**Text S8. Paragraph 1**

We downloaded allele frequency data from the gnomAD dataset15. Specifically, we used gnomAD v3.1.2, which consists of 76,156 whole genome sequences mapped to the GRCh38 reference. In order to allow this assembly to conform to the GRCh37 build used in the UK Biobank GWAS, we used the UCSC in-browser tool LiftOver (<https://genome.ucsc.edu/cgi-bin/hgLiftOver>) to identify and convert the GRCh38 sites of interest to their corresponding GRCh37 positions. Allele count summaries are available for the total sample, as well as stratified by sex chromosomeskaryotype and genetic ancestry groupings: 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 are designated Other (“oth”). Aneuploid individuals (e.g., X or XXY) are not included in the dataset. For the purposes of this study, we again 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>)

**Text S8. Paragraph 3**

Finally, we filtered out sites where difference in male and female allele counts may be partly or fully driven by the mismapping of autosomal reads to sex chromosomes or vice-versa17,18. We follow a similar approach to Kasimatis et al. to identify such sites18. In particular, for every SNP in the UK Biobank GWAS, including all imputed sites, we extract the 301bp sequence surrounding the SNP, with the SNP’s position at the center, from the GRCh37 genome assembly19. We further shorten the 301bp sequence into three 150bp-long subsequences with the SNP’s position at the center, start or end of the sequence. We then use Mega-BLAST through NCBI’s command-line BLAST tool20 to search for regions of high sequence homology to any of the three subsequences. If any of the three was found to have a 90% or greater sequence identity to a sequence on a sex chromosome, we filtered out the site. This filtering was performed agnostic of ancestry.