**The relationship between male-female allele frequency divergence at adulthood and GxSex**

We developed a model relating sex differences in additive effects on a trait at a biallelic locus (and ) and divergence in allele frequencies. Our model resembles that of Cheng and Kirkpatrick (CITEX CK16) who developed a similar model relating allele frequency differences and sex bias in gene expression. In short, we model sexually-antagonistic, post-conception viability selection on a focal complex trait. We assume allele frequencies in adult males, , and adult females, , are at equilibrium. Under these conditions, we derive the relationship

|  |  |  |
| --- | --- | --- |
|  |  |  |

where (CITEX Wright) 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; and is the contribution to phenotypic variance due to GxSex,

|  |  |  |
| --- | --- | --- |
|  |  | ( 2 ) |

where is the mean allele frequency at adulthood. is a parameter that can be thought of as the potential for sexually antagonistic selection acting on genetic variation for the trait in question.

Allele frequencies at the autosomal locus are assumed to be equal in males and female zygotes. If we assume equally-sized male and female subpopulations, at adulthood takes the form

|  |  |  |
| --- | --- | --- |
|  |  | ( 3 ) |

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 for each sex . The selection coefficient and dominance coefficient 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

|  |  |  |
| --- | --- | --- |
|  |  | ( 4 ) |

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

|  |  |  |
| --- | --- | --- |
|  |  | ( 5 ) |

where we neglected terms of order Plugging *eq. 5* into *eq. 3*, the divergence between males and females post-selection is

|  |  |  |
| --- | --- | --- |
|  |  | ( 6 ) |

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,

|  |  |  |
| --- | --- | --- |
|  |  | ( 6 ) |

and make 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,

|  |  |  |
| --- | --- | --- |
|  |  | ( 7 ) |

*Eqs. 6, 7* together imply

|  |  |  |
| --- | --- | --- |
|  |  | ( 8 ) |

Finally, using *eq. 6,*

|  |  |  |
| --- | --- | --- |
|  |  | ( 9 ) |

which together with *eq. 8* gives

|  |  |  |
| --- | --- | --- |
|  |  | ( 9 ) |

We denote the heritability due to GxSex at the locus as and the parameter relating this contribution to the differentiation in allele frequencies as

|  |  |  |
| --- | --- | --- |
|  |  |  |

and plug *eq. 9* into *eq. 6*, we get

|  |  |  |
| --- | --- | --- |
|  |  | ( 10 ) |

**Gathering and Filtering Allele Frequency Data**

We downloaded allele frequency data from the gnomAD dataset [1], a consortium of several ongoing global research projects. Specifically, we used gnomAD v3.1.2 which consists of 76,156 whole genome sequences mapped to the GRCh38 reference. This data is labeled by gnomAD and available for download with information about the sample demographics. The data is divided into ten ancestry groups: African/African American samples (abbreviated “afr” in the gnomAD files); Amish (“ami”); Latino/Admixed American (“amr”); Ashkenazi Jewish (“asj”); East Asian (“eas”); Finnish (“fin”); Non-Finnish European (“nfe”); Middle Eastern (“mid”); South Asian (“sas”); and samples not assigned to any population, designated Other (“oth”). The data is also separately divided into two chromosomal sexes, labeled XX and XY, based on coverage of X and Y chromosomes for the given individual. Aneuploid individuals (e.g., X or XXY) are not included in the dataset. For the purposes of this study, we refer to XX as Female and XY as Male. Total numbers of individuals sampled can be found on the gnomAD website’s help page (https://gnomad.broadinstitute.org/help).

The gnomAD data is available as direct download in variant call format (.vcf) from their online browser (https://gnomad.broadinstitute.org/downloads). We downloaded files for all autosomes and used the command-line program VCFTools [2] to parse the information. The original gnomAD .vcf files contain many more INFO fields than were necessary for the scope of our study, so we selected only the following specific columns: CHROM and POS, the chromosomal location; AN, the Allele Number, the total number of alleles sampled for the given locus; AC, the Allele Count, the number of alternative alleles found in the sample; and AF, the Allele Frequency, the frequency of alternative alleles in the sample (i.e., AC/AN). All of these fields (except for CHROM and POS) are given for each ancestry group and sex and are labeled as such in the .vcf fields e.g., AN\_afr\_XX.

We then passed this data through several filtering steps to remove loci which would not be useful to the analysis. During the gnomAD data acquisition process, we only chose sites for which the alternative alleles were single nucleotide polymorphisms (SNPs) as opposed to structural mutations such as insertions or deletions, and only kept those sites which were bi-allelic. After selecting the specific columns of interest but before performing any other filtering, our whole-genome dataset consisted of 2,288,867 sites. We filtered the data by first simply removing any loci with missing data (labeled as a “?” in the .vcf). This removed 3,698 sites, bringing the total number down from 2,288,867 to 2,285,169 sites. In an effort to avoid confounding results that could arise from population substructure, we split the data into the different ancestry groups labeled by gnomAD and worked with the data in each subpopulation separately from this point forward. We removed any loci where there were less than 1,000 alleles in the sample for that site in each specific ancestry group. The number of loci we removed at this step depended on the sample size of each group, as we removed fewer loci from large populations and removed more loci in small populations (S Table 1). Importantly, from this filtering step we completely removed the Amish and Middle Eastern populations because their low sample sizes meant that there were no loci which had more than 1,000 alleles in the sample. For this reason, we disregarded these populations for all future work.

**Estimating Male-Female FST**

We estimated Male-Female for every site remaining in our dataset following eq(3). We note that this is an upward biased estimator—especially for low levels of , as it only takes non-negative values. To get and , we could directly use the AF\_XX column for each subpopulation. To get , we use (AC\_XX + AC\_XY) / (AN\_XX + AN\_XY), such that is given as the number of alternate alleles in the sample divided by the total number of alleles in the sample.

We removed any loci which resulted in “NA”, indicating that 4pq = 0, or in other words there were no samples with the male or female tag with the alternative allele for that site. This step resulted in a significant reduction in locus number by one or two orders of magnitude (S Table 2).

**Integrating with GWAS Data and Calculating VGxSex**

Having completed these preliminary filtering steps, we integrated the gnomAD data with the GWAS data. We removed loci that were not found in both the gnomAD dataset and the UKB GWAS dataset. The number of loci removed through this step varied greatly across ancestry groups (S Table 3).

We used the point estimates and the standard errors of the sex-stratified GWAS (REFX Table SXX) for 27 physiological or physical traits. We also got an estimate of using eq(2), where and are the GWAS effect estimates and is the total alternate allele frequency as above.

Our final step of filtering involved filtering loci by GWAS p-value. To investigate the contribution of on a trait, we used four different p-value thresholds at 1e-3, 1e-5, 1e-8 and 1 (i.e., all SNPs; see REFXS Table 4 for the number of sites remaining for each p-value threshold). In the main text, we arbitrarily focus on the 1e-5 threshold reasoning that it strikes a good middle ground between sample size and noise. Results for other p-value thresholds are shown in REFX Figs. XXX.

**Estimating of the potential for sexually-antagonistic selection acting on variation in a trait ()**

For each trait and population, we estimated using weighted least squares linear regression of to , with weight inversely proportional to our site-specific estimate of noise in the estimate of .

|  |  |  |
| --- | --- | --- |
|  |  | ( 12 ) |

To simplify the estimation of we treat the allele frequency 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,

|  |  |  |
| --- | --- | --- |
|  |  | ( 13 ) |

and thus the task at hand is estimating . Using the law of total variance,

|  |  |  |
| --- | --- | --- |
|  |  | ( 14 ) |

We begin with the argument of the first term,

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | ( 15 ) | |
| where we denoted | |  |
| for each sex Plugging eq. 15 into the first term of eq. 14, | |  |
|  | |  |
|  |  | ( 16 ) | |

Where the first and second step follow from the fact that is a constant. We can take note of the fact that is Normally distributed around , and in particular that it has no skewness. Therefore,

|  |  |  |
| --- | --- | --- |
|  |  | ( 17 ) |

where is the skewness of . We can also note that

|  |  |  |
| --- | --- | --- |
|  |  | ( 18 ) |

where we defined

and therefore is a Standard Normal and therefore is Chi-squared with one degree of freedom. Eq. 18 now gives

|  |  |  |
| --- | --- | --- |
|  |  | ( 19 ) |

Plugging eq. 17 and eq. 19 into eq. 16, we find

|  |  |  |
| --- | --- | --- |
|  |  | ( 20 ) |

We now turn to the second term of eq. 14. First,

|  |  |  |
| --- | --- | --- |
|  |  | ( 21 ) |

Eq. 17 and 19 again give us

|  |  |  |
| --- | --- | --- |
|  |  | ( 22 ) |

which then gives

|  |  |  |
| --- | --- | --- |
|  |  | ( 23 ) |

Plugging eq. 20 and eq. 23 into eq. 14, we get

|  |  |  |
| --- | --- | --- |
|  |  | ( 24 ) |

Finally, we estimate with the GWAS-derived point estimate of the effect and with its standard error, . Plugging back into eq. 13, we get

|  |  |  |
| --- | --- | --- |
|  |  | ( 25 ) |

To perform this estimation of A on the GWAS and FST data, we used paired FST and VGxS points for all loci which passed all previous stages of filtering. Weights were set by eq(12) and follow eq(25) where and are the GWAS effect estimates as above, and and are the GWAS standard errors (SE) estimates for each locus’ effect size per trait.

To minimize the possibility of LD between sites used in the analysis as much as possible used the approximately independent LD blocks in Europeans [6] as in REFX Section “XXX”. 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 loci which fell within each block; we then randomly sampled one of these loci within each LD block and used this sample of (up to) 1703 sites to perform the weighted linear regression of on 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. The point estimate Z-score presented in REFX Fig. 7B is the mean of the 1,000 replicates, and the error bar is the middle 90% (i.e., the 5th to 95th quantile) of those replicates. 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 groups in gnomAd are more genetically diverged from the UKB White British sample and the GWAS estimates are expected to be less portable (CITEX Bjarni Shai Carmi et al. fine scale portability in 245 traits AJHG 2021; Wang Visscher Yang et al. Nature Genetics 2021), similar to the UK Biobank dataset used to generate the GWAS data.

|  |  |  |  |
| --- | --- | --- | --- |
| **Ancestry** | **# Loci Before Filtering by < 1,000 Alleles in Sample** | **# Loci After Filtering** | **# Loci Removed** |
| African/African American | 2,285,169 | 2,282,394 | 2,775 |
| Amish | 2,285,169 | 0 | 2,285,169 |
| Latino/American Admixed | 2,285,169 | 2,282,394 | 8,910 |
| Ashkenazi Jewish | 2,285,169 | 2,247,203 | 37,966 |
| East Asian | 2,285,169 | 2,255,778 | 29,391 |
| Finnish | 2,285,169 | 2,249,604 | 35,565 |
| Middle Eastern | 2,285,169 | 0 | 2,285,169 |
| Non-Finnish European | 2,285,169 | 2,283,952 | 1,217 |
| South Asian | 2,285,169 | 2,167,009 | 118,160 |
| Other | 2,285,169 | 2,080,336 | 204,833 |

**S Table 1. Results of filtering loci by removing any loci with less than 1,000 alleles in the sample.** We removed all loci with less than 1,000 alleles in the sample in order to narrow our research on loci with sufficiently large sample sizes. The number of sites removed during this step was highly dependent on the overall sample size for each ancestry. The Amish and Middle Eastern sub-populations, having the fewest samples overall, had no loci which passed through this round of filtering. All other ancestries retained a relatively high number of loci.

|  |  |  |  |
| --- | --- | --- | --- |
| **Ancestry** | **# Loci Before Filtering by NA FST Values** | **# Loci After Filtering** | **# Loci Removed** |
| African/African American | 2,282,394 | 924,946 | 1,357,448 |
| Latino/American Admixed | 2,282,394 | 519,496 | 1,756,763 |
| Ashkenazi Jewish | 2,247,203 | 161,271 | 2,085,932 |
| East Asian | 2,255,778 | 300,323 | 1,955,455 |
| Finnish | 2,249,604 | 207,162 | 2,042,442 |
| Non-Finnish European | 2,283,952 | 1,104,008 | 1,179,944 |
| South Asian | 2,167,009 | 330,518 | 1,836,491 |
| Other | 2,080,336 | 267,812 | 1,812,524 |

**S Table 2. Results of filtering loci by removing any loci with NA FST Values.** FST was calculated for all loci using the Allele Count and Allele Number values in the gnomAD data set. We then removed any loci which had an FST value of “NA” in any ancestry, as this would indicate that there were 0 samples with the alternative allele in at least one subpopulation.

|  |  |  |  |
| --- | --- | --- | --- |
| **Ancestry** | **# Loci Before Filtering by GWAS-gnomAD Overlap** | **# Loci After Filtering** | **# Loci Removed** |
| African/African American | 924,946 | 886,467 | 38,479 |
| Latino/American Admixed | 519,496 | 506,760 | 12,736 |
| Ashkenazi Jewish | 161,271 | 160,358 | 913 |
| East Asian | 300,323 | 295,801 | 4,522 |
| Finnish | 207,162 | 204,914 | 2,248 |
| Non-Finnish European | 1,104,008 | 1,051,015 | 52,993 |
| South Asian | 330,518 | 325,444 | 5,074 |
| Other | 267,812 | 264,682 | 3,130 |

**S Table 3. Results of filtering loci by keeping only overlapping loci between the gnomAD dataset and the GWAS dataset.** The gnomAD dataset contains SNP information from a consortium of several ongoing studies, while the GWAS data was gathered from studies using UK Biobank samples, so there was not complete overlap between the loci sampled. Therefore, any loci which did not occur in both the GWAS and gnomAD datasets were excluded.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ancestry** | **P-value Threshold** | **# Loci Before Filtering by NA FST Values** | **# Loci After Filtering** | **# Loci Removed** |
| Ashkenazi Jewish | 1e-3 | 160,358 | 2,749 | 157,609 |
| Ashkenazi Jewish | 1e-5 | 160,358 | 578 | 159,780 |
| Ashkenazi Jewish | 1e-8 | 160,358 | 298 | 160,060 |
| Finnish | 1e-3 | 204,914 | 3,975 | 200,939 |
| Finnish | 1e-5 | 204,914 | 852 | 204,062 |
| Finnish | 1e-8 | 204,914 | 373 | 204,541 |

**S Table 2. Representation of results of filtering by p-value threshold.** We established three thresholds of p-values to further restrict which loci were used for analysis by focusing on sites with strong correlation with the trait of interest. This was performed for all ancestries on all traits. A representative snapshot showing the number of SNPs removed at each threshold is shown here.

**References**

1. Karczewski, K.J., Francioli, L.C., Tiao, G. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581,** 434–443 (2020). <https://doi.org/10.1038/s41586-020-2308-7>
2. Petr Danecek, Adam Auton, Goncalo Abecasis, Cornelis A. Albers, Eric Banks, Mark A. DePristo, Robert E. Handsaker, Gerton Lunter, Gabor T. Marth, Stephen T. Sherry, Gilean McVean, Richard Durbin, 1000 Genomes Project Analysis Group, The variant call format and VCFtools, Bioinformatics, Volume 27, Issue 15, 1 August 2011, Pages 2156–2158, <https://doi.org/10.1093/bioinformatics/btr330>
3. R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
4. Cheng, C., &amp; Kirkpatrick, M. (n.d.). Sex-specific selection and sex-biased gene expression in humans and flies. PLOS Genetics. Retrieved March 18, 2022, from <https://doi.org/10.1371/journal.pgen.1006170>
5. Wright S. The genetical structure of populations. Ann Eugen. 1951 Mar;15(4):323-54. doi: 10.1111/j.1469-1809.1949.tb02451.x. PMID: 24540312.
6. Tomaz Berisa, Joseph K. Pickrell, Approximately independent linkage disequilibrium blocks in human populations, Bioinformatics, Volume 32, Issue 2, 15 January 2016, Pages 283–285, <https://doi.org/10.1093/bioinformatics/btv546>