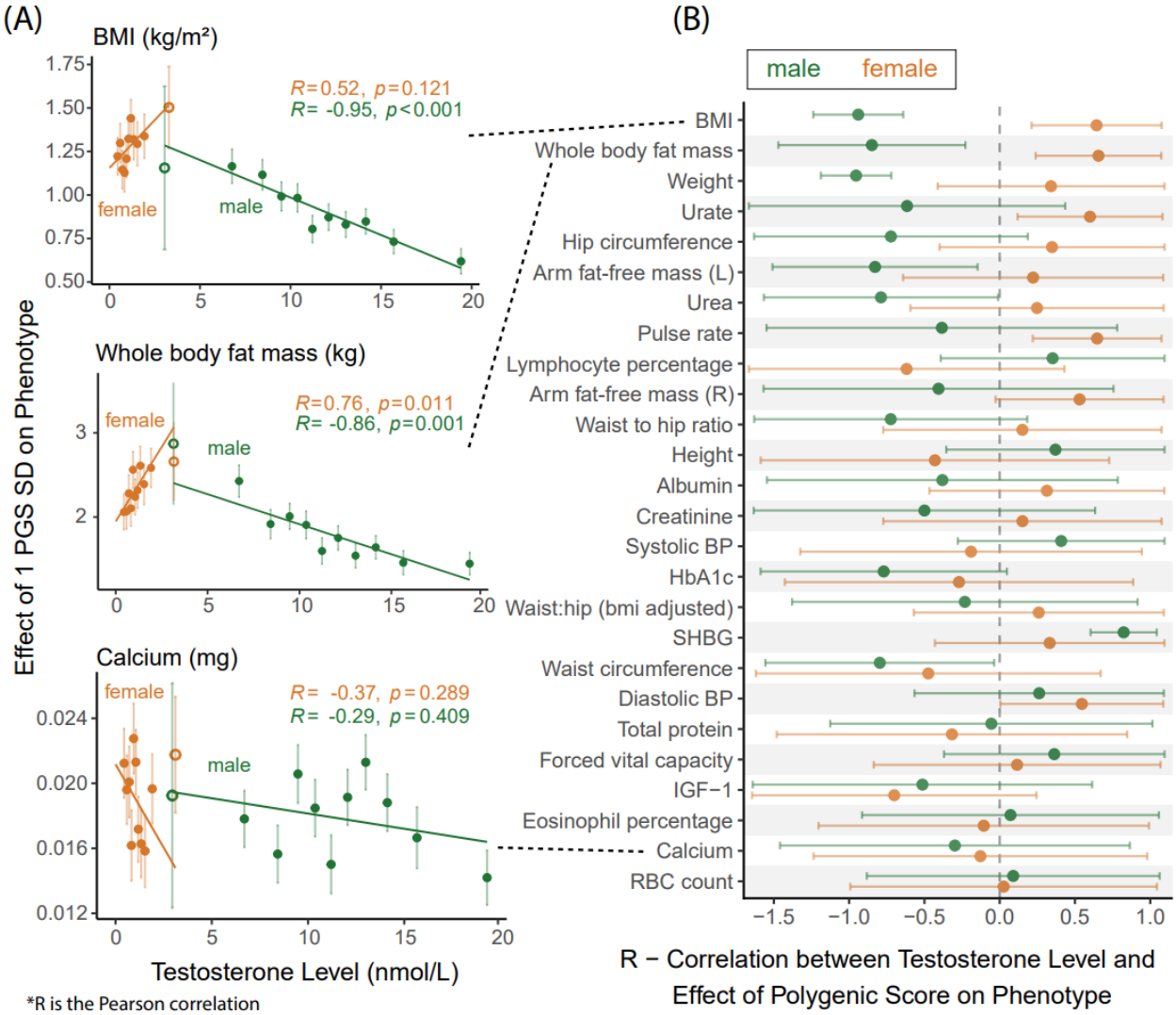


Figure 5. Amplification of total genetic effect in relation to testosterone levels. (A) The relationship between testosterone level bins and estimated magnitude of genetic effect on traits is shown for three traits. The magnitude of genetic effect is estimated using the slope of the regression of phenotypic values to polygenic scores in that bin. The units on the y-axis are effect per standard deviations (SD) of the polygenic scores across all individuals in all bins. The hollow data points are bins with overlapping testosterone ranges between males and females; these are based on fewer individuals (~800 compared to ~2200 in other bins) and not included in the regression. **Fig. S9** show all other traits analyzed. **(B)** The correlation for each sex (90% CI) are shown for all 27 traits. Traits are ordered in descending order of male-female differences in R.

A model of shared polygenic and environmental amplification. Our results thus far suggest that polygenic amplification across sexes is pervasive across traits; and that the ratio of phenotypic variance scales with amplification (**Fig. 4B**). An immediate question of interest is whether the same modulators that act on the magnitude of genetic effects act on environmental effects as well (see also relevant discussion by Domingue et al.[63]). Consider the example of human skeletal muscle. The impact of resistance exercise varies between males and females. Resistance exercise can be considered as an environmental effect since it upregulates multiple skeletal muscle genes present in both males and females such as IGF-1, which in turn is involved in muscle growth[64]. However, after resistance exercise at similar intensities, upregulation of such genes is sustained in males, while levels return sooner to the resting state in females (**Fig. S13**). It is plausible that modulators of the effect of IGF-1, such as insulin or sex hormones, drive a difference in the magnitude of effect of core genes such as IGF-1 in a sex-specific To express this intuition with a model: If amplification mechanisms are shared, then amplification may be modeled as having the same scalar multiplier effect on genetic and environmental effects (**Fig 6A**). In the **Methods** Section, we specify the details of a null model of shared amplification, which yields the prediction that the male-female ratio of genetic variances should equal the respective ratio of environmental variances (blue line in **Fig 6B**).

We tested the fit of this theoretical prediction for the 27 traits analyzed. We used our estimates of sex-specific phenotypic variance and SNP heritabilities to estimate the ratios of genetic and environmental variances. We note that environmental variance is proxied here by all variance not due to additive genetic effects, and caution is advised with interpretation of this proxy.

Twenty of the 27 traits were consistent with the null model of equal amplification (within 90% CI; **Fig. 6B**). This finding may suggest a sharing of pathways between polygenic and environmental effects for these traits (**Fig. 6A**). Interesting exceptions include BMI-adjusted waist:hip ratio, creatinine, arm fat-free mass, and systolic and diastolic blood pressure—which was the strongest outlier ($p = 3.06 \cdot 10^{-12}$, single-sample z-test), excluding testosterone.

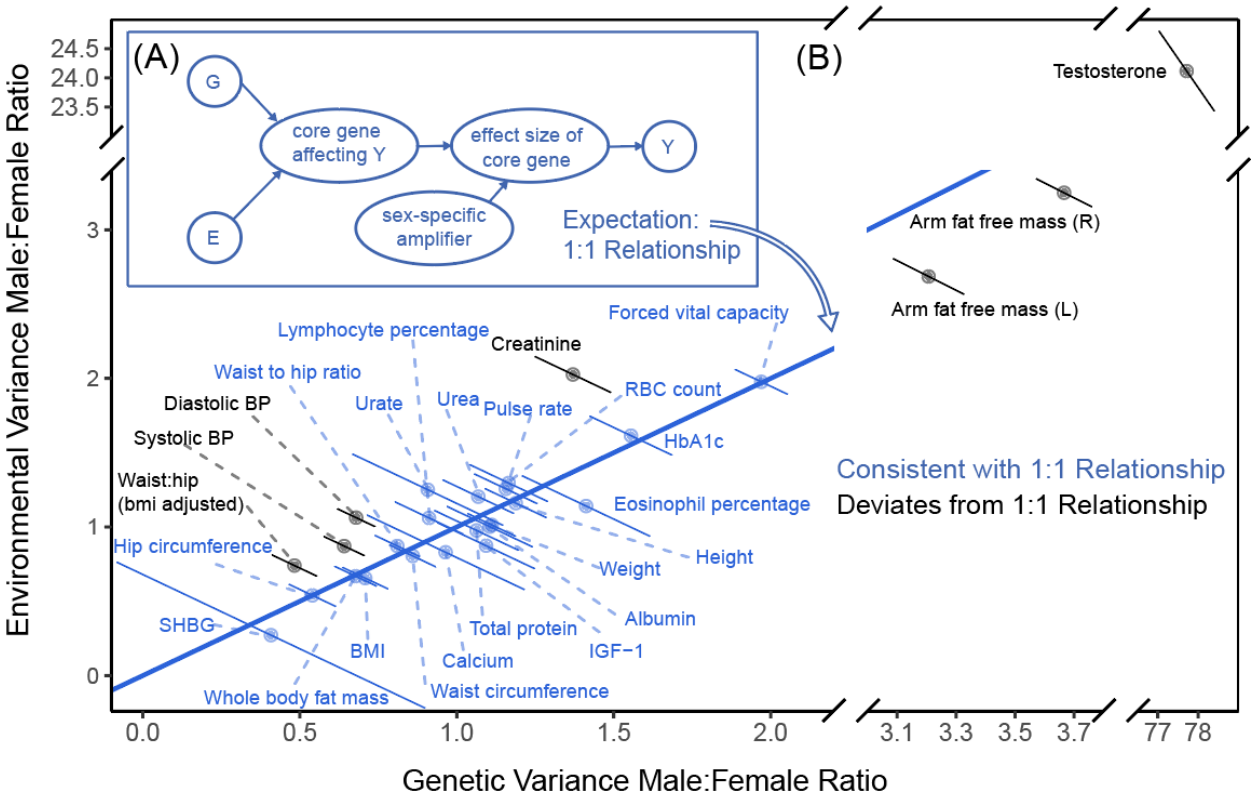


Figure 6. Testing a model of shared amplification between environmental and polygenic effects. **(A)** A model of equal amplification of genetic (G) and environmental (E) effect, that produces the dimorphism shown in the phenotype, Y. G and E both act through a core gene that is amplified in a sex-specific manner. **(B)** The blue 1:1 line depicts the theoretical expectation under a simple model of equal amplification of genetic and environmental effects in males compared to females. Error bars show 90% confidence intervals. Traits in blue are consistent (within 90% CI) with the prediction of a 1:1 relationship.

Sexually-antagonistic selection. A hypothesized cause of sexual dimorphism is sexually-antagonistic selection, where some alleles are beneficial in one sex yet deleterious in the other[23,24,26,44,65]. Sexually-antagonistic selection is difficult to study using traditional population genetics methods because Mendelian inheritance equalizes autosomal allele frequencies between the sexes at conception, thereby erasing informative signals. One way around this limitation is to examine allele-frequency differences between the sexes in the current generation, known as "selection in real time"[26,66,67]. In this section, we consider a model of sexually-antagonistic selection acting on a polygenic trait and use it to estimate the strength of contemporary viability selection acting on the 27 traits we analyzed.

Most theoretical models of sexually-antagonistic selection on a trait under stabilizing selection usually intuit either highly distinct male and female fitness optima or genetic variants affecting traits antagonistically. Our findings on pervasive amplification suggest that variant effects on traits tend to have concordant signs. Yet, under pervasive amplification, a somewhat-surprising intuition arises: Alleles affecting a trait may frequently experience sexually antagonistic selection—both in the case where trait optima for males and females are very distinct (**Fig. 7B**) and for the case that they are similar (**Fig. 7A**).

We developed a theoretical model of sexually-antagonistic viability selection on a single trait which extends on this intuition. The model relates sex-specific effects on a complex trait to the divergence in allele frequency between males and females (measured as F_{ST}[68,69]) due to viability selection "in real time", i.e., acting between conception and the age at the time of sampling. We derive the expected relationship for each site i,

$$F_{ST\,i} \approx AV_{GxS_i}, \tag{3}$$

where

$$V_{\{GxS\}_i} = 2p_i(1-p_i)\left(\beta_i^m - \beta_i^f\right)^2,$$

and p_i, β_i^m and β_i^f are the allele frequency of an allele at site i, its effect on the trait in males and its effect in females, respectively. A is a constant parameter shared across all variants and can therefore be interpreted as the effect of sexually-antagonistic selection on male-female divergence at polymorphic sites associated with variation in that trait (**Methods**). We estimated $F_{ST\,i}$ for all sites i across subsamples of various ancestry groups in the gnomAD dataset [70]. To estimate $V_{\{GxS\}_i}$ at each site and for each trait, we used our sex-stratified GWAS results. Since there is large heterogeneity in uncertainty of GxSex-genetic variance estimates, we use a variance-weighted linear regression to estimate A (see **Methods** and **Supplementary Materials** for the derivation of the variance of $V_{\{GxS\}_i}$ estimates and for further details on the estimation procedure).

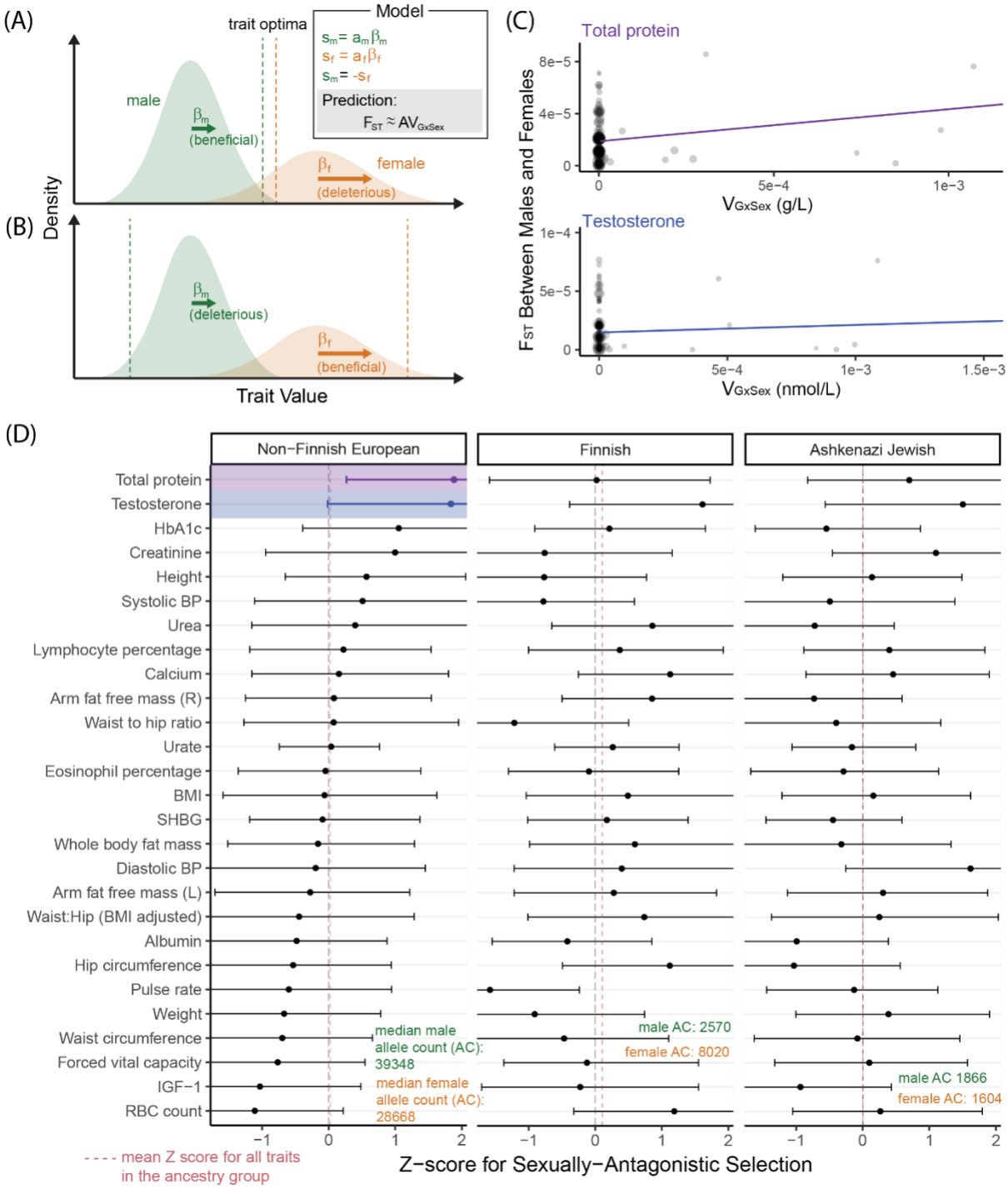


Figure 7. Testing for sexually-antagonistic selection. (A,B) A model of sexually-antagonistic selection. Selection coefficients, s_m and s_f, are linear with the additive effect on the trait in each sex. Sexually-antagonistic selection acts

such that $s_m = -s_f$. The model yields the prediction of **Eq. 3.** In (A), the effect of an allele tends to drive trait values towards the optimum in males, and away from the optimum in females. In (B), the fitness optima are farther in males and females; in both examples, selection on acts antagonistically (i.e., in opposite directions). **(C)** Two examples of the weighted least-squares linear regression preformed to estimate the strength of sexually-antagonistic selection on variants associated with a trait (A in panel A and **Eq. 3**). Each point shows one SNP. Size is proportional to each point's regression weight. **(D)** Z-scores (90% non-parametric bootstrap CI) estimated through 1000 resampling iterations of the weighted linear regression of panel B for each trait. The two colored estimates correspond to the examples in (C).

Recent work has shown that apparent sex differences in autosomal allele frequencies within a sample are often due to a bioinformatic artifact: The mismapping of sequencing reads from autosomes to sex chromosomes or vice versa[43,44]. We identified and excluded sites which are potentially vulnerable to this artifact (**Supplementary Materials**). In **Fig. 7D,** we only show results for gnomAD subsamples that are the closest in their genetic ancestry to our UKB sample[71] (results for other subsamples are shown in **Fig. S15**). Furthermore, given the concerns of study-specific recruitment biases [38,43], we place higher confidence in results that replicate qualitatively across different subsamples, even though we note that subsample-specific selection signals may be real since sexually-antagonistic selection may act heterogeneously across groups.

Across traits and samples, we only find weak evidence for sexually-antagonistic polygenic selection. In the non-Finnish sample, the largest of the three samples, the null hypothesis $H_0: A = 0$ in **Eq. 3** is rejected ($p<0.05$) only marginally for testosterone (Z score=1.8) and total protein (Z=1.9). Notably, while the total protein signal does not replicate, testosterone is among the two strongest signals in all three samples.

Discussion

Departing from previous studies that sought GxSex (or GxE more broadly) through single loci or heritability analysis, we evaluated GxSex through a polygenic lens and modelled it using a mixture of relationships across the genomes. Our analysis supports pervasive context-dependency of genetic effects on complex traits, acting largely through the systematic amplification of effects. Surprisingly, even for some traits such as red blood cell count, previously considered non-sex-specific because of high genetic correlations between sexes and a concordance in top GWAS hits, we find evidence for substantial GxSex through amplification. The strong relationship

between amplification and phenotypic variance further points to its potential importance for sexual dimorphism.

This refined understanding of polygenic GxSex did not carry over to improved predictive ability. Additive models still had higher prediction accuracy for most traits examined. One possible explanation is that the accurate estimation of the additive effect is still more advantageous for prediction than the improved estimation of non-additive effects.

Future efforts localizing amplification in the genome may provide clues to the mechanisms underlying this form of GxSex. Here, we proposed one such modulator, testosterone, and found a correlation between testosterone levels and the magnitude of genetic effect on multiple traits related to body mass. The opposite signs of these correlations in females and males may reflect the discrepant relationship between testosterone and these traits at the phenotypic level.

Finally, we developed a model to consider how GxSex may fuel sexually-antagonistic selection in contemporary populations. Over long evolutionary timescales, the two scenarios depicted in **Fig. 7A,B** lead to different predictions about the maintenance of GxSex genetic variance: With distinct optima in the two sexes, this variance should be maintained, whereas with similar optima, GxSex variance should be purged. Interestingly, however, in both cases, alleles that underlie GxSex may experience sexually-antagonistic selection.

In samples from two of three ancestry groups, we found subtle signals of sexually-antagonistic selection on variation associated with testosterone levels (also see related results by Ruzicka et al.[67]). The signal for our inference of selection is systematic allele frequency differences between adult males and females, which we presume to be due to contemporary viability selection. These signals may therefore point to a related disease that differentially affects lifespan in the two sexes, such as immune system suppression, diabetes, cancers, or hypertension[74–77]. Recently, high testosterone levels have been linked to increased rates of mortality and cancer in women, but decreased rates in men[72,73]. However, it is important to note that subtle recruitment biases affecting male and female participation differently remain an important caveat to our conclusions. The replication of signals of selection in multiple samples may lend credence to our inference, but in medical datasets based on recruitment of participants via referring physicians, participation biases may still plausibly be shared across studies.

In this work, we have shown that amplification of the magnitude of polygenic effects may be important to consider as a driver of sex differences and their evolution. Our approach included

the flexible modelling of genetic effect covariance among the sexes, as well as various subsequent analyses exploring the implications of these covariance structures. We hope this study can inform future work on the context-specificity of genetic effects on complex traits.

Acknowledgments

We thank Doc Edge and the Edge Lab members, Ziyue Gao, Tom Juenger, Jonathan Pritchard, Molly Przeworski, Guy Sella, Jeff Spence and Elliot Tucker-Drob for their helpful comments on the manuscript. We also thank Brian Dilkes, Andrés Bendesky and Jim Fleet for insightful discussions. This work was supported by NIH GM116853-07 to MK. This study has been conducted using the UK Biobank resource under application Number 61666, as approved by the University of Texas at Austin Institutional Review Board, protocol 2019-02-0125.

Methods

UK Biobank sample characteristics. The UK Biobank is an extensive database that contains a wide variety of phenotypic and genotypic information of around half a million participants aged 40-69 at recruitment[78].

In this study, we considered 337,111 individuals who passed quality control (QC) checks, which included the removal of samples identified by the UK Biobank with sex chromosome aneuploidy or self-reported sex differing from sex determined from genotyping analysis. We excluded related individuals (3rd-degree relatives or closer) as identified by the UK Biobank in data field 22020. To reduce potential population structure confounding, we further limited our sample to individuals identified by the UK Biobank as “White British” in data field 22006. These are individuals who both self-identified as White and as British and were additionally very tightly clustered in the genetic principal component space[78,79]. Individuals who had withdrawn from the UK Biobank by the time of this study were removed. For each phenotype, we also removed individuals who had missing data for the specified phenotype. These procedures left us with between 255,426 to 336,551 individuals in the analysis for each trait.

Expectations for sex-specific heritabilities with no GxSex. In the section "The limited scope of analyzing GxSex via heritability differences and genetic correlations," we report our observation that, for most traits examined, sex-specific heritabilities, h_m^2 and h_f^2 are both higher than the heritability in the combined sample, h^2. Here, we explain why this observation is inconsistent with a simple model in which genetic effects are the same across the sexes.

Denote the phenotype again by Y, the polygenic effect by G, the environmental effect by E and the sex by $Z \in \{\text{male}, \text{female}\}$. We assume as before that allele frequencies are highly similar between males and females. Since genetic effects are equal, this implies

$$Var[G|Z = m] \approx Var[G|Z = f].$$

For the environmental variance, we have that

$$Var[E] = \sum_{Z\in\{m,f\}} \int_E dE(E - \mathbb{E}[E])^2\, \mathbb{P}(Z, E) = \sum_{Z\in\{m,f\}} \mathbb{P}(Z) \int_E dE(E - \mathbb{E}[E|Z])^2\, \mathbb{P}(E|Z) = \tag{4}$$
$$\mathbb{P}(Z = m)Var[E|Z = m] + \mathbb{P}(Z = f)Var[E|Z = f] \leq \max_{z\in\{m,f\}} Var[E|Z = z],$$

where the second equality, we have assumed that there are no mean sex differences in the environmental effects (or, in practice, have been corrected for via the correction for the mean phenotypic difference between the sexes, giving

$$\mathbb{E}[E|Z = m] = \mathbb{E}[E|Z = f] = \mathbb{E}[E].$$

Eq. 4 shows that the environmental variance cannot be greater than the larger of the two sex-specific environmental variance. It follows that if the genetic variance is equal in both sexes, then the heritability in the combined sample cannot be smaller than both of the sex-specific heritabilities,

$$h^2 = \frac{Var[G]}{Var[G] + Var[E]} \geq \frac{Var[G]}{Var[G] + \max_{z\in\{m,f\}} Var[E|Z]} = \min_{z\in\{m,f\}} h_Z^2. \tag{5}$$

Multivariate adaptive shrinkage (mash). We used multivariate adaptive shrinkage (*mash*) to examine correlation and differences in magnitude of SNP effects between males and females [61]. *mash* is an adaptive shrinkage method[80] that improves upon previous methods of estimating and

comparing effects across multiple conditions by flexibly allowing for a mixture of effect covariance patterns between conditions and requiring only summary statistics from each condition (including a point estimate of the effect and corresponding standard error for each SNP and condition). The method adapts to patterns of sparsity, sharing, and correlation among the conditions to compute improved effect estimates.

In this study, we set two conditions, male and female, and provided effect estimates and corresponding standard errors from our male-specific and female-specific GWAS. *mash* learns from the data by estimating mixture proportions of various predefined covariance matrices representing different patterns in effects. Using maximum-likelihood, mash assigns low weights to matrices that capture fewer patterns in the data, and higher weights to those that capture more.

Mixture weights for covariance structure between male and female effects. To interpret patterns of SNP effects between males and females, we inputted 66 hypothesis-based covariance matrices (**Fig. S2**) spanning a range of correlations and relative magnitudes of effects between males and females. We used a random subset of all SNPs for mash to learn the covariance mixture weights. In order for the random subset to contain approximately independent SNPs and capture the weight of SNPs with no effect (**Fig. S2**), we created a subset of SNPs for each trait by taking a random SNP from each of 1703 approximately independent LD blocks estimated for Europeans[81]. *mash* can also generate data-driven covariance matrices that capture SNP effects in the data, but we did not use this feature since the data-driven matrices had negligible differences from our hypothesized matrices (in terms of $\ell 2$ norm) and were less interpretable.
For each trait, we repeat this weight-learning step 100 times, sampling the SNPs from the 1703 LD blocks without replacement to fit the mash model and generate mixture proportions. We then take the average proportion for each covariance matrix as an estimate of its weight, effectively treating each of the 100 samples as i.i.d. draws.

Choice of SNPs used to estimate male-female effect covariance. We examined the effect of using a random subset taken from different p-value thresholds [1, 5e-2, 1e-5, 5e-8] while selecting from LD blocks. By doing so, we can examine differences in the distribution of weights across the p-value thresholds. We performed this test on height, BMI, testosterone, and BMI-adjusted waist:hip ratio. For each trait, weight placed on the no-effect matrix decreased as we

reduced the p-value threshold (**Fig. S3**). Patterns of weights for non-null effect matrices varied across the traits (**Fig. S3**). Since *mash* considers the proportion of null effects and sex-specific, SNP-specific noise; together with the fact that for complex traits, less significant associations may still reflect valuable signal, we decided on using the whole set of SNPs to sample from when estimating mixture proportions.

Simulating equal genetic effects and heterogeneous estimation noise among the sexes.

To ensure that *mash* was not mistaking sex differences in estimation noise (e.g.m via differences in the extent of environmental variance) to be differences in the magnitude of genetic effects, we performed a simulation study. In short, samples of males and females were generated under the model given by **Eq. 1**. Genetic effects were set as equal but the environmental variance differed among the sexes. We then perform a GWAS on both samples and input the simulated GWAS results into *mash*, and test whether the estimated mixture weights spuriously suggest the presence of GxSex. We performed this simulation on a grid of parameters, including heritabilities in males set to either 5% or 50%, female to male environmental variance ratio of 1, 1.5 or 5; and 100, 1,000 or 10,000 causal SNPs.

First, we created a sample of 300K individuals with randomly assigned sex. We then created a matrix of genotypes for all individuals consisting of 20K SNPs by sampling from the observed distribution of allele frequencies from UK Biobank's imputed data[82]. From the 20K SNPs, we portioned out the predetermined number of causal SNPs and assigned effect sizes by sampling from a Standard Normal distribution. We estimated the male environmental variance for each causal SNP using the equation,

$$Var[E_m] = \frac{Var[G](1 - h_m^2)}{h_m^2} = \frac{(\sum_{i=0} \beta_i^2 2p_i(1 - p_i))(1 - h_m^2)}{h_m^2} \quad (6)$$

where $Var[E_m]$ is the simulated environmental variance for males, G is a vector of the genetic effects, h_m^2 is the heritability in males and β_i and p_i are the effect size and allele frequency at site i. We multiplied $Var[E_m]$ by the environmental variance ratio to obtain the environmental variance for females, $Var[E_f]$. Afterwards, for each individual, we sampled environmental effects from a Normal distribution with mean zero, and variance set as $Var[E_m]$ or $Var[E_f]$ according to the sex of the individual. Phenotypes were then simulated using the following additive model,

$$y_j = \sum_{i=0} \beta_i x_{ij} + E_j \quad (7)$$

where y_j is the phenotypic value for individual j and x_{ij} is the number of effect allele copies at the ith causal SNP for the jth individual. With the phenotype, genotype and environmental effect, we obtained the estimated effect sizes, $\{\widehat{\beta}_i\}$, using least squares simple linear regression for all 20K SNPs and used the estimated effect sizes and corresponding standard errors as input into *mash*. For nearly all parameters, out of the weights on matrices with effect, the vast majority was placed on the matrix for perfect correlation, equal magnitude (**Fig S6**). As the number of causal SNPs increased, the weight on the no-effect covariance matrix decreased accordingly. These results suggest that *mash* was not grossly mistaking differences in environmental variance as amplification.

Testosterone as an amplifier. We tested a model of testosterone as a modulator of magnitude differences in males and females. We first split individuals by sex and for each sex, created 10 bins of testosterone levels. We adjusted one of the 10 bins to have testosterone levels overlap between males and females. The overlapping testosterone bin was based on fewer individuals (~800) compared to the other bins (~2200). For each trait, each of the sexes, and within each bin, we performed a simple linear regression of trait values to the PGS for the trait (using a PGS based on both-sex summary statistics (**Supplementary Materials**)). We interpret the estimated coefficient for the effect of the PGS as a proxy of the magnitude of polygenic effect. Finally, we summarized the relationship between testosterone level and magnitude of polygenic effect across bins using the Pearson correlation between the two.

To mitigate the possible effects of confounding (of testosterone and magnitude of polygenic effect) or reverse causation (the magnitude of polygenic effect on the focal trait causally affecting testosterone levels) we employed a version of Mendelian Randomization[83,84] of the same analysis (**Fig. S11**). Namely, we replaced testosterone levels of each individual with their PGS for testosterone. Here, given the near zero genetic correlation between males and females, we used our sex-specific PGS for each sex; otherwise, the analysis is unchanged.

We also examined whether participants' age may have confounded the relationship between testosterone and polygenic effect. In this analysis, instead of using the polygenic effect

as the response variable across bins, we used the polygenic effect residualized for mean age in the bin, and examined the effect of an individual's polygenic score on the residual (**Fig. S12**).

Model of Shared Amplification. Here, we suggest a null model in which amplification is shared between genetic and environmental effects. We then suggest a prediction that the model yields and explain how we tested this prediction across traits (**Fig. 6**).

If an amplifier is shared, it may be modeled as having the same scalar multiplier effect on genetic and environmental effects. Consider the within-sex additive model of **Eq. 1** in the section "The limited scope of analyzing GxSex via heritability differences and genetic correlations" above. For a phenotype value Y_z in sex $z \in \{\mathrm{m}, \mathrm{f}\}$

$$Y_z = G_z + E_z, \tag{8}$$

where E_z is the environmental effect and

$$G_z = \sum_{site\ i} x_i \beta_i^z \tag{9}$$

is the polygenic effect where β_i is the effect of the an allele at site i (say the minor allele) and x_i is the number of copies of the allele. We assume here for simplicity that male genetic effects relate to female effects solely through a shared polygenic amplification constant, α,

$$\beta_i^m = \alpha \beta_i^f \qquad \forall i;\ \ \alpha > 0. \tag{10}$$

Allele frequencies are once again assumed to be close to equal between males and females, since due to random segregation of alleles during meiosis, genotype frequencies at autosomal sites are independent of sex; and further assuming no substantial interaction between genotype and sex affecting participation in UKB[38]. Consequently, differences in polygenic effect distributions between males and females are solely based on GxSex, and thus:

$$Var[G_m] = \alpha^2 Var[G_f]. \tag{11}$$

The model we would like to test is one where the amplification of environmental effects can also be simplified to the same scalar multiplier,

$$E_m = \alpha E_f, \text{ and} \tag{12}$$
$$Var[E_m] = \alpha^2 Var[E_f].$$

Hence, with equal amplification,

$$\frac{Var[G_m]}{Var[G_f]} = \frac{Var[E_m]}{Var[E_f]} \tag{13}$$

To test the model of shared amplification between environmental and polygenic effects (**Eq. 8**) we obtained the genetic and environmental variance for males and females based on the following relationships,

$$Var[G_z] = h^2 Var[Y_z] \tag{14}$$

and

$$Var[E_z] = (1 - h^2) Var[Y_z], \tag{15}$$

where $Var[G_z], Var[E_z]$, and $Var[G_z]$ are the additive genetic, environmental, and phenotype variances, respectively. The sex-specific heritabilities, h_z^2, were obtained from previous estimates using LD Score Regression (**Supplementary Materials**).

Representing male genetic or environmental variance as x, and the corresponding female variance as y, we derived standard errors for the ratio of male to female variance using the 2nd-order Taylor approximation for the standard error of a ratio of estimators of x and y,

$$SE\left[\frac{\hat{x}}{\hat{y}}\right] = \sqrt{Var\left[\frac{\hat{x}}{\hat{y}}\right]} \cong \frac{E[\widehat{x}]}{E[\hat{y}]}\sqrt{\frac{Var[\hat{x}]}{E[\hat{x}]^2} + \frac{Var[\hat{y}]}{E[\hat{y}]^2} - \frac{2Cov[\hat{x},\hat{y}]}{E[\hat{x}]E[\hat{y}]}} \approx \frac{\hat{x}}{\hat{y}}\sqrt{\frac{SE[\hat{x}]^2}{\hat{x}^2} + \frac{SE[\hat{y}]^2}{\hat{y}^2}} \tag{16}$$

assuming independence between $\hat{x}$ and $\hat{y}$ since they are statistics of independent sampling distributions (independent samples of males and females). The standard errors of the genetic and environmental variance were estimated using the law of total variance for a product of two random variables. For $\hat{a}$ and $\hat{b}$, unbiased estimators of the two parameters a and b, respectively, we get

$$SE[\hat{a}\hat{b}] = \sqrt{SE[\hat{a}]^2 SE[\hat{b}]^2 + E[\hat{a}]^2 SE[\hat{b}]^2 + E[\hat{b}]^2 SE[\hat{a}]^2}.$$

Plugging in the point estimate $\hat{a}$ for $E[\hat{a}] = a$ and the point estimate $\hat{b}$ for $E[\hat{b}] = b$,

$$\widehat{SE}[\hat{a}\hat{b}] = \sqrt{SE[\hat{a}]^2 SE[\hat{b}]^2 + \hat{a}^2 SE[\hat{b}]^2 + \hat{b}^2 SE[\hat{a}]^2}. \tag{17}$$

In this case, a represents the phenotypic variance, and b represents either h_z^2 for estimation of genetic variance or $(1 - h_z^2)$ for estimation of environmental variance. Lastly, to obtain the standard error of the phenotypic variance, we used bootstrapping with 100 samples (with replacement) of estimates of the phenotypic variance in sex z,

$$\widehat{SE}[\hat{a}] = \sqrt{\frac{\sum_{i=1}^{100}\left(Var[Y]_i - \widehat{Var}[Y_z]_i\right)^2}{100 - 1}}$$

Finally, for each trait, we estimated $\tilde{\mathrm{Z}}$, the ratio of the two male-female ratios (environmental and genetic, y and x axes in **Fig. 6**, respectively), and its standard error, $SE[\tilde{\mathrm{Z}}]$, using the same method as in **Eq. 16**. Under the null hypothesis of equal environmental and genetic amplification (**Eq. 8**),

$$H_0: Z = 0, \tag{18}$$

where

$$\mathrm{Z} = \frac{\tilde{\mathrm{Z}} - 1}{SE[\tilde{\mathrm{Z}}]}.$$

In **Fig. 6,** we approximated 90% confidence intervals on Z by treating it as a Z score, i.e., further treating Z as a Standard Normal.

A Model of Sexually-Antagonistic Selection. We developed a model relating sex differences in additive effects on a trait at a biallelic locus ($\beta_m$and β_f) and divergence in allele frequencies. Our model resembles that of Cheng and Kirkpatrick[26] who developed a similar model relating allele-frequency differences and sex bias in gene expression. In short, we modelled sexually antagonistic, post-conception viability selection on a focal complex trait. We assumed allele frequencies in adult males, p_m, and adult females, p_f, are at equilibrium, i.e. do not change in consecutive generations. Under these conditions, we derive the relationship

$$F_{ST} \approx AV_{GxSex},$$

where F_{ST}[68] is the fixation index with respect to the male and female subpopulations, i.e., the proportion of heterozygosity in the population that is due to allelic divergence between the sexes. V_{GxSex} is defined as

$$V_{GxSex} := 2p(1 - p)\left(\beta_m - \beta_f\right)^2, \tag{19}$$

where p is the allele frequency in zygotes. A is a parameter that, importantly, is shared across all variants affecting the trait and can be thought of as the intensity of sexually-antagonistic selection acting on genetic variation for the trait in question.

In our model, allele frequencies at the autosomal locus are assumed to be equal in males and female zygotes. F_{ST} at adulthood takes the form

$$F_{ST} := \frac{Var_z[p_z]}{\bar{p}(1-\bar{p})} = \frac{E[p_z^2] - \bar{p}^2}{\bar{p}(1-\bar{p})} = \frac{p_m^2 + p_f^2 - \left(\frac{p_m + p_f}{2}\right)^2}{\bar{p}(1-\bar{p})} = \frac{\left(p_m - p_f\right)^2}{4\bar{p}(1-\bar{p})}, \tag{20}$$

where

$$\bar{p} = \frac{p_m + p_f}{2}.$$

If we further assume a near-1:1 sex ratio such that $\bar{p} \approx p$,

$$F_{ST} \approx \frac{\left(p_m - p_f\right)^2}{4p(1-p)}. \tag{21}$$

Sexually-antagonistic selection acting on viability will cause divergence in allele frequencies between adult males and females. We write the relative viabilities of the homozygote for the reference allele, the heterozygote and the homozygote for the effect allele as $1 :: 1 + d_z S_z :: 1 + S_z$ for each sex $z \in \{m, f\}$. The selection coefficient S_z and dominance coefficient d_z can be frequency-dependent, in which case these coefficients take their values at equilibrium. We can write the additive selection coefficient of the effect allele as

$$s_z = [p + (1-2p)d_z]S_z. \tag{22}$$

Assuming that zygotes are at Hardy-Weinberg equilibrium, the allele frequency in each sex at adulthood is

$$p_z \approx p + p(1-p)s_z, \tag{23}$$

where we neglected terms of order s_z^2 [85]. Plugging **Eq. 23** into **Eq. 21**, the divergence between males and females post-selection is

$$F_{ST} \approx \frac{1}{4}p(1-p)\left(s_m - s_f\right)^2. \tag{24}$$

We model the strength of viability selection acting on males and females as linear with the additive effect on a focal trait in each sex,

$$s_z = a_z \beta_z, \tag{25}$$

and recalling the simplifying assumption that allele frequencies are at equilibrium under sexually-antagonistic viability selection at the locus, such that selection favoring an allele in one sex is balanced by selection against that allele in the other sex,

$$s_f = -s_m. \tag{26}$$

If $\beta_m = \beta_f$, then **Eq. 24** simplifies to

$$F_{ST} \approx p(1-p)\left(a_f\beta_f\right)^2 = \frac{a_f^2}{2}V_G. \tag{27}$$

where

$$V_G = 2p(1-p){\beta_f}^2. \tag{28}$$

is the additive genetic variance. However, when β_m does not strictly equal β_f, **Eq. 25, 26** together imply

$$\beta_m + \beta_f = \frac{\beta_m + \beta_f}{\beta_m - \beta_f}(\beta_m - \beta_f) = \frac{\frac{s_m}{a_m} - \frac{s_m}{a_f}}{\frac{s_m}{a_m} + \frac{s_m}{a_f}}(\beta_m - \beta_f) = \frac{a_f - a_m}{a_f + a_m}(\beta_m - \beta_f). \tag{29}$$

Finally, using **Eq. 25**,

$$s_m - s_f = a_m\beta_m - a_f\beta_f = \frac{1}{2}[\left(a_m + a_f\right)\left(\beta_m - \beta_f\right) + \left(a_m - a_f\right)\left(\beta_m + \beta_f\right)], \tag{30}$$

which together with **Eq. 29** gives

$$s_m - s_f = \frac{1}{2}\left[\left(a_m + a_f\right) + \frac{\left(a_m - a_f\right)\left(a_f - a_m\right)}{a_f + a_m}\right]\left(\beta_m - \beta_f\right) = \frac{2a_m a_f}{a_m + a_f}\left(\beta_m - \beta_f\right). \tag{31}$$

We denote the heritability due to GxSex at the locus as $V_{GxSex} \coloneqq 2p(1-p)(\beta_m - \beta_f)^2$ and the parameter relating this contribution to the differentiation in allele frequencies as

$$A \coloneqq 2\left(\frac{a_m a_f}{a_m + a_f}\right)^2, \tag{32}$$

and plug **Eq. 31** into **Eq. 24**, we get

$$F_{ST} \approx AV_{GxSex}. \tag{33}$$

as given by **Eq. 3** in **Results**.

Estimating the potential for sexually-antagonistic selection on standing variation (A). For each trait and gnomAD subsample (**Supplementary Materials**), we estimated A using weighted

least squares linear regression of our estimate of F_{ST} ($\widehat{F_{ST}}$) to our estimate of V_{GxSex} ($\hat{V}_{GxSex}$), with weight w inversely proportional to our site-specific estimate of noise in the estimate of V_{GxSex}.

$$w = \frac{1}{\widehat{Var[\hat{V}_{GxSex}]}}. \tag{34}$$

To simplify the estimation of $Var[\hat{V}_{GxSex}]$, we treated the allele frequency p as perfectly estimated, and as independent of the allele frequency in the GWAS sample—as different data are used in the GWAS (UK Biobank) and in the allele frequency estimation (gnomAD). Under these assumptions,

$$\widehat{Var[\hat{V}_{GxSex}]} = Var[2p(1-p)\widehat{D^2}] = [2p(1-p)]^2 Var[(\hat{\beta}_m - \hat{\beta}_f)^2], \tag{35}$$

and thus the task at hand is estimating $Var[(\hat{\beta}_m - \hat{\beta}_f)^2]$. Using the law of total variance,

$$Var[(\hat{\beta}_m - \hat{\beta}_f)^2] = Var_{\hat{\beta}_f}\left[E_{\hat{\beta}_m}\left[(\hat{\beta}_m - \hat{\beta}_f)^2 \middle| \hat{\beta}_f\right]\right] + E_{\hat{\beta}_f}\left[Var_{\hat{\beta}_m}\left[(\hat{\beta}_m - \hat{\beta}_f)^2 \middle| \hat{\beta}_f\right]\right]. \tag{36}$$

We begin with the argument of the first term,

$$E_{\hat{\beta}_m}\left[(\hat{\beta}_m - \hat{\beta}_f)^2 \middle| \hat{\beta}_f\right] = E_{\hat{\beta}_m}[\hat{\beta}_m^2 - 2\hat{\beta}_m\hat{\beta}_f + \hat{\beta}_f^2 | \hat{\beta}_f] = \mu_m^2 + \sigma_m^2 - 2\mu_m\hat{\beta}_f + \hat{\beta}_f^2, \tag{37}$$

where we denote

$$\mu_z = E[\hat{\beta}_z]; \tag{38}$$
$$\sigma_z^2 = Var[\hat{\beta}_z]$$

for each sex $z \in \{m, f\}$. Plugging **Eq. 37** into the first term of **Eq. 36**,

$$Var_{\hat{\beta}_f}\left[E_{\hat{\beta}_m}\left[(\hat{\beta}_m - \hat{\beta}_f)^2 \middle| \hat{\beta}_f\right]\right] = Var_{\hat{\beta}_f}[\mu_m^2 + \sigma_m^2] + Var_{\hat{\beta}_f}[\hat{\beta}_f^2 - 2\mu_m\hat{\beta}_f] = \tag{39}$$
$$0 + Var_{\hat{\beta}_f}[\hat{\beta}_f^2 - 2\mu_m\hat{\beta}_f] = Var_{\hat{\beta}_f}[\hat{\beta}_f^2] + 4Var_{\hat{\beta}_f}[\mu_m\hat{\beta}_f] - 4\mu_m Cov_{\hat{\beta}_f}[\hat{\beta}_f^2, \hat{\beta}_f],$$

where the first and second step follow from the fact that $\mu_m^2 + \sigma_m^2$ is a constant. We can take note of the fact that $\hat{\beta}_z$ is Normally distributed around β_z, and in particular that it has no skewness. Therefore,

$$Cov_{\hat{\beta}_z}[\hat{\beta}_z^2, \hat{\beta}_z] = E[\hat{\beta}_z^3] - E[\hat{\beta}_z]E[\hat{\beta}_z^2] = (\mu_z^3 + 3\mu_z\sigma_z^2 + \gamma_z\sigma_z^3) - \mu_z(\mu_z^2 + \sigma_z^2) = 2\mu_z\sigma_z^2, \tag{40}$$

where $\gamma_z = 0$ is the skewness of $\hat{\beta}_z$. We can also note that

$$Var_{\hat{\beta}_z}[\hat{\beta}_z^2] = Var_{\hat{\beta}_z}[(\sigma_z b_z + \mu_z)^2], \tag{41}$$

where we defined

$$b_z = \frac{\hat{\beta}_z - \mu_z}{\sigma_z},$$

and therefore b_z is a Standard Normal and therefore b_z^2 is Chi-squared with one degree of freedom. **Eq. 41** now gives

$$\begin{aligned} Var_{\widehat{\beta}_z}[\hat{\beta}_z^2] &= Var_{\widehat{\beta}_z}[\sigma_z^2 b_z^2 + 2\sigma_z\mu_z b_z] \\ &= Var_{\widehat{\beta}_z}[\sigma_z^2 b_z^2] + Var[2\sigma_z\mu_z b_z] + Cov[\sigma_z^2 b_z^2, 2\sigma_z\mu_z b_z] \\ &= Var[b_z^2]\sigma_z^4 + 4Var[b_z]\mu_z^2\sigma_z^2 + 0 = 2\sigma_z^4 + 4\mu_z^2\sigma_z^2. \end{aligned} \tag{42}$$

Plugging **Eq. 40** and **Eq. 42** into **Eq. 39**, we find

$$Var_{\widehat{\beta}_f}\left[E_{\widehat{\beta}_m}\left[\left(\hat{\beta}_m - \hat{\beta}_f\right)^2 \middle| \hat{\beta}_f\right]\right] = 2\sigma_f^4 + 4\mu_f^2\sigma_f^2 + 4\mu_m^2\sigma_f^2 - 8\mu_m\mu_f\sigma_f^2. \tag{43}$$

We now turn to the second term of **Eq. 36**. First,

$$\begin{aligned} Var_{\widehat{\beta}_m}\left[\left(\hat{\beta}_m - \hat{\beta}_f\right)^2 \middle| \hat{\beta}_f\right] &= Var\left[\hat{\beta}_m^2 + 2\hat{\beta}_m\hat{\beta}_f \middle| \hat{\beta}_f\right] \\ &= Var\left[\hat{\beta}_m^2\right] + 4\sigma_m^2\hat{\beta}_f^2 - 4\hat{\beta}_f Cov\left[\hat{\beta}_m, \hat{\beta}_m^2\right]. \end{aligned} \tag{44}$$

Eq. 40 and **42** again give us

$$Var_{\widehat{\beta}_m}\left[\left(\hat{\beta}_m - \hat{\beta}_f\right)^2 \middle| \hat{\beta}_f\right] = 2\sigma_m^4 + 4\mu_m^2\sigma_m^2 + 4\sigma_m^2\hat{\beta}_f^2 - 8\mu_m\sigma_m^2\hat{\beta}_f, \tag{45}$$

which then gives

$$E_{\widehat{\beta}_f}\left[Var_{\widehat{\beta}_m}\left[\left(\hat{\beta}_m - \hat{\beta}_f\right)^2 \middle| \hat{\beta}_f\right]\right] = 2\sigma_m^4 + 4\mu_m^2\sigma_m^2 + 4\sigma_m^2(\sigma_f^2 + \mu_f^2) - 8\mu_m\mu_f\sigma_m^2. \tag{46}$$

Plugging **Eq. 43** and **Eq. 46** into **Eq. 36**, we obtain

$$\begin{aligned} Var\left[\left(\hat{\beta}_m - \hat{\beta}_f\right)^2\right] &= \\ &= 2\left(\sigma_m^4 + \sigma_f^4\right) + 4\sigma_m^2\sigma_f^2 + 4\left(\mu_m^2\sigma_m^2 + \mu_f^2\sigma_f^2\right) + 4(\sigma_m^2\mu_f^2 + \sigma_f^2\mu_m^2) \\ &- 8\mu_m\mu_f(\sigma_m^2 + \sigma_f^2). \end{aligned} \tag{47}$$

Finally, we estimate μ_z with the GWAS-derived point estimate of the effect $\hat{\beta}_z$ and σ_z with its standard error, $\hat{\sigma}_z = [\hat{\beta}_z]$. Plugging back into **Eq. 35**, we obtain

$$\begin{aligned} Var\left[\widehat{\hat{V}_{GxSex}}\right] &= [2p(1-p)]^2[2\left(\hat{\sigma}_m^4 + \hat{\sigma}_f^4\right) + 4\hat{\sigma}_m^2\hat{\sigma}_f^2 + 4\left(\hat{\beta}_m^2\sigma_m^2 + \hat{\beta}_f^2\sigma_f^2\right) + 4(\hat{\sigma}_m^2\hat{\beta}_f^2 + \hat{\sigma}_f^2\hat{\beta}_m^2) \\ &- 8\hat{\beta}_m\hat{\beta}_f(\hat{\sigma}_m^2 + \hat{\sigma}_f^2)]. \end{aligned} \tag{48}$$

To perform this estimation of A on the GWAS and F_{st} data, we used paired v and V_{GxSex} points for all sites which passed all previous stages of filtering. Weights were set by **Eq. 34** and follow **Eq. 48** where $\hat{\beta}_m$ and $\hat{\beta}_f$ are the GWAS effect estimates as above, and $\hat{\sigma}_m$ and $\hat{\sigma}_f$ are the GWAS standard errors (SE) estimates for the effect size of each site per trait.

To minimize the possibility of LD between sites used in the analysis as much as possible, we used the approximately independent LD blocks in Europeans[81] as in Section "Mixture weights for covariance structure between male and female effects". Namely, we subdivided the genome into 1703 LD blocks such that we expect sites from different LD blocks will have little to no LD between them. We iterated over the 1703 blocks and identified all post-filtering sites which fell within each block; we then randomly sampled one of these sites within each LD block and used this sample of (up to) 1703 sites to perform the weighted linear regression of F_{ST} on $V_{G\times Sex}$. The slope and SE of this regression line was our estimate of A. We replicated this estimation process 1,000 times to generate 1,000 estimates of A. We then used the estimates of slope divided by the SE of these estimates to generate 1,000 Z-scores for A. Each point estimate Z-score presented in **Fig. 7D** is the mean of the 1,000 samples of one site per LD block. We also show 90% bootstrap confidence intervals in the figure. These intervals correspond to the 5th and 95th percentile of 1000 replicates. In each replicate, we sample with replacement *m* LD blocks from the *m* LD blocks which had at least one site within them post-filtering (**Supplementary Materials**); we then sample one site per resampled block. In the main text, we focus on the results performed this estimation for Ashkenazi Jewish, Finnish, and Non-Finnish European populations as the other ancestry group-stratified subsamples in gnomAD are further diverged from the UKB White British sample and therefore our GWAS estimates are expected to be less portable[71,86].

Code and Data Availability

Computer code used in this study can be found at:

https://github.com/harpak-lab/GxSex

Sex-specific GWAS summary statistics are available at:

https://www.harpaklab.com/data

References

1. Ge, T., Chen, C.-Y., Neale, B. M., Sabuncu, M. R. & Smoller, J. W. Phenome-wide heritability analysis of the UK Biobank. *PLoS Genet* **13**, e1006711 (2017).

2. Mostafavi, H. *et al.* Variable prediction accuracy of polygenic scores within an ancestry group. *Elife* **9**, (2020).

3. Young, A. I., Benonisdottir, S., Przeworski, M. & Kong, A. Deconstructing the sources of genotype-phenotype associations in humans. *Science* **365**, 1396–1400 (2019).

4. Patel, R. A. *et al.* Effect sizes of causal variants for gene expression and complex traits differ between populations. *bioRxiv* (2021) doi:10.1101/2021.12.06.471235.

5. Helgeland, Ø. *et al.* Characterization of the genetic architecture of infant and early childhood body mass index. *Nature Metabolism* **4**, 344–358 (2022).

6. Hudson, A. I. *et al.* Analysis of genotype-by-environment interactions in a maize mapping population. *G3 (Bethesda)* **12**, (2022).

7. des Marais, D. L., Hernandez, K. M. & Juenger, T. E. Genotype-by-environment interaction and plasticity: Exploring genomic responses of plants to the abiotic environment. *Annual Review of Ecology, Evolution, and Systematics* vol. 44 5–29 (2013).

8. Long, Q. *et al.* Massive genomic variation and strong selection in Arabidopsis thaliana lines from Sweden. *Nat Genet* **45**, 884–890 (2013).

9. MacQueen, A. H. *et al.* Mapping of genotype-by-environment interactions in phenology identifies two cues for flowering in switchgrass (Panicum virgatum). *bioRxiv* (2021).

10. Huang, W. *et al.* Context-dependent genetic architecture of Drosophila life span. *PLOS Biology* **18**, e3000645 (2020).

11. Huang, W., Carbone, M. A., Lyman, R. F., Anholt, R. R. H. & Mackay, T. F. C. Genotype by environment interaction for gene expression in Drosophila melanogaster. *Nature Communications* **11**, 5451 (2020).

12. Falconer, D. S. & Mackay, T. F. C. *Introduction to Quantitiative Genetics* . (Harlow, 1996).

13. Lynch, M. & Walsh, B. *Genetics and Analysis of Quantitative Traits*. (Sinauer Associates, 1998).

14. Duncan, L. *et al.* Analysis of polygenic risk score usage and performance in diverse human populations. *Nat Commun* **10**, 3328 (2019).

15. Mars, N. *et al.* Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. *Nat Med* **26**, 549–557 (2020).

16. Young, A. I., Wauthier, F. L. & Donnelly, P. Identifying loci affecting trait variability and detecting interactions in genome-wide association studies. *Nat Genet* **50**, 1608–1614 (2018).

17. Mills, M., Barban, N. & Tropf, F. C. *An introduction to Statistical Genetic Data Analysis*. (The MIT Press, 2020).

18. Nagpal, S., Tandon, R. & Gibson, G. Canalization of the Polygenic Risk for Common Diseases and Traits in the UK Biobank Cohort. *Molecular Biology and Evolution* (2022) doi:10.1093/molbev/msac053.

19. Wolak, M. E., Roff, D. A. & Fairbairn, D. J. Are we underestimating the genetic variances of dimorphic traits? *Ecology and Evolution* **5**, 590–597 (2015).

20. van Doorn, G. S. Intralocus sexual conflict. *Ann N Y Acad Sci* **1168**, 52–71 (2009).

21. Arnqvist, G. & Rowe, L. *Sexual Conflict*. (Princeton University Press, 2005). doi:10.1515/9781400850600.

22. Barson, N. J. *et al.* Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature* **528**, 405–408 (2015).

23. Kidwell, J. F., Clegg, M. T., Stewart, F. M. & Prout, T. Regions of stable equilibria for models of differential selection in the two sexes under random mating. *Genetics* **85**, 171–83 (1977).

24. Connallon, T., Cox, R. M. & Calsbeek, R. Fitness consequences of sex-specific selection. *Evolution* **64**, 1671–82 (2010).

25. Harrison, P. W. *et al.* Sexual selection drives evolution and rapid turnover of male gene expression. *Proceedings of the National Academy of Sciences* **112**, 4393–4398 (2015).

26. Cheng, C. & Kirkpatrick, M. Sex-Specific Selection and Sex-Biased Gene Expression in Humans and Flies. *PLOS Genetics* **12**, e1006170 (2016).

27. Schroderus, E. *et al.* Intra- and Intersexual Trade-Offs between Testosterone and Immune System: Implications for Sexual and Sexually Antagonistic Selection. *The American Naturalist* **176**, E90–E97 (2010).

28. Power, R. A. *et al.* Fecundity of Patients With Schizophrenia, Autism, Bipolar Disorder, Depression, Anorexia Nervosa, or Substance Abuse vs Their Unaffected Siblings. *JAMA Psychiatry* **70**, 22 (2013).

29. Mokkonen, M. & Crespi, B. J. Genomic conflicts and sexual antagonism in human health: insights from oxytocin and testosterone. *Evolutionary Applications* **8**, 307–325 (2015).

30. Harper, J. A., Janicke, T. & Morrow, E. H. Systematic review reveals multiple sexually antagonistic polymorphisms affecting human disease and complex traits. *Evolution (N Y)* **75**, 3087–3097 (2021).

31. Oliva, M. *et al.* The impact of sex on gene expression across human tissues. *Science (1979)* **369**, eaba3066 (2020).

32. Sinnott-Armstrong, N., Naqvi, S., Rivas, M. & Pritchard, J. K. Gwas of three molecular traits highlights core genes and pathways alongside a highly polygenic background. *Elife* **10**, 1–35 (2021).

33. Carole Hooven. *T: The Story of Testosterone, the Hormone that Dominates and Divides Us*. (Henry Holt and Co., 2021).

34. Flynn, E. *et al.* Sex-specific genetic effects across biomarkers. *Eur J Hum Genet* **29**, 154–163 (2021).

35. Bernabeu, E. *et al.* Sex differences in genetic architecture in the UK Biobank. *Nature Genetics* **53**, 1283–1289 (2021).

36. Berg, J. J. *et al.* Reduced signal for polygenic adaptation of height in UK Biobank. *Elife* **8**, (2019).

37. Sohail, M. *et al.* Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *Elife* **8**, (2019).

38. Benonisdottir, S. & Kong, A. The Genetics of Participation: Method and Analysis. *bioRxiv* (2022) doi:10.1101/2022.02.11.480067.

39. Coop, G. & Przeworski, M. Lottery, luck, or legacy. A review of “The Genetic Lottery: Why DNA matters for social equality.” *Evolution (N Y)* **76**, 846–853 (2022).

40. Mills, M. C. & Tropf, F. C. Sociology, Genetics, and the Coming of Age of Sociogenomics. *Annual Review of Sociology* **46**, 553–581 (2020).

41. Coop, G. Reading tea leaves? Polygenic scores and differences in traits among groups. (2019).

42. Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am J Epidemiol* **186**, 1026–1034 (2017).

43. Pirastu, N. *et al.* Genetic analyses identify widespread sex-differential participation bias. *Nature Genetics* **53**, 663–671 (2021).

44. Kasimatis, K. R. *et al.* Evaluating human autosomal loci for sexually antagonistic viability selection in two large biobanks. *Genetics* **217**, (2021).

45. Muir, W., Nyquist, W. E. & Xu, S. Alternative partitioning of the genotype-by-environment interaction. *Theoretical and Applied Genetics* **84**, 193–200 (1992).

46. Robertson, A. The Sampling Variance of the Genetic Correlation Coefficient. *Biometrics* **15**, 469 (1959).

47. Falconer, D. S. The Problem of Environment and Selection. *The American Naturalist* **86**, 293–298 (1952).

48. Fry, J. D. The Mixed-Model Analysis of Variance Applied to Quantitative Genetics: Biological Meaning of the Parameters. *Evolution (N Y)* **46**, 540 (1992).

49. Yamada, Y. Genotype by Environment Interaction and Genetic Correlation of the same Trait under Different Environments. *The Japanese Journal of Genetics* **37**, 498–509 (1962).

50. Brown, B. C., Asian Genetic Epidemiology Network Type 2 Diabetes Consortium, Ye, C. J., Price, A. L. & Zaitlen, N. Transethnic Genetic-Correlation Estimates from Summary Statistics. *Am J Hum Genet* **99**, 76–88 (2016).

51. Galinsky, K. J. *et al.* Estimating cross-population genetic correlations of causal effect sizes. *Genet Epidemiol* **43**, 180–188 (2019).

52. Ni, G., Moser, G., Schizophrenia Working Group of the Psychiatric Genomics Consortium, Wray, N. R. & Lee, S. H. Estimation of Genetic Correlation via Linkage Disequilibrium Score Regression and Genomic Restricted Maximum Likelihood. *Am J Hum Genet* **102**, 1185–1194 (2018).

53. Shi, H., Mancuso, N., Spendlove, S. & Pasaniuc, B. Local Genetic Correlation Gives Insights into the Shared Genetic Architecture of Complex Traits. *The American Journal of Human Genetics* **101**, 737–751 (2017).

54. Bulik-Sullivan, B. K. *et al.* An atlas of genetic correlations across human diseases and traits. *Nature Genetics* **47**, 1236–1241 (2015).

55. DiMarco, M., Zhao, H., Boulicault, M. & Richardson, S. S. Why “sex as a biological variable” conflicts with precision medicine initiatives. *Cell Reports Medicine* **3**, 100550 (2022).

56. Lumish, H. S., O’Reilly, M. & Reilly, M. P. Sex Differences in Genomic Drivers of Adipose Distribution and Related Cardiometabolic Disorders: Opportunities for Precision Medicine. *Arteriosclerosis, thrombosis, and vascular biology* vol. 40 45–60 (2020).

57. Boyle, E. A., Li, Y. I. & Pritchard, J. K. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* **169**, 1177–1186 (2017).

58. Shi, H., Kichaev, G. & Pasaniuc, B. Contrasting the Genetic Architecture of 30 Complex Traits from Summary Association Data. *Am J Hum Genet* **99**, 139–53 (2016).

59. Sella, G. & Barton, N. H. Thinking About the Evolution of Complex Traits in the Era of Genome-Wide Association Studies. *Annual Review of Genomics and Human Genetics* **20**, 461–493 (2019).

60. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics* **47**, 291–295 (2015).

61. Urbut, S. M., Wang, G., Carbonetto, P. & Stephens, M. Flexible statistical methods for estimating and testing effects in genomic studies with multiple conditions. *Nature Genetics* **51**, 187–195 (2019).

62. Pasquali, R. Obesity and androgens: facts and perspectives. *Fertil Steril* **85**, 1319–40 (2006).

63. Domingue, B. W., Kanopka, K., Mallard, T. T., Trejo, S. & Tucker-Drob, E. M. Modeling Interaction and Dispersion Effects in the Analysis of Gene-by-Environment Interaction. *Behavior Genetics* **52**, 56–64 (2022).

64. Liu, D. *et al.* Skeletal muscle gene expression in response to resistance exercise: sex specific regulation. *BMC Genomics* **11**, 659 (2010).

65. Ryan, M. *The Genetics of Political Behavior*. (Routledge, 2020).

66. Ruzicka, F., Holman, L. & Connallon, T. Polygenic signals of sexually antagonistic selection in contemporary human genomes. *bioRxiv* (2021) doi:10.1101/2021.09.20.461171.

67. Ruzicka, F. *et al.* The search for sexually antagonistic genes: Practical insights from studies of local adaptation and statistical genomics. *Evolution Letters* **4**, 398–415 (2020).

68. Wright, S. The genetical structure of populations. *Ann Eugen* **15**, 323–54 (1951).

69. Weir, B. S. & Cockerham, C. C. Estimating F-Statistics for the Analysis of Population Structure. *Evolution (N Y)* **38**, 1358 (1984).

70. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).

71. Privé, F. *et al.* Portability of 245 polygenic scores when derived from the UK Biobank and applied to 9 ancestry groups from the same cohort. *The American Journal of Human Genetics* **109**, 12–23 (2022).

72. Arthur, R. S., Dannenberg, A. J. & Rohan, T. E. The association of prediagnostic circulating levels of cardiometabolic markers, testosterone and sex hormone-binding globulin with risk

of breast cancer among normal weight postmenopausal women in the <scp>UK</scp> Biobank. *International Journal of Cancer* **149**, 42–57 (2021).

73. Wang, J. *et al.* Sex-specific associations of circulating testosterone levels with all-cause and cause-specific mortality. *European Journal of Endocrinology* **184**, 723–732 (2021).

74. Khan, S. S. *et al.* Association of Body Mass Index With Lifetime Risk of Cardiovascular Disease and Compression of Morbidity. *JAMA Cardiology* **3**, 280 (2018).

75. Vazquez, G., Duval, S., Jacobs, D. R. & Silventoinen, K. Comparison of Body Mass Index, Waist Circumference, and Waist/Hip Ratio in Predicting Incident Diabetes: A Meta-Analysis. *Epidemiologic Reviews* **29**, 115–128 (2007).

76. Ning, Y., Wang, L. & Giovannucci, E. L. A quantitative analysis of body mass index and colorectal cancer: findings from 56 observational studies. *Obesity Reviews* **11**, 19–30 (2010).

77. Brown, C. D. *et al.* Body Mass Index and the Prevalence of Hypertension and Dyslipidemia. *Obesity Research* **8**, 605–619 (2000).

78. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).

79. Haworth, S. *et al.* Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. *Nature Communications* **10**, 333 (2019).

80. Stephens, M. False discovery rates: a new deal. *Biostatistics* kxw041 (2016) doi:10.1093/biostatistics/kxw041.

81. Berisa, T. & Pickrell, J. K. Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics* **32**, 283–5 (2016).

82. Sudlow, C. *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Medicine* **12**, e1001779 (2015).

83. Davies, N. M., Holmes, M. v & Davey Smith, G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* k601 (2018) doi:10.1136/bmj.k601.

84. Davey Smith, G. & Ebrahim, S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?*. *International Journal of Epidemiology* **32**, 1–22 (2003).

85. Gillespie, J. H. *Population Genetics: A Concise Guide*. (Johns Hopkins University Press, 2004).

86. Wang, Y. *et al.* Theoretical and empirical quantification of the accuracy of polygenic scores in ancestry divergent populations. *Nature Communications* **11**, 3865 (2020).