Local Genetic Sex Differences in Quantitative Traits

Emil Uffelmann[1], Christiaan de Leeuw[1], Danielle Posthuma[1,2]

Affiliations:

[1] Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Amsterdam Neuroscience, Vrije Universiteit Amsterdam

[2] Department of Child and Adolescent Psychiatry and Pediatric Psychology, Section Complex Trait Genetics, Amsterdam Neuroscience, Vrije Universiteit Medical Center, Amsterdam University Medical Center, Amsterdam, The Netherlands.

Address correspondence to Emil Uffelmann, Department of Complex Trait Genetics, Vrije Universiteit Amsterdam, De Boelelaan 1085, 1081 HV, Amsterdam, The Netherlands; E-mail: e.uffelmann@vu.nl.

Introductory Paragraph

Recently, large-scale sex-stratified Genome-Wide Association Studies (GWASs) have been conducted that compare heritabilities and estimate global genetic correlations between males and females[1–5]. These studies have identified a small number of traits that show genetic sex differences on a global level. However, such global analyses cannot elucidate the local architecture of genetic sex differences[6]. To address this gap and gain insight into local (dis-)similarities in genetic signal between males and females, we estimated local genetic correlations and tested whether these differ from one, compared local heritability estimates, and tested for equality of local genetic effects in 157 quantitative traits. While the vast majority of loci do not significantly differ between males and females, almost every trait we studied had at least one locus that did. We show that these loci can highlight trait-relevant biology and may partly explain observed phenotypic sex differences.

Main Text

Sex-stratified GWAS summary statistics were downloaded from the Nealelab (URLs). We analyzed 157 quantitative traits across 13 ICD/ICF domains with a combined maximum sample size of 360,564 (varies across traits, see Supplementary Table 1). Subjects were of British ancestry (see Methods for details).

The autosomal genome was divided into 2495 semi Linkage Disequilibrium (LD)-independent loci of approximately 1Mb in size (see Methods and ref [7]). We computed local heritability estimates (h^2_{local}) for all traits and found 146 loci across 47 traits to significantly differ between males and females at a Bonferroni-corrected significance threshold of $p < 0.05 / (2495 \times 157) = 1.28\text{e-}07$ (see Figure 1). 55% of these loci (80/156) have larger h^2_{local}s in females, which is not significantly different from 50% ($p = 0.28$). Across all traits, h^2_{local}s for loci that are significant in both males and females correlate strongly ($r = 0.98$, $p < 1.00\text{e-}300$). After the exclusion of loci with very large h^2_{local}s (i.e. lipoprotein A, direct bilirubin, and total bilirubin), the correlation estimate decreases but remains high ($r = 0.89$, $p < 1.00\text{e-}300$).

Approximately 62% of all significant h^2_{local}s are significant in only one of the sexes. Across all traits, we identify many more loci with significant h^2_{local}s in females compared to males (10614 loci across 153 traits vs. 8767 loci across 151 traits). In part this is due to the sample sizes of the sex-stratified GWASs being nearly always larger in females, increasing the power of h^2_{local} analyses in females relative to that in males. This is a consequence of the female sampling bias in the UK Biobank. Mostly blood biomarker traits show strong h^2 differences in magnitude for individual loci. For example, locus 2181 (17:7264459:8554763) for testosterone has a h^2_{local} of 4.3% (p < 1.00e-300) for males, but 0.06% (p = 3.79e-05) for females. Locus 963 for Rheumatoid factor (6:32454578:32539567) has a h^2_{local} of 1% (p = 3.19e-19) for females, but 0.02% (p = 0.4) for males. For Urate, four neighboring loci (4:8882617:11050119) have significant and large h^2_{local}s for both males and females, but the female estimates are 4-5 times larger (female range: 0.5% - 10%; male range: 0.1% - 2.7%). SLC2A9, a Urate solute carrier and main GWAS hit[8], is located within these loci, which suggests sex differences in the relative importance of this gene for Urate blood levels.

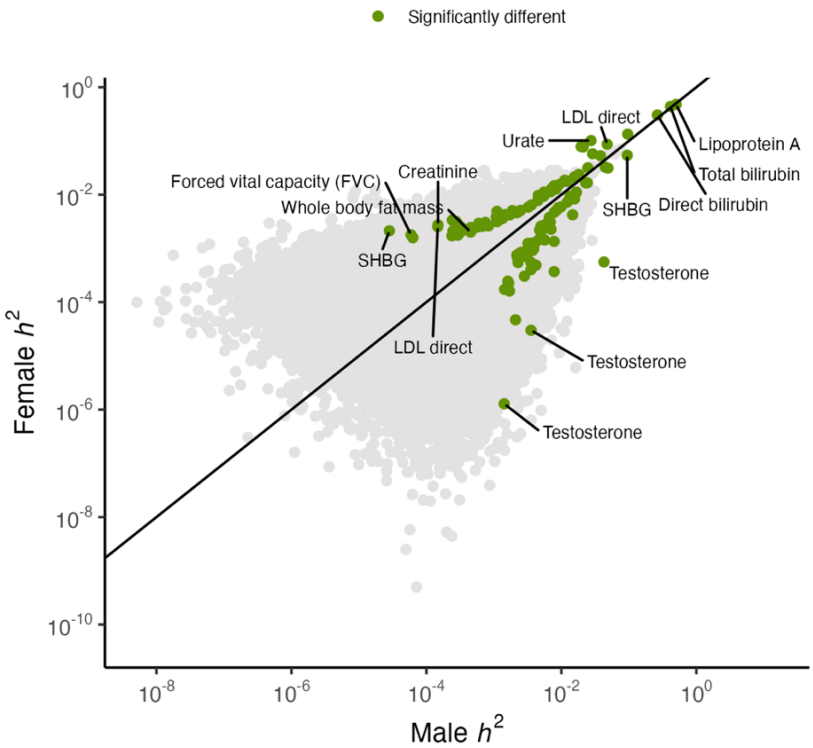

Figure 1. Local heritability estimates for 2495 loci across 157 traits in males and females. To test for significantly different heritability estimates, a Bonferroni-corrected threshold of p = 0.05 / (2495 x 157) = 1.28e-07 was used.

To test for local genetic sex differences, we computed local genetic correlations ($r_{g,\ local}$) with LAVA for every trait and locus with sufficient genetic signal (i.e. we used a h^2_{local} p-value threshold of $p < 1.00e\text{-}04$) for both males and females ($N_{loci} = 11259$) and tested if the $r_{g,\ local}$ is significantly different from one at a Bonferroni-corrected significance threshold of $p < 0.05 / 11259 = 4.44e\text{-}06$. We found 118 traits to have at least one locus where the $r_{g,\ local}$ is significantly different from one (see Figure 2 and Supplementary Figure 1). Again, we find blood biomarker traits (metabolic and immunological traits) to exhibit the largest number of differences between the sexes. IGF-1 and testosterone have the largest number of loci whose correlation is significantly different from one, namely 12 and 11, respectively (Supplementary Figure 1). For some traits, such as Urate, Total and Direct bilirubin and Lipoprotein A, we found loci where the $r_{g,\ local}$ was significantly different from one while being very close to one (see Figure 2). These are loci that explain much of the total h^2 in both males and females, thus having more precise $r_{g,\ local}$ estimates and higher statistical power to detect subtle deviations from one.

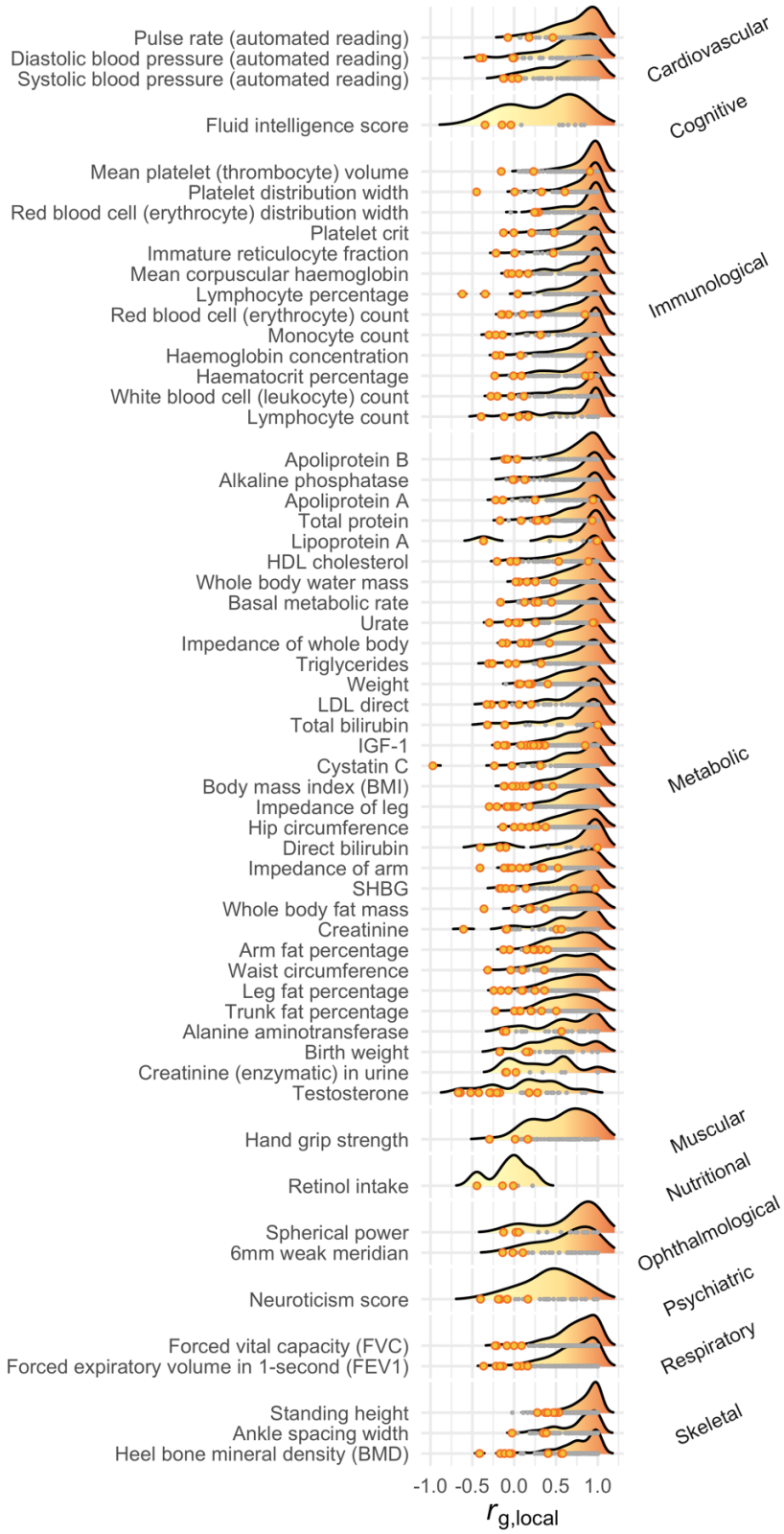

Figure 2. Density plots of $r_{g,\ local}$s. Depicted are all traits with at least three loci that are significantly different from one at a Bonferroni-corrected significance threshold of $p < 0.05 / 11259 = 4.44e\text{-}06$ (this minimum number of significant loci was chosen to aid visualization; see Supplementary Figure 1 for the number of significant loci for all traits). Grey dots are loci not significantly different from one and orange dots are loci significantly different from one.

In addition to local genetic correlations, we computed global genetic correlations ($r_{g,\ global}$) using LD Score Regression[9] and compared them to the weighted mean of $r_{g,\ local}$s across all loci for which $r_{g,\ local}$s could be computed using LAVA (see Figure 3; see Methods). 20.13% (30/149) of traits had LDSC-$r_{g,\ global}$s significantly different from one. However, most LDSC-$r_{g,\ global}$s are very close one, with the notable exception of testosterone. To ensure some reliability of LAVA mean $r_{g,\ local}$s, we only considered traits for which at least ten $r_{g,\ local}$s could be computed. Overall, there is good agreement between both methods (mean absolute difference = 0.12). For many traits with LDSC-$r_{g,\ global}$ close to one, we estimate some $r_{g,\ local}$s that are closer to zero and in some instances negative. For example, Body Mass Index (BMI) has an LDSC-$r_{g,\ global}$ of 0.93, but locus 1727 (11:122589225:123398882) has a $r_{g,\ local}$ of -0.12 (p = 4.03e-06) (see Figure 4). Similarly, testosterone has an LDSC-$r_{g,\ global}$ of zero but has several loci that are strongly positive and negative. For example, locus 2020 (14:93386329:94892240) has a $r_{g,\ local}$ of -0.67 and is significantly different from one (p < 1.00e-300), while locus 963 (6:32454578:32539567) has a $r_{g,\ local}$ of 0.85 which is significantly different from zero (p = 2.37e-08) but not from one (p = 0.17).

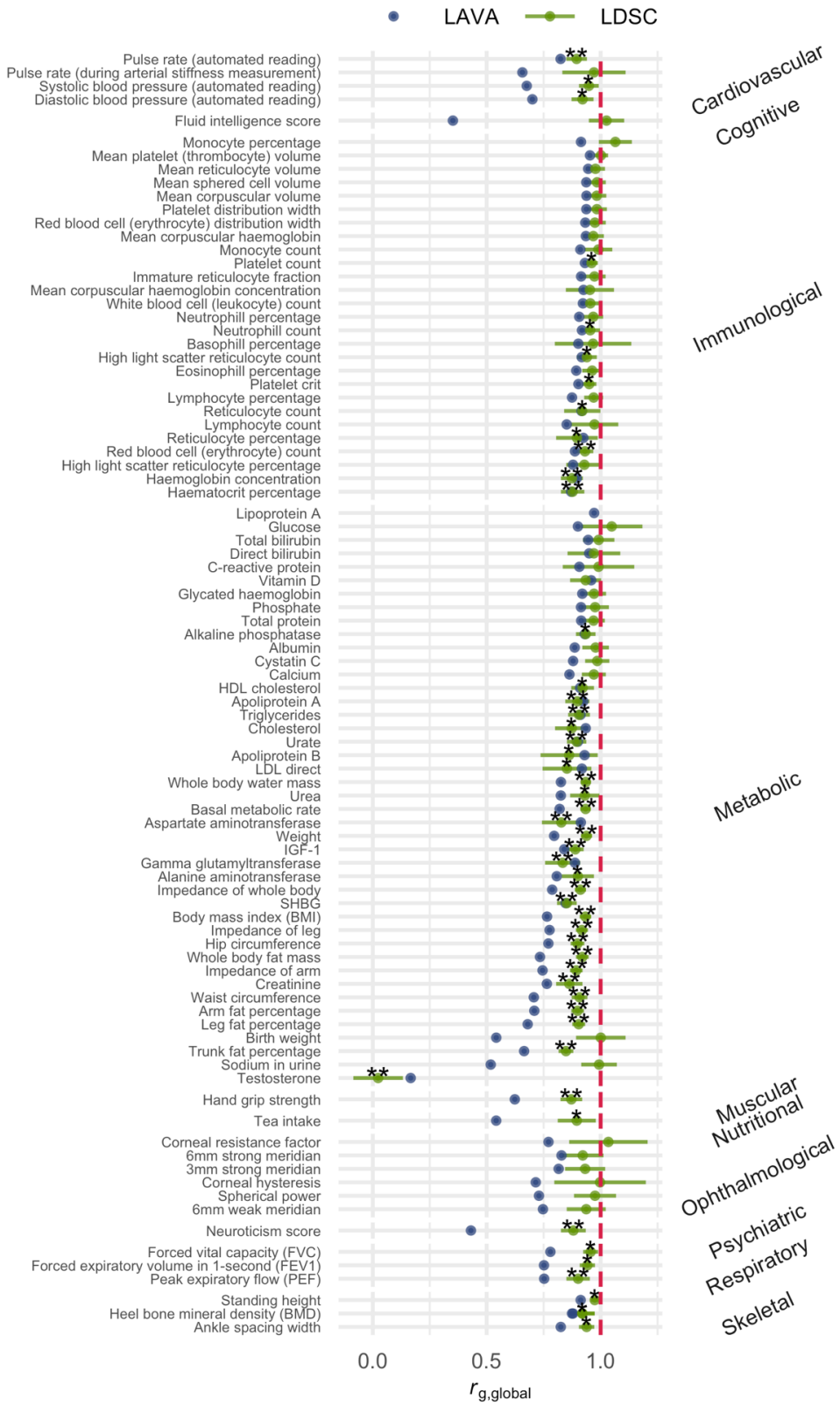

Figure 3. Global genetic correlations ($r_{g,\ global}$) estimated with LD Score Regression and the variance-weighted mean of local genetic correlations estimated for LAVA. The error bars for LDSC estimates depict standard errors. The depicted LAVA weighted means are not true estimates of $r_{g,\ global}$s and therefore have no associated p-values or standard errors.
* nominal significance at $p < 0.05$
** Bonferroni-corrected significance at 0.05 / 149 = 3.34e-04.

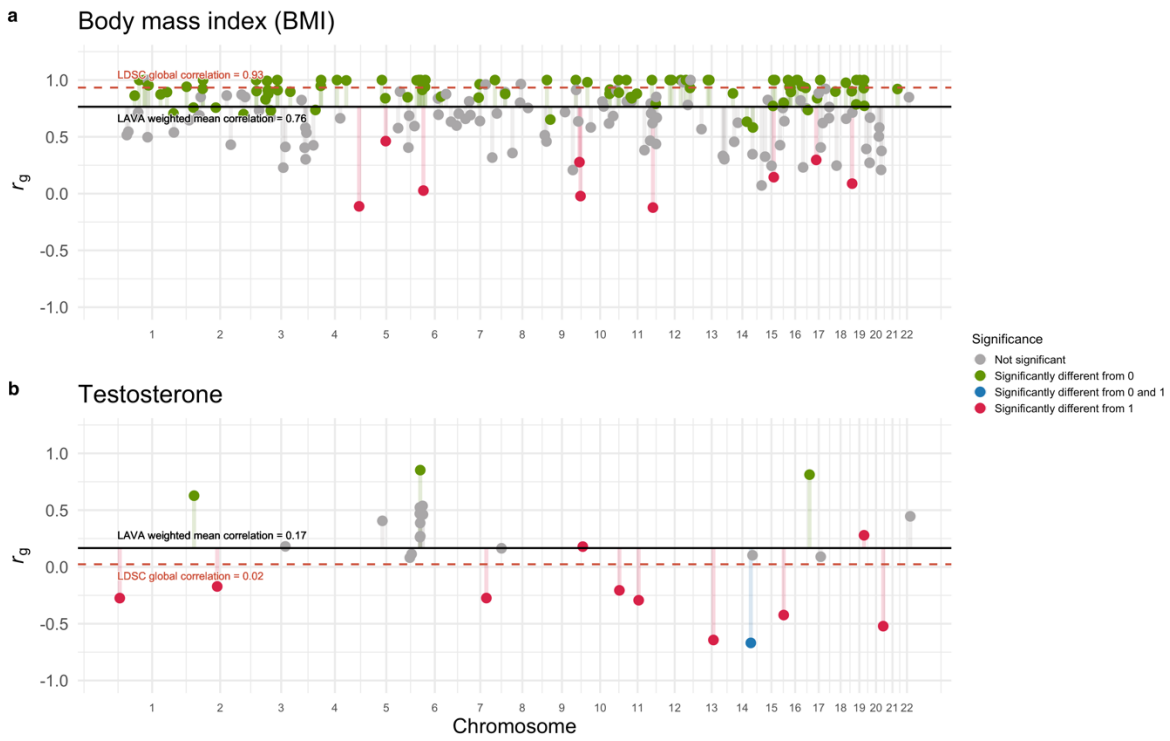

Figure 4. $r_{g,\ local}$ results for BMI (a) and testosterone (b). While the LDSC- $r_{g,\ global}$ estimate for BMI is close to one, many loci have $r_{g,\ local}$s that strongly and significantly diverge from one. In contrast, the testosterone LDSC- $r_{g,\ global}$ estimate is close to zero, while some loci are strongly negative and others positive. Similar plots for all other traits can be found in our Zenodo repository (see URLs).

A recent study suggested that differences in the absolute magnitude of genetic effects could strongly contribute to phenotypic sex differences[10]. As such, we extended LAVA to test for equality of genetic effects (see Methods for details). This is a more specific test for differences in genetic architecture because it not only tests for equality of the direction of genetic effects and their relative magnitudes (i.e. correlations), but also for equality of their absolute magnitude. For example, a locus in which SNP effects have the same direction and pattern of effects between the sexes but where the effects are twice as large in males would have a $r_{g,\ local}$ of one, even though the SNP effects all differ. Equality of genetic effects implies a $r_{g,\ local}$ of one and that the SNP effects are the same. As such, a test of equality of genetic effects is more specific than testing $r_{g,\ local}$ equal to one. The equality test is performed on two effect-size scales, the natural scale of the phenotype and a standardized scale where the effect sizes are scaled by the standard deviation of the phenotype. Here, we perform the equality test on both scales and report the maximum p-value, using a Bonferroni-

corrected significance threshold of 0.05 / (2495 x 157) = 1.28e-07. Out of 157 traits, 150 had at least one locus that is significantly different between males and females (see Figure 5 for traits with at least three significant loci). For example, locus 2181 (17:7264459:8554763) for testosterone has a $r_{g,\,local}$ not significantly different from 1 ($r_{g,\,local}$ = 0.81, p = 0.1), but the equality test is strongly significant (p < 1.00e-300), which can be explained by a much higher heritability estimate for males (h^2_{male} = 4.3%, h^2_{female} = 0.06%). As such, the equality test captures a larger number of loci that differ between males and females compared to testing only for $r_{g,\,local}$ that are significantly different from one.

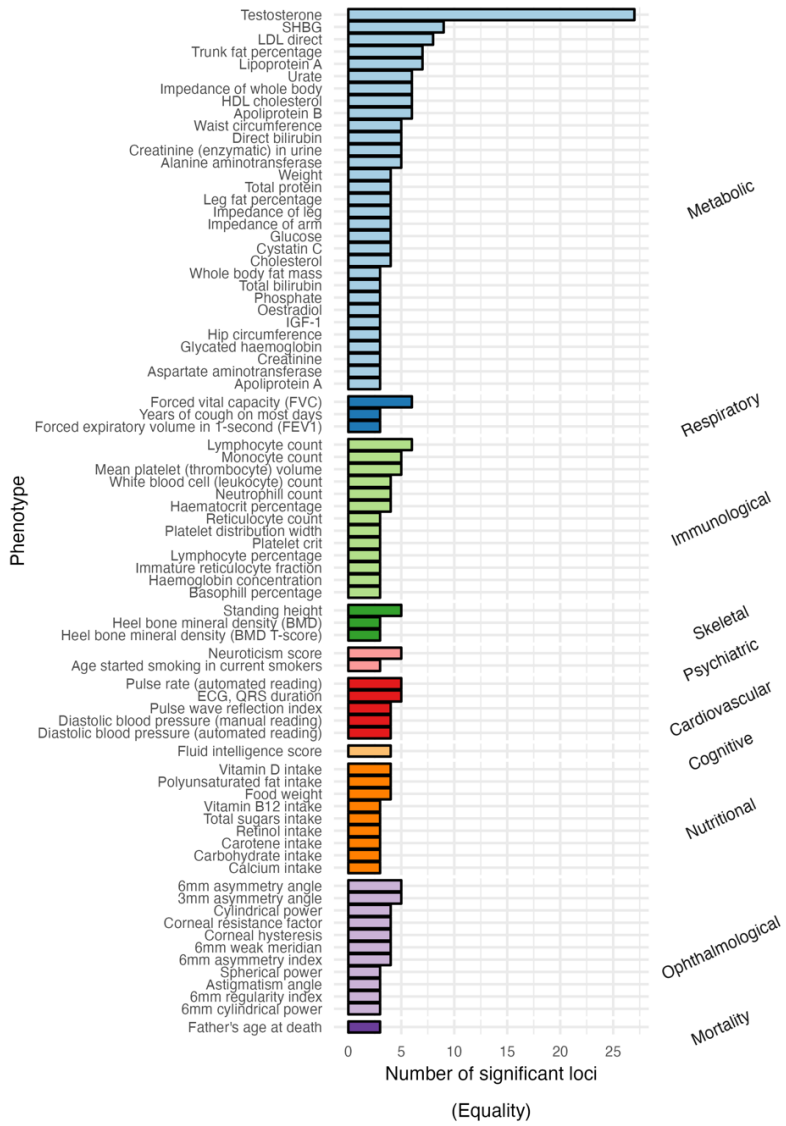

Figure 5. Bar plot of the number of significant loci in the equality test at a Bonferroni-corrected significance threshold of 0.05 / (2495 x 157) = 1.28e-07. Depicted are traits with at least three significant loci (this minimum number of significant loci was chosen to aid visualization; see Supplementary Figure 2 for a plot with all traits).

Next, we tested whether significant loci in the equality test can highlight known biological differences between the sexes. As a positive control, we highlight results for testosterone because its biology is well-understood and known to be markedly different between males and females[5,8]. We ran a MAGMA[11] gene analysis for testosterone in males and females separately, using a Bonferroni-corrected significance threshold for the number of tested genes 0.05 / 19288 = 2.60e-06 (see Figure 6). We highlighted all genes with known function in testosterone biology and that lie within loci that are significant in the equality test (referred to as inequality loci; Bonferroni-corrected for the number of loci tested: 0.05 / 2495 = 2.00e-05). As expected, genes in inequality loci that are associated with testosterone in the female sample are almost exclusively related to the KEGG pathway 'synthesis and metabolism' of testosterone (except for Sex Hormone Binding Globulin)[8]. In the male sample, genes in inequality loci highlight a wider range of functions, namely 'synthesis and metabolism', 'gonad development', 'HPG signaling', 'Serum homeostasis' and 'regulation of SHBG levels'. While the MAGMA results in inequality loci highlight genes in the 'synthesis and metabolism' pathway in both males and females, the specific sets of genes differ, again reflecting known sex differences[8]. For example, *CYP3A7* and *CYP3A5* are related to the hydroxylation of testosterone and estrone and are strongly associated with female but not male testosterone levels[8].

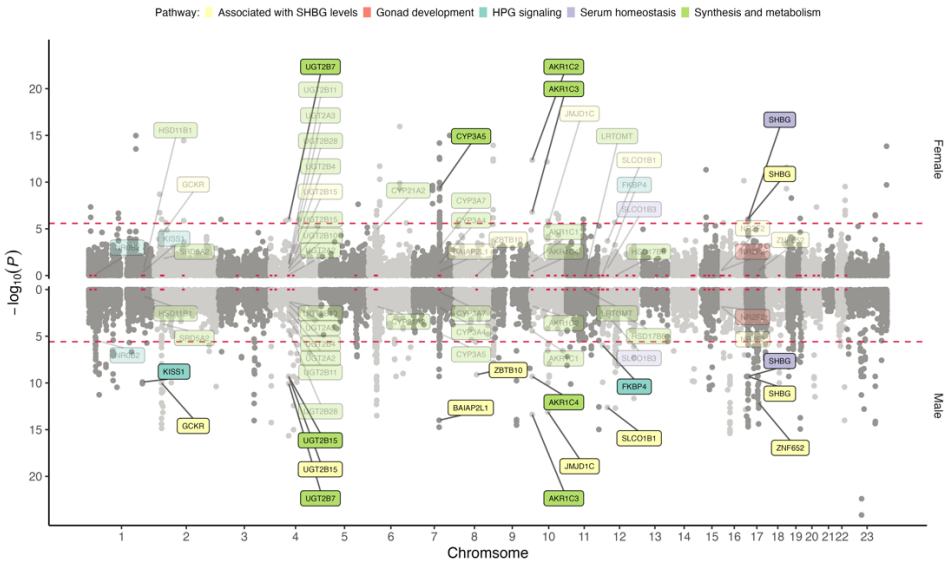

Figure 6. Miami plot of a MAGMA gene analysis for testosterone. Red bars at the base of the Manhattan plots mark significant loci from the equality test (inequality loci). Genes related to the biology of testosterone are highlighted only when located inside inequality loci. Transparent genes are not significant (Bonferroni-corrected for the number of genes tested, namely 19288).

For Cystatin C, a biomarker for kidney function, locus 2094 (15:94565522:95662425) has a near-perfect negative $r_{g,\ local}$ between males and females ($r_{g,\ local}$ = -0.97) which is significantly different from one (p < 1.00e-08) and 0 (p = 1.16e-06). This locus harbors the gene MCPT2, which was previously found to be associated with Diabetic Kidney Disease[12,13] and may explain some of the sex differences found in levels of Cystatin C.

For low-density lipoprotein levels (LDL-direct), locus 2351 (19:45040933:45893307) shows a pronounced sex difference. While the equality test is highly significant (p = 7.76e-72), the $r_{g,\ local}$ is not significantly different from one ($r_{g,\ local}$ = 0.99; p = 0.025), suggesting that the magnitude of effect sizes differ. The heritability estimate for this locus is almost twice as large for females (h^2_{female} = 9%; h^2_{male} = 5%). Interestingly, APOE is located in this locus, which has been shown to be differentially associated with LDL levels in males and females[14] and may explain why the Alzheimer's Disease risk conferred by APOE also differs by sex[15].

In conclusion, using publicly available sex-stratified GWAS summary statistics, we provide a broad overview of local genetic sex differences across a range of quantitative traits. Previous studies of genetic sex differences estimated global estimates of genetic correlations[1–5]. Using LAVA, we additionally computed $r_{g,\ local}$ and h^2_{local} estimates. Moreover, to move beyond a mere comparison of the direction and relative magnitude of effects, we developed the equality test to probe whether or not genetic effects are of equal size. For example, only focusing on genetic correlations would have missed the APOE locus that affects LDL-direct levels differently between the sexes. Most loci for most traits do not show significant differences, explaining why some previous genome-wide comparisons did not yield much insight into genetic sex differences[16]. Given our results, males and females are genetically much more alike than different. Nonetheless, almost every trait we studied had at least one locus that significantly differed between males and females. Our analyses for testosterone illustrate how these loci flag biologically-relevant regions that harbor key genes. As such, these loci may provide important clues as to why males and females sometimes differ in disease prevalences and mean trait values.

Methods

Quality control of summary statistics

Summary statistics for 157 traits (see Supplementary Table 1) were downloaded from the Nealelab (Data availability and URLs). Only summary statistics for untransformed (i.e. raw) and quantitative phenotypes were downloaded. Phenotypes that were deemed too similar were filtered out, such that only one phenotype remained (e.g. hand grip strength left vs. right). Additionally, phenotypes were removed that were suspected to have minimal genetic basis (e.g. 'age at recruitment'). Details on how the GWAS was conducted and which quality control filters were applied can be found on the Nealelab website (URLs). Briefly, sex-combined and sex-stratified GWASs were performed in 337,199 individuals of British ancestry using a linear regression model in Hail[17]. For the sex-stratified analyses, the first 20 principal components, age and age^2 were included as covariates. For initial quality control, all SNPs with imputation quality score below 0.8, minor allele frequency (MAF) below 0.1% (except for VEP[18] annotated SNPs) and Hardy-Weinberg-Equilibrium p-values below 1.00e-10 (except for VEP[18] annotated SNPs with MAFs below 0.1%) were removed. Considering that this database includes summary statistics that were generated on a large-scale without trait-specific quality control procedures, we additionally applied a strict minor allele frequency filter of 10% to ensure only well-imputed and reliable SNPs went into the analysis. This resulted in approximately five million SNPs for each trait. Manhattan and qq-plots, as well as LDSC Regression intercepts and Genomic Control statistics were additionally inspected to ensure high quality of summary statistics.

Local heritability, genetic correlation, and equality of genetic effects analysis

LAVA v0.0.7 scripts (URLs) were used to process all loci and compute univariate h^2_{local} and bivariate $r_{\text{g, local}}$ estimates. All LAVA analyses were performed on 2495 semi linkage disequilibrium independent blocks of approximately 1Mb in size (Data availability). Local genetic correlations were only estimated for loci with sufficient evidence of heritability in both males and females. To this end, we applied a h^2_{local} threshold of $p < 1.00\text{e-}04$. Because the summary statistics were based on individuals

of British ancestry, we used the European sample of phase 3 of 1000 Genomes[19] as the Linkage Disequilibrium reference sample (Data availability). All genomic coordinates refer to human genome build 37. Finally, it was tested if the h^2_{local}s differed, if $r_{\text{g, local}}$s significantly differed from one and if the genetic effects were significantly differed between the sexes.

LAVA model

A brief overview of the LAVA model is given here, a full description can be found in Werme et al. (2022)[7]. LAVA assumes a linear model of the form $Y_p = X\alpha_p + \varepsilon_p = W\delta_p + \varepsilon_p$ for each continous phenotype p, with standardized genotype matrix X of SNPs in the region being analyzed, standardized phenotype vector Y_p, and residual variance $\eta_p^2 = \text{var}(\varepsilon_p)$. The predictor matrix W contains standardized principal components obtained from X, and these are used to deal with the high degree of collinearity in X. In practice, estimates for δ_p and η_p^2 are obtained by reconstructing the linear regression model from GWAS summary statistics, and the LD structure of X from genotype reference data. Under the model, the estimates $\hat{\delta}_p$ are distributed as $\hat{\delta}_p \sim \text{MVN}(\delta_p, \sigma_p^2 I_K)$, with sampling variance $\sigma_p^2 = \frac{\eta_p^2}{N_p - 1}$, sample size N_p, and with K the number of principal components.

The local genetic component $G_p = X\alpha_p = W\delta_p$ is defined for each phenotype, which are combined into a matrix G of local genetic components for all phenotypes in the analysis. In the context of this study, the same phenotype for each sex is treated as two separate phenotypes. The quantity of interest is the local genetic covariance matrix $\Omega = \text{cov}(G)$, from which the local genetic components can be computed as $\rho_{pq} = \frac{\omega_{pq}}{\sqrt{\omega_p^2 \omega_q^2}}$, for each pair of phenotypes p and q. Since the phenotype vectors are assumed to be standardized, the local heritability for each phenotype p, which is the explained variance of the linear regression model, is equal to the variance of G_p, which are the diagonal elements of Ω. A method of moments estimator is used to obtain an estimate $\widehat{\Omega}$.

Testing equality of genetic effects

Strict homogeneity of the local genetic structure of the phenotype across sexes can be tested using a null hypothesis $H_0: \delta_M = \delta_F$, with δ_M and δ_F the genetic effect vectors for men and women for that phenotype respectively (this is also equivalent to testing $H_0: \alpha_M = \alpha_F$). Under this null model, the difference $\widehat{D} = \hat{\delta}_M - \hat{\delta}_F$ is distributed $\widehat{D} \sim \mathrm{MVN}(0, (\sigma_M^2 + \sigma_F^2)I_K)$, and it therefore follows that the test statistic $\frac{\widehat{D}^T\widehat{D}}{\sigma_M^2+\sigma_F^2}$ has a χ_K^2 distribution, which can be used to obtain a p-value.

Because these genetic effect vectors are defined on a standardized scale, the null hypothesis $H_0: \delta_M = \delta_F$ implies that the correlations of all of the SNPs in X with the phenotype (as well as the local heritabilities generally) are the same for each gender. However, homogeneity of genetic effects can also be defined on the natural scale of the phenotype instead, which implies that a change in genotypes results in the same amount of change (on the natural scale) in the phenotype. Writing $Y_p^* = Y_p S_p$, with Y_p^* the phenotype on its natural scale and S_p its standard deviation, the corresponding genetic effects on this scale would be $\delta_p^* = \delta_p S_p$, and equality on this scale can be tested using the null hypothesis $H_0: \delta_M^* = \delta_F^*$, which is equivalent to testing $H_0: S_M\delta_M = S_F\delta_F$. The corresponding difference vector $\widehat{D^*} = S_M\hat{\delta}_M - S_F\hat{\delta}_F$ is distributed $\widehat{D^*} \sim \mathrm{MVN}(0, (S_M^2\sigma_M^2 + S_F^2\sigma_F^2)I_K)$ under this null, and the test statistic $\frac{\widehat{D^*}^T\widehat{D^*}}{S_M^2\sigma_M^2+S_F^2\sigma_F^2}$ again has a χ_K^2 distribution.

If $S_M = S_F$ then these null hypotheses are equivalent, but if the phenotypic variances differ between sexes then the two null hypotheses cannot be true at the same time. As such, for this study we test both null hypotheses, and only declare a result significant if the p-value is below the significance threshold for both. This ensures that loci are not interpreted as showing heterogeneity of the genetic effects across sexes when this may be due only to the choice of scale.

Testing equality of local heritability

As the local heritability h_p^2 of a phenotype p equals the variance of G_p, it can be expressed as $h_p^2 = \mathrm{var}(G_p) = \frac{\delta_p^T W^T W \delta_p}{N_p - 1} = \delta_p^T \delta_p$. Moreover, $\eta_p^2 = 1 - h_p^2$. Since $\hat{\delta}_p \sim \mathrm{MVN}(\delta_p, \sigma_p^2 I_K)$, the statistic $T_p = \frac{\hat{\delta}_p^T\hat{\delta}_p}{\sigma_p^2}$ has a noncentral χ_K^2 distribution with

noncentrality parameter $\lambda_p = \frac{\delta_p^T \delta_p}{\sigma_p^2} = (N_p - 1)\frac{h_p^2}{\eta_p^2} = (N_p - 1)\frac{h_p^2}{1-h_p^2}$. Defining $C_p = \frac{h_p^2}{1-h_p^2}$, the expected value of T_p can be expressed as $E[T_p] = K + \lambda_p = K + (N_p - 1)C_p$.

Under the null hypothesis $H_0: h_M^2 = h_F^2 = h^2$ for a shared h^2, the expected value of the difference in test statistics is $E[T_M - T_F] = E[T_M] - E[T_F] = K + (N_M - 1)C - (K + (N_F - 1)C) = C(N_M - N_F)$, with $C = \frac{h^2}{1-h^2}$. To test this null hypothesis, we therefore define the test statistic $T_D = T_M - T_F$, and generate draws of T_D by separately sampling T_M and T_F from χ_K^2 distributions with noncentrality parameters of $(N_M - 1)C$ and $(N_F - 1)C$, and taking their difference. An empirical p-value is then computed as $2 \times \Pr(T_D^{(\text{draw})} < T_D^{(\text{obs})})$ if $T_D^{(\text{obs})} < C(N_M - N_F)$, and $2 \times \Pr(T_D^{(\text{draw})} > T_D^{(\text{obs})})$ otherwise. An adaptive sampling procedure was used for this, starting at an initial 10,000 draws of T_D, and increasing these up to a maximum of 100 million draws for lower p-values.

Note that the null hypothesis of equal local heritabilities, $H_0: h_M^2 = h_F^2$, is also implied by the null hypothesis $H_0: \delta_M = \delta_F$ of equality of standardized genetic effects, though the reverse is not true. By contrast, under the null hypothesis $H_0: S_M \delta_M = S_F \delta_F$ of equality of natural scale genetic effects, heritabilities can only be equal if $S_M = S_F$.

Testing perfect correlation of local genetic signal

To test the null hypothesis $H_0: \rho_{pq} = \rho_0$ of a perfect local genetic correlation between two phenotypes, we generalized the base LAVA model. To do so, we defined the test statistic $T_\rho = \widehat{\omega}_{pq} - \sqrt{\widehat{\omega}_p^2 \widehat{\omega}_q^2}\rho_0$, ie. the estimate of the local genetic covariance $\widehat{\omega}_{pq}$ minus its estimated expected value given a specific null value ρ_0 for the local genetic correlation. This reverts back to the $T_\rho = \widehat{\omega}_{pq}$ used in the original LAVA implementation when $\rho_0 = 0$. The full matrix estimate $\widehat{\Omega}$ has a noncentral Wishart distribution with K degrees of freedom, scale parameter $\Sigma = \begin{pmatrix} \sigma_p^2 & 0 \\ 0 & \sigma_q^2 \end{pmatrix}$ and noncentrality parameter $\Lambda = \Sigma^{-0.5} \Omega \, \Sigma^{-0.5}$ (see also Werme (2022)[7]). Filling in these parameters with the null value ρ_0 and the sample estimates, draws of T_ρ can therefore be generated by sampling values of $\widehat{\Omega}$ and computing the corresponding values of T_ρ. An empirical p-value can then be computed as $\Pr\left(\left|T_\rho^{(draw)}\right| > \left|T_\rho^{(obs)}\right|\right)$. An adaptive

sampling procedure was used for this, starting at an initial 1,000 draws of T_D, and increasing these up to a maximum of 100 million draws for lower p-values.

For this study, we specifically tested the null hypothesis $H_0: \rho_{MF} = 1$, to determine whether the sex-stratified local genetic signals for a phenotype were perfectly correlated. This null hypothesis is true if $\delta_M = c\delta_F$ for any arbitrary value $c > 0$, and is therefore implied by the null hypothesis of equal genetic effects at both the standardized ($c = 1$) as well as natural ($c = \frac{S_F}{S_M}$) scale, which shows that the $\rho_{MF} = 1$ test can be seen as a generalization of the equality of genetic effect tests. There is no logical relationship between this null hypothesis and the null hypothesis of equal local heritability however, since a local genetic correlation of one does not require local heritabilities to be equal, and equality of local heritabilities can exist at any value of the local genetic correlation.

Global genetic correlations

We used LDSC Regression v1.0.1[9] to compute $r_{g,\ global}$ between males and females for 157 traits to describe the average effect of pleiotropy across the whole genome. LD scores based on the European sample of 1000 Genomes were used (Data availability). Out of 157 traits, 8 had insufficient h^2 estimates for correlations to be computed. With LAVA, we computed the inverse-variance weighted mean of $r_{g,\ local}$s. Only loci that exceeded a h^2_{local} p-value threshold of 1.00e-04 were used to compute this mean.

Data availability

Sex-stratified GWAS summary statistics:

https://github.com/Nealelab/UK_Biobank_GWAS

LAVA locus definition file: https://github.com/josefin-werme/LAVA/blob/dbd3c4697fd3a139d6a2e8afdaf181eeac4a26cd/support_data/blocks_s2500_m25_f1_w200.GRCh37_hg19.locfile

1000 Genomes LD reference file for LAVA: https://ctg.cncr.nl/software/lava

Pre-computed LD scores can be downloaded at: s3://broad-alkesgroup-ukbb-ld/UKBB_LD/

Results from this study: https://doi.org/10.5281/zenodo.7896637

Code availability

Scripts for testing equality of genetic effects, equality of local heritability and perfect correlation of local genetic signal can be downloaded from:
https://doi.org/10.5281/zenodo.7896637

Acknowledgement

D.P. is supported by the Netherlands Organization for Scientific Research - Gravitation project 'BRAINSCAPES: A Roadmap from Neurogenetics to Neurobiology' (024.004.012) and D.P., E.U. and C.d.L. are supported by the European Research Council advanced grant 'From GWAS to Function' (ERC-2018-ADG 834057).

Author contributions

D.P. conceived of the project. E.U. performed all analyses. C.d.L. wrote the R code for the equality, local genetic correlations, and h^2-difference test and performed simulations to probe their statistical properties. E.U. wrote the paper and C.d.L. contributed to the Methods section. All authors discussed the results and commented on the paper.

Competing interests

C.d.L. was funded by Hoffman-La Roche until July 2022. The other authors declare no competing interests.

URLs

Nealelab UK Biobank GWAS (https://github.com/Nealelab/UK_Biobank_GWAS)

LAVA (https://github.com/josefin-werme/LAVA)

LD-Score Regression (https://github.com/bulik/ldsc)

MAGMA (https://ctg.cncr.nl/software/magma)

Zenodo repository (https://doi.org/10.5281/zenodo.7896637)

References

1. Bernabeu, E. *et al.* Sex differences in genetic architecture in the UK Biobank. *Nat. Genet.* **53**, 1283–1289 (2021).
2. Blokland, G. A. M. *et al.* Sex-Dependent Shared and Nonshared Genetic Architecture Across Mood and Psychotic Disorders. *Biol. Psychiatry* **0**, (2021).
3. Martin, J. *et al.* Examining Sex-Differentiated Genetic Effects Across Neuropsychiatric and Behavioral Traits. *Biol. Psychiatry* **89**, 1127–1137 (2021).
4. Trzaskowski, M. *et al.* Quantifying between-cohort and between-sex genetic heterogeneity in major depressive disorder. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **180**, 439–447 (2019).
5. Flynn, E. *et al.* Sex-specific genetic effects across biomarkers. *Eur. J. Hum. Genet.* **29**, 154–163 (2021).
6. van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H. & Wray, N. R. Genetic correlations of polygenic disease traits: from theory to practice. *Nat. Rev. Genet.* **20**, 567–581 (2019).
7. Werme, J., van der Sluis, S., Posthuma, D. & de Leeuw, C. A. An integrated framework for local genetic correlation analysis. *Nat. Genet.* **54**, 274–282 (2022).
8. 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**, e58615 (2021).
9. ReproGen Consortium *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
10. Zhu, C. *et al.* Amplification is the primary mode of gene-by-sex interaction in complex human traits. *Cell Genomics* 100297 (2023).

11. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Comput. Biol.* **11**, e1004219 (2015).

12. Salem, R. M. *et al.* Genome-Wide Association Study of Diabetic Kidney Disease Highlights Biology Involved in Glomerular Basement Membrane Collagen. *J. Am. Soc. Nephrol.* **30**, 2000 (2019).

13. Sandholm, N. *et al.* New Susceptibility Loci Associated with Kidney Disease in Type 1 Diabetes. *PLOS Genet.* **8**, e1002921 (2012).

14. Fu, J. *et al.* Effects of Sex on the Relationship Between Apolipoprotein E Gene and Serum Lipid Profiles in Alzheimer's Disease. *Front. Aging Neurosci.* **14**, (2022).

15. Altmann, A., Tian, L., Henderson, V. W., Greicius, M. D. & Investigators, A. D. N. I. Sex modifies the APOE-related risk of developing Alzheimer disease. *Ann. Neurol.* **75**, 563–573 (2014).

16. Stringer, S., Polderman, T. J. C. & Posthuma, D. Majority of human traits do not show evidence for sex-specific genetic and environmental effects. *Sci. Rep.* **7**, 8688 (2017).

17. Hail Team, H. T. Hail. (2023).

18. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).

19. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).