

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26



Supplementary Materials for

The pivotal role of the X-chromosome in the genetic architecture of the human brain  
**Authors:** Zhiwen Jiang, Patrick F. Sullivan, Tengfei Li, Bingxin Zhao, Xifeng Wang, Tianyou Luo, Shuai Huang, Peter Y. Guan, Jie Chen, Yue Yang, Jason L. Stein, Yun Li, Dajiang Liu, Lei Sun, and Hongtu Zhu\*

\*Corresponding author: htzhu@email.unc.edu

**The PDF file includes:**

- Materials and Methods
- Supplementary Text
- Figs. S1 to S21
- Captions of Tables S1 to S23
- References 115-142

**Other Supplementary Materials for this manuscript include the following:**

- Tables S1 to S23

## 1 **Materials and Methods**

### 2 Image acquisition and processing

3 The structural MRI (sMRI), diffusion MRI (dMRI), resting-state functional MRI (rfMRI), and task-  
4 evoked functional MRI (tfMRI) images were obtained from the UK Biobank  
5 (<http://www.ukbiobank.ac.uk/resources/>) with application 22783. These images have been subject to  
6 appropriate preprocessing by the UKB team, including reconstruction, gradient distortion correction, and  
7 quality control (QC). We briefly summarized the acquisition steps for each modality in Supplementary  
8 Text. Comprehensive information for the image acquisition and preprocessing is available at  
9 [https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain\\_mri.pdf](https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf). We included 1,045 brain imaging traits in  
10 the study, including 230 sMRI traits for cortical and subcortical structures, 635 diffusion tensor imaging  
11 (DTI) traits from dMRI for white matter (WM) microstructure; and 90 rfMRI (tfMRI) traits for intrinsic  
12 (extrinsic) brain functions. For each trait and continuous covariate variable (discussed later), we removed  
13 values greater than five times the median absolute deviation from the median value.

14  
15 We processed the sMRI locally using consistent procedures via advanced normalization tools (ANTs,  
16 <http://stnava.github.io/ANTs>). Specifically, we performed N4 bias correction, registration-based brain  
17 extraction, and a prior-based N4-Atropos 6 tissue segmentation (oasis template). The quality of  
18 segmentation was visually checked by experts. Subsequently, we conducted multi-atlas cortical  
19 parcellation based on the Desikan-Killiany-Tourville (DKT) cortical labeling protocol (115). After  
20 excluding three ROIs (5th ventricle, left lesion, right lesion) due to over 99% missing rate, 98 ROIs  
21 remained, with 62 of those capturing cortical regions. As a result, 101 traits for regional BVs were  
22 generated, including 98 regional and three global traits – gray matter volume (GMV), white matter  
23 volume (WMV), and total BV. The reproducibility, measured by the Pearson correlation between BV  
24 traits across two repeated scans from 2,944 retest subjects, was 0.92. For CT, we generated a global mean  
25 trait as well as regional traits from 62 cortical ROIs. Moreover, because the ANTs pipeline cannot  
26 generate surface area (SA) traits, we utilized 66 SA traits provided by the UKB (Category 193), which  
27 were derived based on the Desikan-Killiany (116) parcellation scheme using the FreeSurfer  
28 (<https://surfer.nmr.mgh.harvard.edu>) software. We provide details for the differences between DKT atlas  
29 and Desikan-Killiany atlas in Supplementary Text.

30  
31 DTI evaluated dMRI in a tensor model and characterized water molecular diffusions in all directions. Five  
32 metrics of DTI: Axial diffusivity (AD), fractional anisotropy (FA), mean diffusivity (MD), mode of  
33 anisotropy (MO), and radial diffusivity (RD) were applied to each voxel of the image. The five DTI  
34 metrics can reflect different patterns of water diffusion in WM tracts. Specifically, AD is the eigenvalue  
35 of the principal direction; FA is related to directionality; MD quantifies the magnitude of absolute  
36 directionality; MO is the third moment of a tensor; and RD is the average of the eigenvalues of secondary  
37 diffusion directions. Given a metric, a tract-mean trait was generated by taking the average of all voxels  
38 in a tract. As a result, we obtained 110 traits, including 105 tract-mean traits for all tract-metric pairs and  
39 5 overall-mean traits across all tracts ( $21 \times 5 + 5 = 110$ ). We also applied functional principal component  
40 analysis (FPCA) to the voxels in a tract and extracted the top five functional PCs for each tract-metric  
41 pair. The PCs preserved the strongest variation in voxel-level measurement and are expected to provide  
42 more microstructural details on axonal organization and integrity. We generated 525 ( $= 5 \times 5 \times 21$ )  
43 functional PC traits for all tracts and metrics. The tracts were labeled by the ENIGMA-DTI pipeline (117,  
44 118). Check Table S1 for the full names of the 21 tracts, and Supplementary Text for details of trait  
45 generation based on ENIGMA-DTI pipeline and FPCA.

46  
47 The shape, face, and shape-face activation contrasts were the major interests in UKB for tfMRI. We  
48 extracted the time series from the whole scan of the tfMRI data including blocks of both shape and face  
49 activations (UKB Category 106). When we calculated the functional connectivity measurements in

1 downstream analysis, we considered the shape and face contrasts together for capturing a more  
2 comprehensive view of the brain's functional dynamics.

3  
4 We applied parcellation-based methods with Glasser360 (38) atlas (also called HCP-MMP) to generate 90  
5 mean amplitude and functional connectivity traits each for rfMRI and tfMRI. The Glasser360 atlas is  
6 originally a surface-based parcellation for the cerebral cortex and has been transformed to a volumetric  
7 atlas to be compatible with the UKB volume-based data. We first projected the rfMRI and tfMRI data  
8 onto the Glasser360 atlas and generated 360×360 functional connectivity matrices. Then the 360  
9 functional areas were grouped into 12 functional networks including the primary visual, secondary visual,  
10 auditory, somatomotor, cingulo-opercular, default mode, dorsal attention, frontoparietal, language,  
11 posterior multimodal, ventra multimodal, and orbito-affective networks (119). As a result, 12 mean  
12 amplitude traits and 78 (=12+11×12/2) mean pairwise functional connectivity traits were generated for  
13 the 12 functional networks. Refer to the Supplementary Text for introduction of fMRI acquisition and  
14 detailed steps of the parcellation-based dimension reduction procedure.

#### 15 16 Discovery and replication data processing for association analyses

17 In our X-chromosome association analysis (XWAS), we utilized UK Biobank (UKB) phase 1-3 imaging  
18 data (up to February 2020) for the discovery phase, which included approximately 36,000 subjects. For  
19 replication, we employed phase 4 imaging data comprising an additional 3,100 subjects. We utilized the  
20 imputed genetic data (version 3). Details regarding genotyping and imputation are available in the UKB  
21 documentation. Genetic data processing in our study was conducted using PLINK2 (52) (v2.00a3LM,  
22 <https://www.cog-genomics.org/plink/2.0/>) separately for each imaging trait category.

23  
24 In the discovery phase, we focused on subjects with non-Hispanic white ancestries (Field 21000).  
25 Following UKB's quality control guidelines, we excluded subjects with excessive heterozygosity (Field  
26 ID 22027), discrepancies between reported and genetic gender (Field ID 22001), potential sex  
27 chromosome anomalies (Field ID 22019), and a genotype missing rate over 5% (--mind 0.05). We  
28 excluded SNPs with an imputation score less than 0.3 (Resource 1967), a minor allele frequency (MAF)  
29 less than 0.0003 (--maf 0.0003) for heritability analysis (more on this later), and further excluded SNPs  
30 with an imputation score less than 0.6, a MAF less than 0.005 for XWAS. Moreover, SNPs with a Hardy-  
31 Weinberg equilibrium test p-value less than  $10^{-6}$  (--hwe 0.000001) and multiallelic sites/indels (--snps-  
32 only just-acgt) were also excluded. To address potential relatedness among subjects, we employed GCTA  
33 (42) (v1.93.2 beta, <https://yanglab.westlake.edu.cn/software/gcta/#Overview>) to compute the genetic  
34 relationship matrix (GRM) using autosomes and excluded one of a pair of related subjects with the  
35 threshold 0.05 (--grm-cutoff 0.05), which removed approximately 1,800 subjects.

36  
37 The refined discovery dataset consisted of 33,591 subjects (15,939 males and 17,652 females) varying  
38 across imaging trait categories from 29,078 to 35,793. We analyzed 289,866 non-pseudo-autosomal  
39 region (NPR) SNPs and 11,508 pseudo-autosomal region (PAR) SNPs on the X-chromosome. For the  
40 replication phase, we included phase 4 non-Hispanic white subjects (UKBE, n = 4,181), phases 1-4 South  
41 Asian and Chinese subjects (UKBSAC, n = 462), and phases 1-4 African subjects (UKBA, n = 295). To  
42 enhance our sample size, we reintegrated white subjects previously excluded due to relatedness with the  
43 phase 4 white subjects (1,800 + 3,100), then conducted another round of relatedness pruning (--grm-  
44 cutoff 0.05). The same threshold was applied to Asian and African subjects for relatedness pruning. All  
45 other quality control steps for the replication phase were consistent with the discovery phase.

#### 46 47 Dosage compensation and heritability analysis

48 We employed the GREML analysis tool in GCTA (42) for heritability analysis on NPR SNPs on the X-  
49 chromosome. We explored three model assumptions for GRMs on the X-chromosome: full DC (--dc 1),  
50 no DC (--dc 0), and equal variance (without specifying --dc) (42). These models account for differences  
51 in genetic variance between sexes due to their unique coding schemes. Specifically, females are

1 consistently coded as (0, 1, 2), while males are coded as (0, 2) for full DC, (0, 1) for no DC, and  $(0, \sqrt{2})$   
2 for equal variance.

3  
4 To identify the most suitable DC model for each brain imaging trait, we produced GRMs under each  
5 assumption. As default, we presumed a consistent allele frequency distribution between causal and  
6 genotyped SNPs and adjusted for imperfect LD (--grm-adj 0). Both GRMs for autosomes and the X-  
7 chromosome were included in a single model to jointly estimate heritability, enhancing the total  
8 heritability captured compared to separate analyses (16).

9  
10 For sMRI and dMRI traits, we controlled for the indicator of release phase (1 if the subject was released  
11 in phase 3 and 0 if the subject was released in phase 1-2), assessment center (Field 54), genotype batch  
12 (Field 22000), top 40 genetic principal components (Field 22009), age at imaging, age-squared, sex, age-  
13 sex interaction, and age-squared-sex-interaction. Additional covariates were adjusted for non-global sMRI  
14 traits, including total BV, mean CT, and left/right total SA. For fMRI traits, following Alfaro-Almagro et  
15 al. (120), we additionally adjusted for head size (Field ID 25000), scan position X (Field ID 25756), scan  
16 position Y (Field ID 25757), scan position Z (Field ID 25758), scan table position (Field ID 25759), mean  
17 rfMRI head motion (Field ID 25741), and mean tfMRI head motion (Field ID 25742), as well as scan  
18 position X squared, scan position Z squared, mean rfMRI head motion squared and mean tfMRI head  
19 motion squared.

20  
21 We employed the likelihood-ratio-test (LRT) with the null distribution  $0.5\chi_1^2 + 0.5\chi_0^2$  to assess the X-  
22 chromosome's variance component in GREML. The Akaike information criterion (AIC) facilitated the  
23 comparison of model assumptions, with the smallest AIC indicating the optimal DC model.

#### 24 Enrichment analysis for heritability

25  
26 Following the approach of Mallard et al. (16), we characterized enrichment of X-linked heritability ( $h_X^2$ )  
27 by comparing it to the overall heritability ( $h_{all}^2$ ) relative to the proportion of NPR variants on the X-  
28 chromosome. Variant counts for each chromosome were obtained from the Genome Reference  
29 Consortium Human Build 37 (GRCh37 release 13), available at

30 [https://www.ncbi.nlm.nih.gov/assembly/GCF\\_000001405.13](https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.13). Specifically, the enrichment for  $h_X^2$  is

31 defined as  $\frac{h_X^2/h_{all}^2}{length_x/length_{all}}$ , where  $length_x$  and  $length_{all}$  represent the genomic lengths of the NPR on  
32 the X-chromosome and the entire genome, respectively. To determine whether  $h_X^2$  was enriched or

33 depleted for each trait, we applied a two-sided Z-test with the statistic  $\frac{observed\ h_X^2 - expected\ h_X^2}{standard\ error\ of\ h_X^2}$ , where

34  $expected\ h_X^2 = \frac{h_{all}^2 \times length_x}{length_{all}}$ , and the standard error of  $h_X^2$  was derived by using GREML. For significant

35 results, the Z-statistic greater than 0 indicates enrichment of  $h_X^2$ , while less than 0 suggests depletion.

#### 36 Sex-stratified heritability and phenotypic variance analysis

37  
38 We conducted a sex-stratified heritability analysis to directly examine the disparities in  $h_X^2$  between males  
39 and females. Utilizing GREML, we incorporated the same set of covariates used in our sex-agnostic  
40 analysis, except for sex and its interaction terms. We adopted the equal variance model for all traits,  
41 which assumes that males and females have equivalent heritability. The GREML output provides the  
42 phenotypic variance, denoted as  $V_p$ . We subsequently calculated the phenotypic variance ratio for each

43 trait as  $\frac{V_{p,male}}{V_{p,female}}$ .

#### 44 XWAS, sex-stratified XWAS, and meta-analysis

45  
46 We conducted association analysis for NPR and PAR SNPs using PLINK2 (v2.00a3LM). For NPR SNPs,  
47 we adopted the full DC model (--xchr-model 2) for traits that favor full DC or equal variance, and the no

1 DC model (--xchr-model 1) for others. The direction of effect size corresponded to the minor allele in the  
2 data. However, the minor allele for a particular variant might differ across imaging modalities. In the  
3 discovery phase, we adjusted for the same covariates as did in the heritability analysis. In the replication  
4 phase, the adjusted covariates included an indicator that the subject was included in phase 3 or phase 4,  
5 the assessment center, the top 10 genetic PCs, and all other imaging-related covariates. We applied the  
6 wild bootstrap approach to calibrate correlations between test statistics for imaging traits. The detailed  
7 procedure is provided in Supplementary Text. After multiple-comparison adjustment, SNPs achieving a  
8 genome-wide threshold of  $5 \times 10^{-8}$  were considered as significant. Separate XWAS for males and  
9 females in the discovery cohort were conducted, maintaining consistent covariate adjustments, excluding  
10 sex and its interaction terms.

11  
12 We employed FUMA (v1.4.1, <https://fuma.ctglab.nl/>) to group individual significant SNPs into genomic  
13 loci. The LD between SNPs were internally handled by FUMA using white subjects in 1000 Genome  
14 dataset. SNPs in LD > 0.6 were grouped under one independent significant SNP. Subsequently, those in  
15 LD > 0.1 were categorized under a single top SNP. LD blocks indexed by adjacent independent  
16 significant SNPs within 250 Kb of each other were merged into one genomic locus.

17  
18 We employed a meta-analysis approach to integrate separate XWAS results using METAL (version  
19 released on 05.05.2020, [https://genome.sph.umich.edu/wiki/METAL\\_Documentation](https://genome.sph.umich.edu/wiki/METAL_Documentation)). Inputs to the  
20 software included effect alleles, effect sizes, p-values, and sample sizes from individual analyses. The  
21 default procedure was adopted. Initially, p-values were transformed into Z-statistics. Then, the effect  
22 alleles and the direction of the effect sizes were adopted to align all studies to a consistent reference  
23 allele. An overarching Z-statistic was derived by taking a sample-size weighted sum of each individual  
24 statistic. The weighting was derived from the square root of sample size in each study. In our research, the  
25 meta-analysis combined the outcomes of sex-stratified XWAS, replication XWAS from UKBE,  
26 UKBSAC, and UKBA cohorts, as well as discovery and replication XWAS from UKBE participants.

#### 27 Gene-level analysis and biological annotation

28  
29 We conducted a series of gene-level analysis through FUMA (v1.4.1). We first performed a gene-based  
30 association analysis targeting 747 protein-coding genes on the X-chromosome using MAGMA (v1.08). In  
31 the analysis, SNPs were mapped to genes based on their physical locations in GRCh37, without including  
32 upstream and downstream regions of genes. Additionally, significant SNPs and other SNPs in LD > 0.6  
33 (including those not in XWAS but from the 1000 Genome dataset) were passed to functional consequence  
34 mapping (ANNOVAR (121), version 2017-01-11), eQTL mapping (with reference database:  
35 CommonMind Consortium (20) and GTEx v8 brain), and 3D chromatin interaction mapping (built-in  
36 chromatin interaction data: adult cortex, fetal cortex, dorsolateral prefrontal cortex, and hippocampus  
37 (122); annotate enhancer/promoter regions: E053-E082 (brain)).

38  
39 All significant genes prioritized in the functional mapping analyses were annotated using the DAVID  
40 Bioinformatics Database (94) (<https://david.ncifcrf.gov/home.jsp>) and SynGO (<https://syngoportal.org/>),  
41 with all parameters retained as default.

42  
43 We further employed H-MAGMA (123) to map significant SNPs to long non-coding RNAs using  
44 chromatin interaction profiles. An annotation file specific to the X-chromosome was created following  
45 established protocols (124) and applied MAGMA (v1.08) for H-MAGMA execution. Minor adjustments  
46 were made to the R code to focus exclusively on the X-chromosome.

#### 47 Summary data-based Mendelian randomization (SMR)

48  
49 The summary statistics of the XWAS and the summary statistics of eQTL analysis using CAGE whole-  
50 blood data provided by Sidorenko et al. (44) were utilized in the SMR analysis. The genetic data in  
51 XWAS was used as a reference for LD estimation. There were 1,639 probes for genes in the NPR of the

1 X-chromosome, and significant trait-gene pairs were identified by controlling FDR at 0.05 level. The  
2 significance indicates that the gene expression level may have a causal effect on the trait. Then we  
3 performed the HEterogeneity In Dependent Instrument (HEIDI) test to distinguish the pleiotropy of  
4 causal SNPs from linkage for the significant trait-gene pairs. A non-significant result at the nominal level  
5 ( $p$ -value  $> 0.05$ ) corresponds to no linkage effect, thereby supporting the hypothesis of pleiotropy.

### 6 Sex differences in genetic effect

7 Sex differences in genetic effects of NPR SNPs were tested by a two-sided z-test (95). The null  
8 hypothesis is  $H_0: \beta_m = \beta_f$ , where  $\beta_m$  and  $\beta_f$  are true per-allele genetic effects for males and females,  
9 respectively. We coded males as (0, 2) for a full DC trait, and as (0, 1) for a no DC trait while always  
10 coded females as (0, 1, 2). The test statistic is defined as  $z = \frac{b_m - b_f}{\sqrt{se_m^2 + se_f^2}}$ , regardless of DC assumption,

11 where  $b$  is a genetic effect estimate and  $se$  is the corresponding standard error. The raw SNP effect size  
12 and its standard error were divided by the trait's standard deviation derived from the phenotypic variance  
13 to correct for scale effects that could act as confounders in the study.

### 14 Differences in genetic profiles between subjects classified by phenotypic quantiles

15 We initially segregated the data by sex, considering significant confounding effects from sex-related  
16 phenotypic differences. For every trait, we calculated both the upper and lower 10th percentile scores.  
17 Using these scores, we selected subjects based on their quantile rankings. For instance, within a specific  
18 trait category (e.g., BV), if a subject's scores for over 25% of the traits surpassed the upper 10th percentile  
19 for those traits, then the subject was categorized into the "upper outlier" group. Conversely, if a subject's  
20 scores for more than 25% of the traits fell below the lower 10th percentile, then the subject was placed in  
21 the "lower outlier" group. We then extracted the genetic profiles of these subjects using PLINK2 (--geno-  
22 counts). Male genetic profiles were coded as 0 or 2, while female profiles were coded as 0, 1, or 2. We  
23 employed Fisher's exact test to compare genetic profiles between the "upper outlier" and "lower outlier"  
24 groups for each sex.

## 25 **Supplementary Text**

### 26 An introduction to imaging acquisition in the UK Biobank

27 The acquisition of T1-weighted structural scans was conducted using Siemens Skyra 3T scanners,  
28 employing a 3D MPRAGE sequence to achieve high-resolution imaging. These scans featured 1-mm  
29 sagittal slices and a 1×1-mm in-plane resolution, with in-plane acceleration (iPAT) set to 2, alongside  
30 prescan normalization to ensure consistent image quality across the dataset. To maintain the  
31 confidentiality and anonymity of study participants, the T1 images underwent a defacing process,  
32 removing any facial features that could potentially identify individuals.

33 The UKB dMRI data collection protocol was designed to capture high-resolution brain microstructures,  
34 utilizing a 2×2×2 mm spatial resolution with anterior-to-posterior phase-encoding and a multiband  
35 acceleration factor of three. The imaging employed two nonzero b-values of 1000 and 2000 s/mm<sup>2</sup>, each  
36 with 50 diffusion-encoding directions to enhance the detection of microstructural properties of the brain's  
37 WM. The diffusion preparation followed a standard Stejskal-Tanner pulse sequence, characterized by an  
38 echo time (TE) of 92 ms and a repetition time (TR) of 3600 ms. This choice aimed to improve the signal-  
39 to-noise ratio (SNR) by opting for a shorter TE compared to a twice-refocused (bipolar) sequence, albeit  
40 with a trade-off of increased susceptibility to eddy current distortions. For a more detailed description of  
41 the dMRI data acquisition protocol and its specifications, refer to Section 2.8 of the UKB brain imaging  
42 documentation.

43 The UKB rfMRI and tfMRI imaging data were collected using Siemens Skyra 3T scanners, following  
44 standardized protocols. For rfMRI, data was captured over 490 time points during a six-minute span, with  
45

1 a spatial resolution of  $2.4 \times 2.4 \times 2.4$  mm and dimensions of  $88 \times 88 \times 64$ . The TE was set at 39 ms, and the  
2 TR at 735 ms. Additionally, a single-band reference scan, matching the time series data's geometry but  
3 with enhanced between-tissue contrast, was acquired for head motion correction and modality alignment  
4 purposes (125).

5  
6 The tfMRI data utilized the Hariri faces/shapes emotion task, adapted from the Human Connectome  
7 Project (HCP) with modifications for duration and stimulus block repeats.  
8 (<https://biobank.ndph.ox.ac.uk/showcase/label.cgi?id=106>). Participants engaged with blocks of trials,  
9 identifying matching faces or shapes. This task was conducted over 332 time points within four minutes,  
10 maintaining the same spatial resolution, dimensions, and scanning parameters as the rfMRI sessions. The  
11 shape, face, and shape-face activation contrasts were the major interests in UKB.

12  
13 Preprocessing of both rfMRI and tfMRI datasets was conducted by the UKB brain imaging team (125).  
14 The full pipeline can be found in Section 3 of the UKB Brain Imaging Documentation. Overall, the  
15 pipeline encompassed image cleaning, image registration, and representative time series generation. The  
16 detailed preprocessing pipeline is documented by the UKB, with source codes accessible at  
17 [https://git.fmrib.ox.ac.uk/falmagro/UK\\_biobank\\_pipeline\\_v\\_1](https://git.fmrib.ox.ac.uk/falmagro/UK_biobank_pipeline_v_1).

#### 18 19 The differences between DKT atlas and Desikan-Killiany atlas in brain anatomy

20 In general, the DKT atlas is very similar to Desikan-Killiany atlas as the former parcellates each  
21 hemisphere into 31 regions while the latter parcellates it into 32. The 31 cortical labels in the DKT atlas  
22 are actually a modification of the Desikan-Killiany atlas to improve cortical labeling consistency and  
23 enhance FreeSurfer's cortical classification. Specifically, the DKT atlas excluded the regions "banksst"  
24 and "frontal pole" but included the "insula" region. All other ROIs exactly match in two atlases.

#### 25 26 The ENIGMA-DTI pipeline and functional principal component analysis for DTI

27 Our analysis employed the ENIGMA-DTI pipeline, a standardized set of procedures for processing DTI  
28 data across various datasets (<http://enigma.ini.usc.edu/protocols/dti-protocols/>). The process began with  
29 linear registration, aligning each FA image to the ENIGMA FA template, which is set at a  $1 \times 1 \times 1$  mm  
30 spatial resolution in the MNI-ICBM-152 standard space. We then applied nonlinear registration  
31 techniques to further refine the alignment of the FA images to the standard space. The registered FA  
32 images were then masked using the template brain mask to isolate brain tissue from non-brain elements  
33 within the images. Subsequently, we transferred the individual images of AD, MD, MO, and RD into the  
34 FA template space. The ENIGMA skeleton, representing the major WM pathways, was projected onto the  
35 registered images. Finally, we extracted tract-based statistics for each DTI metric, including tract-  
36 averaged means and functional PCs.

37  
38 To generate PCs for DTI metrics, we employed a "smooth-first-then-estimate" method, detailed at  
39 <https://doi.org/10.5281/zenodo.4549465>, using FA as an example. Initially, we extracted skeletonized FA  
40 images from 17,706 subjects in UKB phases 1 and 2, smoothing them with FSL software using a 2 mm  
41 Gaussian kernel. Smoothed voxel-wise FA values for each tract were then centralized by subtracting the  
42 dataset's group average. Subsequently, these values were reshaped into an  $n \times p$  matrix  $A$ , where  $n$  is the  
43 sample size and  $p$  is the size of the tract. We performed singular value decomposition  $A = UDV^T$ , where  $V$   
44 is a  $p \times p$  orthogonal matrix,  $U$  is an  $n \times n$  orthogonal matrix, and  $D$  is a diagonal matrix. The top 5 FA  
45 PCs were extracted, explaining over 70% of the FA variation on average.

#### 46 47 Parcellation-based network level feature extraction using the Glasser360 atlas

48 The Glasser360 atlas(38), a comprehensive parcellation developed from high-resolution functional and  
49 anatomical MRI data from Human Connectome Project (HCP) participants, offers a detailed  
50 representation of the cortex with 360 distinct cortical areas (180 per hemisphere). For our analysis, we  
51 utilized the volumetric version of the Glasser360 atlas, available at NeuroVault



1 (https://identifiers.org/neurovault.collection:1549), which had been transformed from its original surface-  
2 based format into a volumetric atlas via FreeSurfer. This conversion allows for direct application to the  
3 rfMRI and tfMRI datasets subject to appropriate preprocessing.

4  
5 The Glasser360 atlas was applied to these datasets to extract regional time series for all 360 cortical areas.  
6 From these time series, we derived amplitude traits based on standard deviation and full connectivity  
7 matrices through Gaussianized temporal correlations between all pairs of node time series. Further, by  
8 adopting the network definitions from Ji et al. (119), which split the 360 areas into 12 networks, we  
9 generated network-level traits. This segmentation resulted in 78 network-level groups, encompassing 12  
10 within-network and 66 between-network groupings.

#### 11 Adjusting for multiple hypothesis testing in XWAS by wild bootstrap

12 Suppose we have  $Q$  traits in total, and we want to adjust for multiple hypothesis testing considering the  
13 number of traits by using wild bootstrap approach, which indicates that we still apply the genome-wide  
14 threshold  $5 \times 10^{-8}$  for significance to the adjusted p-values. The specific steps are as follows,

- 15 1. For each trait  $Y_q$ , fit a null model  $Y_q = X_q \alpha_q + \varepsilon_q$ , where  $Y_q$  is a vector of trait;  $X_q$  is the  
16 covariate matrix for trait  $q$ . We get  $\hat{\alpha}_q$  and  $\hat{\varepsilon}_q$ .
- 17 2. Let  $v_i^{(b)} \sim N(0,1)$ , generate a sample  $b$  for each trait and each subject  $i$ :  $Y_{iq}^{(b)} = X'_{iq} \hat{\alpha}_q + v_i^{(b)} \hat{\varepsilon}_{iq}$ .  
18 Note that  $v_i^{(b)}$  are the same for all traits but different among subjects and bootstrap samples.
- 19 3. For a SNP  $k$ ,  $k = 1, \dots, N_g$ , do XWAS on all traits with models  $Y_q^{(b)} = X_q \alpha_q^{(b)} + Z_{qk} \beta_{qk}^{(b)} +$   
20  $\varepsilon_q^{(b)}$ ,  $q = 1, \dots, Q$ , where  $Z_{qk}$  are genotype vectors with different sample sizes across traits. Then  
21 we get  $F$  statistics for all  $Q$  traits  $(F_{1k}^{(b)}, \dots, F_{Qk}^{(b)})'$ . Take the max  $F$  statistic  $F_k^{(b)*}$  over all traits.
- 22 4. Repeat the step 2, 3 for  $B = 150$  samples, we get  $(F_k^{(1)*}, \dots, F_k^{(B)*})'$ . Then we approximate a  $\chi^2$   
23 distribution for SNP  $k$  by moment matching. Specifically, let  $k_1$ ,  $k_2$  and  $k_3$  denote the mean,  
24 variance, and skewness of  $(F_k^{(1)*}, \dots, F_k^{(B)*})'$ . We compute parameters  $a = k_3 / (4 \times k_2)$ ,  
25  $b = k_1 - 2k_2^2 / k_3$ , and  $d = 8k_2^3 / k_3^2$ . The adjusted p-value for SNP  $k$  on trait  $q$  can be  
26 approximated by  $1 - pchisq(\frac{F_{qk}^{obs} - b}{a}, d)$ , where  $pchisq$  is the c.d.f. of  $\chi^2$  distribution, and  $F_{qk}^{obs}$   
27 is the observed  $F$  statistic computed from the real data. Note the more bootstrap samples, the  
28 more accurate the approximation. If the raw p-values are extremely small, such as  $10^{-20}$ ,  $B$  may  
29 increase to 200 or 300.

30  
31 Since the computational burden of implementing the procedure for all SNPs was intractable, we split the  
32 SNPs into five groups based on MAF: [0.005, 0.01), [0.01, 0.05), [0.05, 0.10), [0.10, 0.25), and  
33 [0.25, 0.50], and randomly picked 10 SNPs in each group. Within a group, we combined  $(F_k^{(1)*}, \dots, F_k^{(B)*})$   
34 generated from the 10 SNPs to approximate the  $\chi^2$  distribution for that group. That implies we implicitly  
35 assumed that the distribution parameters were identical for SNPs in the same MAF group.

#### 36 Methods for sex phenotypic difference analysis

37 We used linear models including sex and other covariates used in XWAS to test sex differences. Sex was  
38 the term of interest (males coded as 1 and females coded as 2) and the significance was determined via an  
39 F test. A negative effect size of sex indicated males had greater measurement than females. Note that we  
40 also included the interaction terms between sex and age (and age squared) in the model, but we only  
41 focused on the main effect of sex in assessment.

#### 42 Additional results on dosage compensation and X-linked heritability for DTI and fMRI traits



1 Out of the 110 tract-mean traits, 93 exhibited significant  $h_X^2$ , with an average  $h_X^2$  of 1.2% (se = 0.39%)  
2 (Fig. S16), whereas 330 out of 525 DTI PC traits were significant, with an average  $h_X^2$  of 1.3% (se =  
3 0.42%) (Fig. S17). The atlas of enrichment for tract-mean DTI traits are present in Fig. S18.

4  
5 For each pair of tract and DTI metric, there was a correspondence between the number of PCs favoring  
6 full DC and the DC status of the tract-mean trait (Fig. S19). Most tract-mean traits in favor of full DC had  
7 more than three PCs also favoring full DC. In contrast, those favoring no DC typically had no more than  
8 two PCs in agreement. However, we found some outliers. For example, FA of the fornix had the second  
9 to fifth PC favoring no DC, but the tract-mean trait favored full DC. Similarly, MO of the posterior  
10 corona radiata had the second to fifth PC favoring full DC but the tract-mean trait favored no DC. The  
11 latter case is probably because the first PC had extremely high  $h_X^2$  of 4.6%, thus dominating the DC  
12 status. As a result, the top five PCs can capture the major genetic information underlying the DC  
13 mechanism for DTI traits. The heritability analysis for DTI traits on the X-chromosome shows that  
14 functional PCs can provide different dimensions from the tract-mean traits to explain genetic control.

15  
16 Distinct patterns in  $h_X^2$  emerged between rfMRI and tfMRI traits. For rfMRI, language, secondary visual,  
17 and frontoparietal networks had the highest  $h_X^2$  among the mean amplitude traits ( $h_X^2 = 1.67\% \sim 1.79\%$ ),  
18 while the average functional connectivity within primary and secondary visual networks and the one  
19 between these two networks surpassed all other connectivity traits ( $h_X^2 = 2.46\% \sim 2.86\%$ ) (Fig. S20A).  
20 However, tfMRI traits presented a different hierarchy of heritability. The mean amplitudes of  
21 frontoparietal, posterior-multimodal, and dorsal-attention networks were the most heritable ( $h_X^2 =$   
22  $1.13\% \sim 1.53\%$ ), while the average functional connectivity traits between frontoparietal and posterior-  
23 multimodal networks, and that between language and posterior multimodal networks, as well as that  
24 within the language network surpassed all other tfMRI traits ( $h_X^2 = 1.78\% \sim 1.94\%$ ) (Fig. S20B).

#### 25 26 Gene-based analysis and Functional gene mapping

27 MAGMA (126) (v1.08) was applied to XWAS summary statistics for gene-based analysis, which  
28 aggregated individual variant effects to evaluate association signals of genes. We identified 45 genes in  
29 NPR located at 23 genomic regions, associated with 137 traits by controlling for FDR at 0.05 level (Table  
30 S20). *CLIC2*, *DUSP9*, *RAB39B*, *TMLHE*, *VBP1*, and *PJAI* were previously seen as being associated with  
31 regional SA phenotypes (16). *ZNF275*, *DACH2*, and *VMA21* were associated with brain connectivity  
32 measurement (127) and brain shape (55). *FAAH2* was associated with neuroticism measurement (18).  
33 Many detected genes were involved in intellectual disability, educational attainment, and neuropsychiatric  
34 disorders. For instance, *RENBP* (128), *TKTL1* (58), and *MAP7D2* (58) were for schizophrenia; *USP51*  
35 was for educational attainment (5, 60); *DACH2* was for AD (129). *IRAK1* and *TMEM187* were for  
36 Internet addiction disorder (130). Some genes were linked to subtypes of autism. For instance, *TMLHE*  
37 deficiency, leading to a defect in carnitine biosynthesis, was a risk factor for nondysmorphic autism (91,  
38 92). *DACH2* was associated with ASD (131).

39  
40 We mapped the significant SNPs to genes based on functional sequences, resulting in 38 unique genes  
41 associated with 72 traits. These genes exhibited diverse association patterns with SNPs that influenced  
42 brain measurements, neuropsychiatric disorders as well as cognitive abilities (Table S21). For instance,  
43 *DCAF8L1* was linked to educational attainment (60) and neuroticism (18), *EFNB1* was linked to  
44 schizophrenia (61), Parkinson's disease (64) and educational attainment(60), while both *RN7SKP31* and  
45 *RNU6-985P* were linked to neuroticism (132), cognitive function (133) and educational attainment (60),  
46 through SNPs in nearby intergenic regions. Moreover, *SRPX* was tagged by rs35318931 in its exonic  
47 region, correlating with CSF volume, which was a risk factor for anorexia nervosa (134).

48  
49 X-chromosome inactivation (XCI) status (6) was determined for all protein-coding genes identified by  
50 FUMA (Table S22). Among the 96 genes with available XCI information, 65 (67.7%) were previously

1 reported inactive, 12 (12.5%) were escaping, and 19 (19.8%) were variable. When compared with the  
2 reference provided by Tukiainen et al (6), the identified genes were not enriched in any XCI category  
3 (hypergeometric test, p-value > 0.05/3).

4  
5 Long non-coding RNAs (lncRNAs) have emerged as pivotal regulators. For example, *XIST* and *TIST* not only  
6 initiate XCI but also participate in subsequent complex processes (135, 136). We utilized two approaches  
7 – eQTL mapping and H-MAGMA (123) to map significant SNPs to lncRNAs. H-MAGMA incorporates  
8 a chromatin interaction profile to aggregate SNPs to the nearest genes (Methods, Table S23). SNPs  
9 associated with 82 traits were mapped to 14 lncRNAs across 10 genomic regions by controlling FDR at  
10 the 0.05 level.

### 11 Sex phenotypic differences in complex brain imaging traits

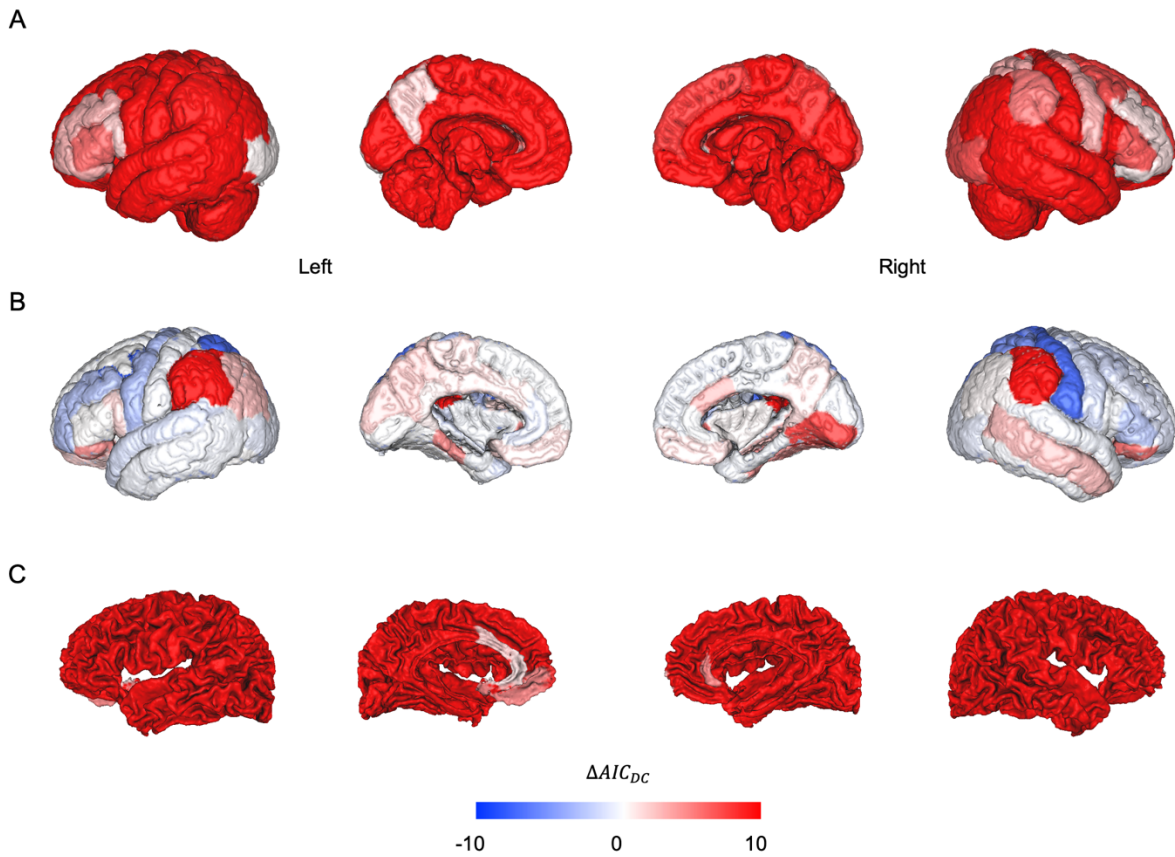
12 Sex differences in the human brain with respect to cortical and subcortical structures, WM  
13 microstructures, and functional connectivity were observed in plenty of studies (97). We evaluated the sex  
14 differences using a linear model and adjusted for potential confounding factors that may bias the sex  
15 effect, including age, genetic principal components, head motion and positions, etc (Supplementary Text).  
16 Out of the total 1,045 traits, 920 (88.0%) were significantly different between sexes after controlling for  
17 FDR at 0.05 level (Table S14). Notably, each category of brain imaging traits maintained a comparable  
18 proportion of significant differences. We then focused on these significant traits in the following  
19 analysis.

20  
21 Set males as reference (males were coded as 1, females were coded as 2), a negative effect size of sex  
22 means males had a greater measurement. The mean effect size of the significant traits was -0.05,  
23 indicating males registered a 5% standard deviation increase compared to females. And 60.5% significant  
24 traits had a negative effect size for sex (proportion test against 0.5, p-value = 0.03). The proportion of  
25 negative effect size varied across trait categories (Fig. S21). For instance, both CT and BV had more than  
26 58% traits with negative effects for sex, but the proportion for SA was only 47.4%. All DTI traits  
27 exhibited the proportion greater than 50%, where 71.8% tract-mean traits exhibited larger values for  
28 males. Remarkably, 79.3% rfMRI traits were larger for males, while the proportion for tfMRI traits was  
29 decreased to 67.8%. In parallel with the proportion results, tract-mean DTI traits and rfMRI traits  
30 displayed the strongest sex bias with the mean effect size less than -0.1.

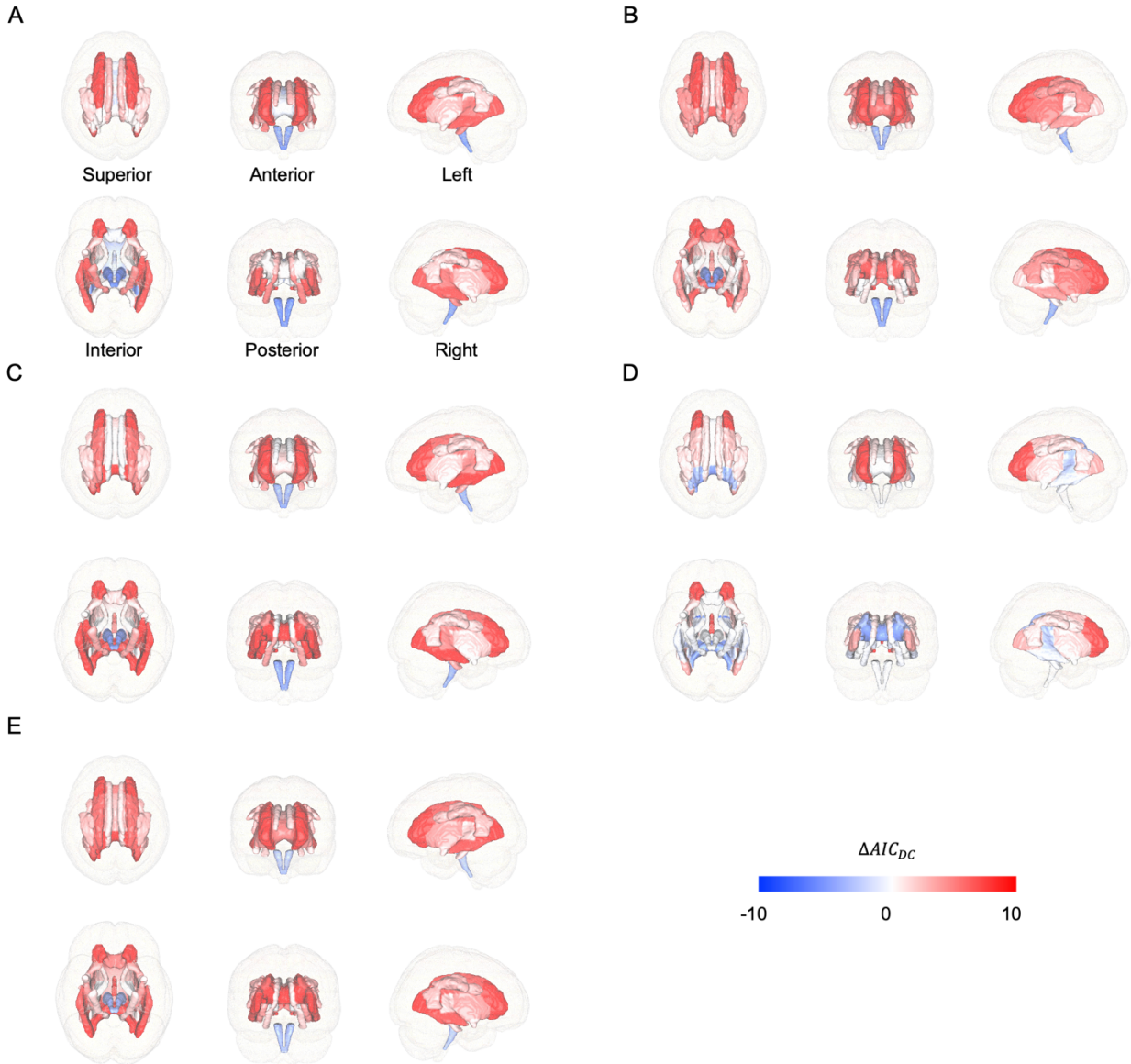
31  
32  
33 By closely examining the sex differences within each modality, we found more significant disparities. For  
34 example, males had larger total BV, cerebrospinal fluid volume, and total SA relative to females, but had  
35 smaller gray and white matter volume and mean CT. Overall DTI mean traits were larger for males except  
36 for RD. The external capsule and the ratio of fornix to stria terminalis were the only two tracts with  
37 negative sex effect evaluated by AD. For rfMRI and tfMRI traits, we discovered that males had greater  
38 mean amplitude for all networks, but females had greater mean functional connectivity related to the  
39 ventral-multimodal network. Most of these discoveries were consistent with previous studies (30, 137-  
40 141).

41  
42 Interestingly, we found significant negative correlation between the sex differences and  $h_X^2$  ( $r = -0.17$ , p-  
43 value =  $3.8 \times 10^{-7}$ ), which means traits with larger measurements in males tend to have greater  $h_X^2$ . We  
44 speculate that this correlation may be explained by larger phenotypic variance in males. We also observed  
45 significant correlation between male:female variance ratio and  $h_X^2$  ( $r = 0.17$ , p-value =  $1.6 \times 10^{-7}$ ).  
46 Estimation of the phenotypic variance ratio is provided in Methods and the results can be found in Table  
47 S2. However, there was no significant signals between sex phenotypic differences and variance ratio ( $r =$   
48  $-0.03$ , p-value = 0.38). We found no significant disparities in sex phenotypic differences among the DC  
49 status (Wilcoxon rank sum test, p-value > 0.19), nor did we find significant disparities in sex phenotypic  
50 differences among  $h_X^2$  enrichment groups (Wilcoxon rank sum test, p-value > 0.22).

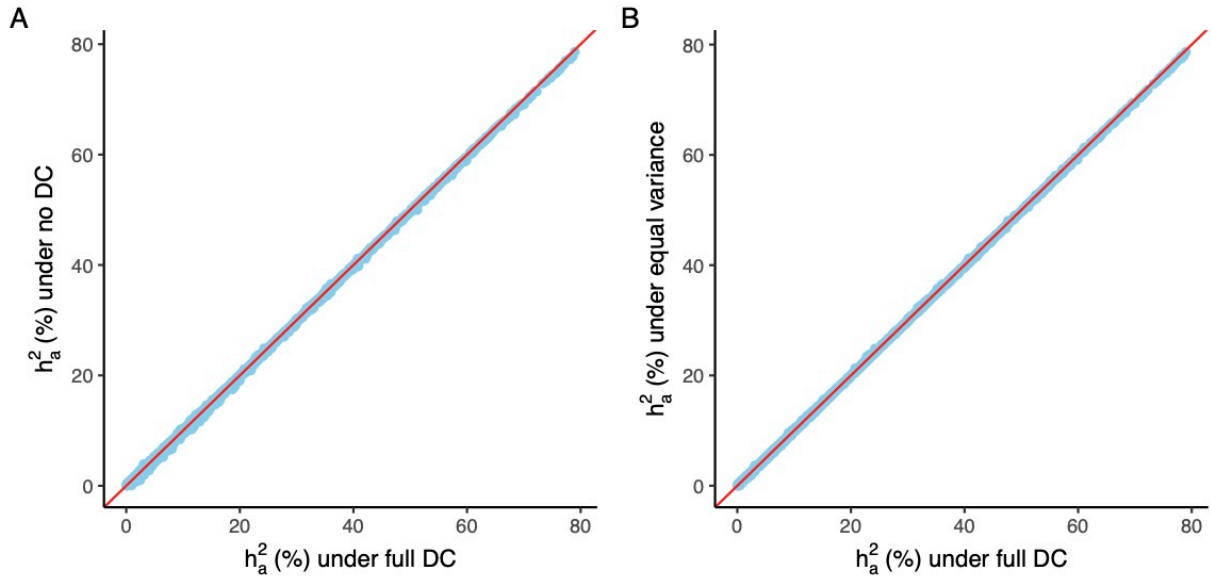
1 There is increasing evidence that both the X-chromosome and sex steroid exposures across the lifespan  
2 can potentially alter the human brain from many aspects (4, 8, 32, 142), and therefore, we are motivated  
3 to investigate genetic underpinnings of human brain with respect to the X-chromosome.



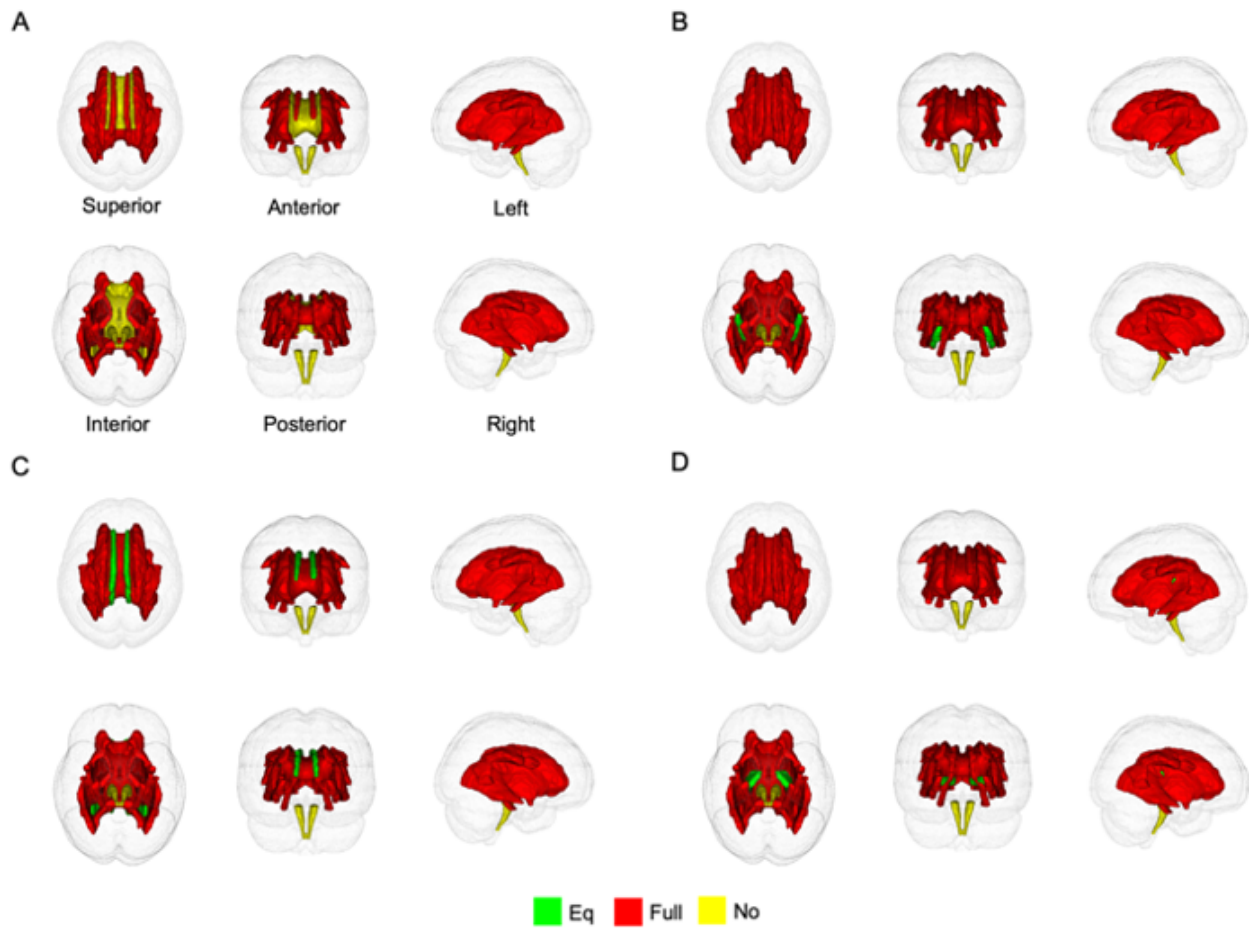
1  
 2 **Fig. S1: Continuous map showing regional variation in the strength of statistical evidence for X-**  
 3 **chromosome dosage compensation in sMRI traits, indexed by  $\Delta AIC_{DC}$ .** There is evidence for full DC  
 4 if  $\Delta AIC_{DC}$  is greater than zero, while it indicates no DC if  $\Delta AIC_{DC}$  is less than zero. It equals zero when  
 5 the equal variance model fits best. Any  $\Delta AIC_{DC}$  that are greater than 10 or less than -10 are truncated to  
 6 10 and -10, respectively, for illustration. A) BV; B) CT; and C) SA.



1  
2 **Fig. S2: Continuous map showing regional variation in the strength of statistical evidence for X-**  
3 **chromosome dosage compensation in DTI traits, indexed by  $\Delta AIC_{DC}$ .** There is evidence for full DC if  
4  $\Delta AIC_{DC}$  is greater than zero, while it indicates no DC if  $\Delta AIC_{DC}$  is less than zero. It equals to zero when  
5 the equal variance model fits best. Any  $\Delta AIC_{DC}$  that are greater than 10 or less than -10 are truncated to  
6 10 and -10, respectively, for illustration. A) AD; B) FA; C) MD; D) MO; and E) RD.

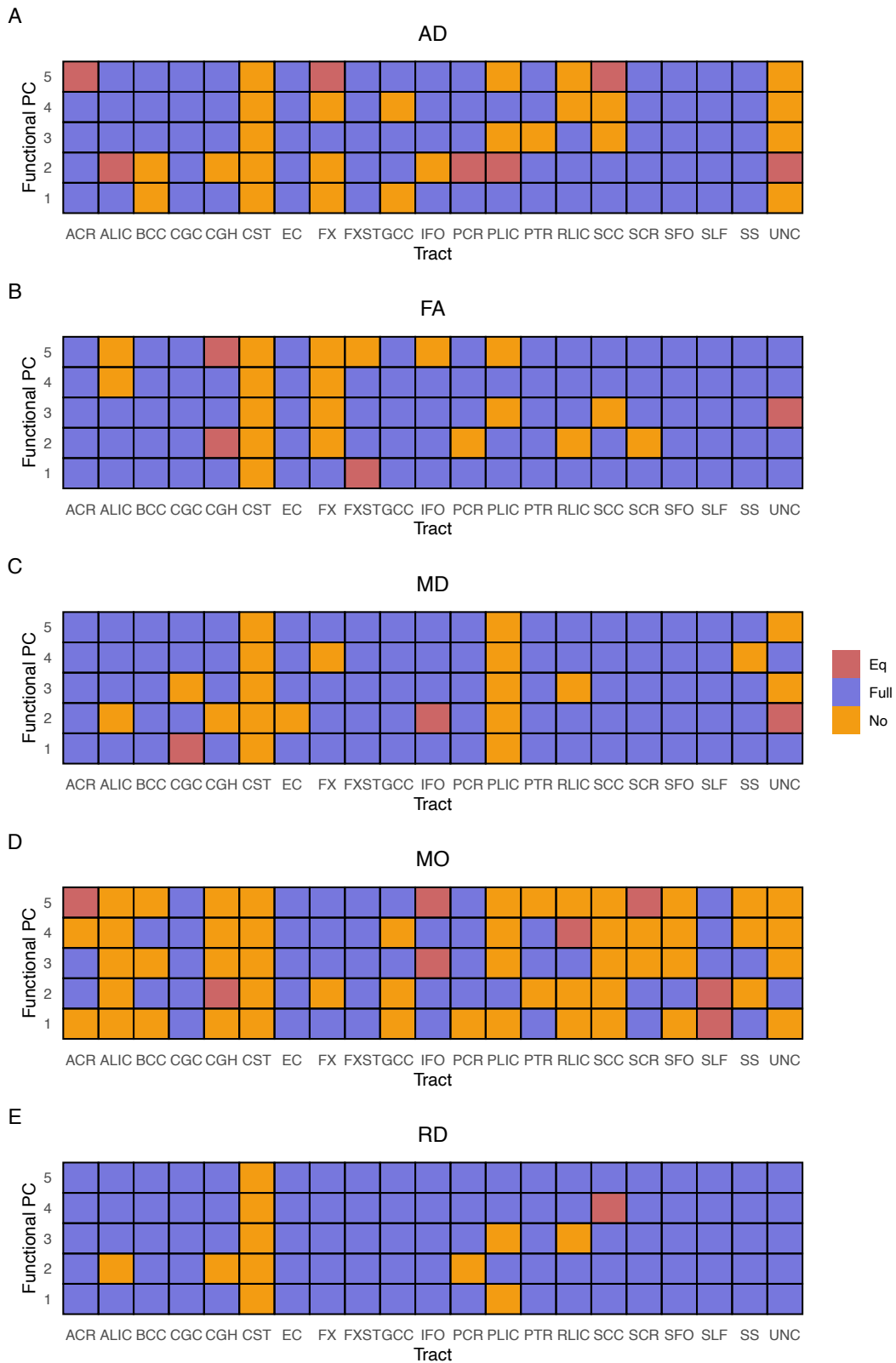


1  
 2 **Fig. S3: Comparisons of autosomal heritability ( $h_a^2$ ) among DC groups.** The  $h_a^2$  is estimated by the  
 3 model including both GRMs of the autosomes and the X-chromosome in GCTA. A) Comparison between  
 4 using full DC and no DC. B) Comparison between using full DC and equal variance.

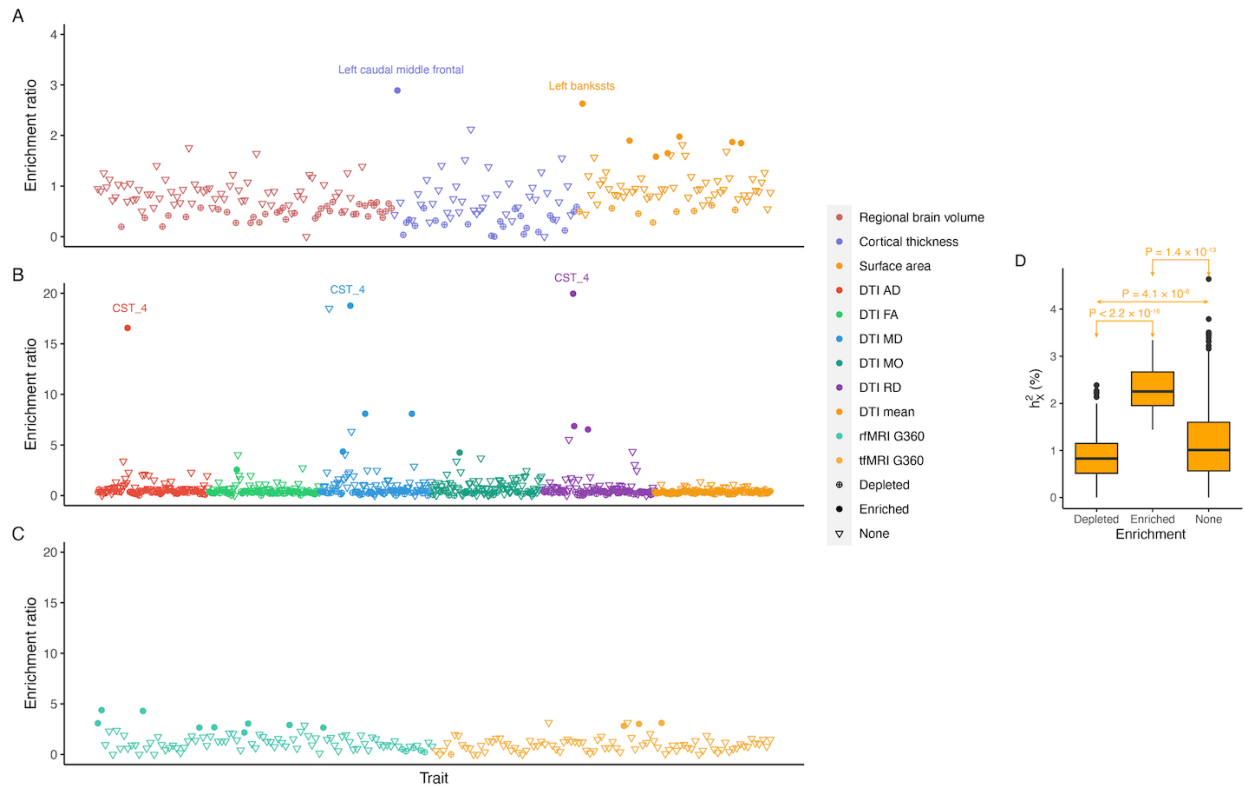


1  
 2 **Fig. S4: Atlases of DC in DTI tract-mean traits evaluated by different metrics. A) AD; B) FA; C)**  
 3 **MD; and D) RD.**

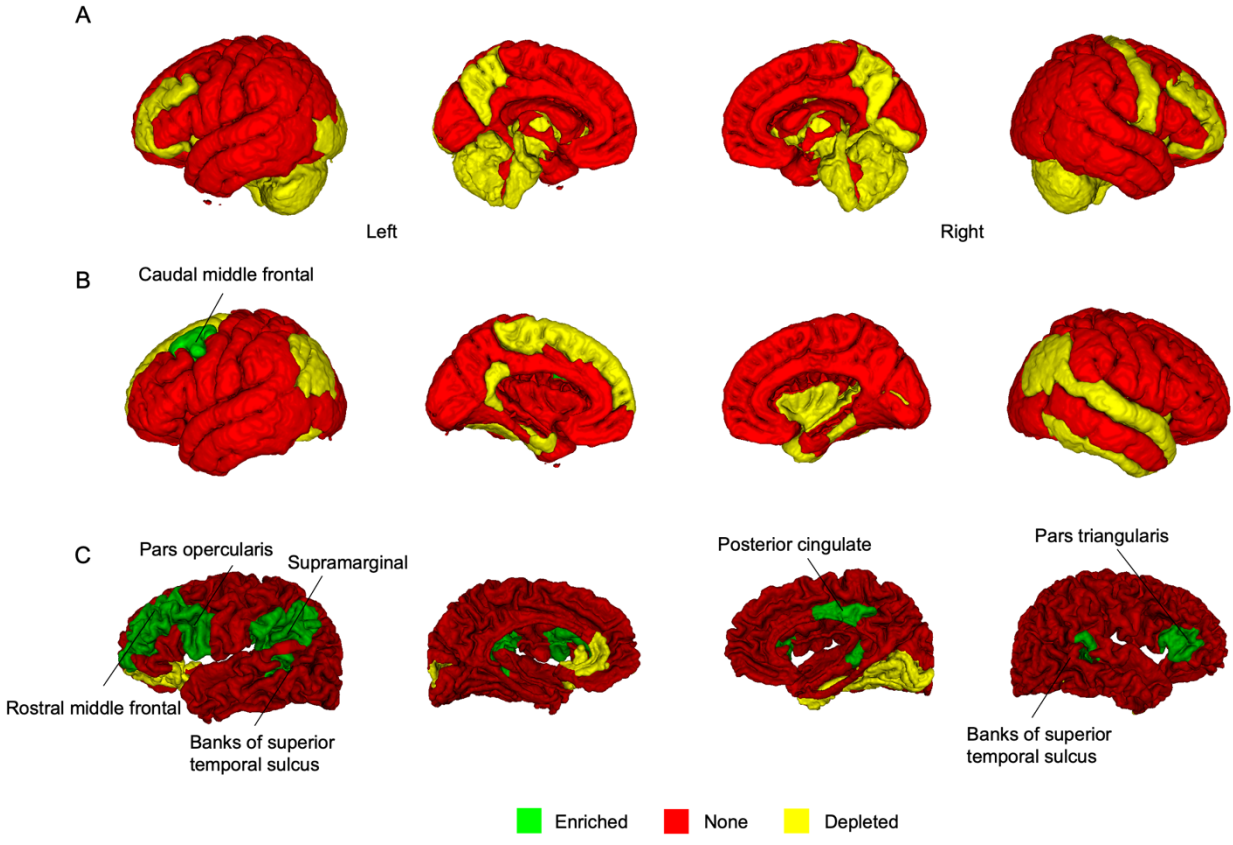




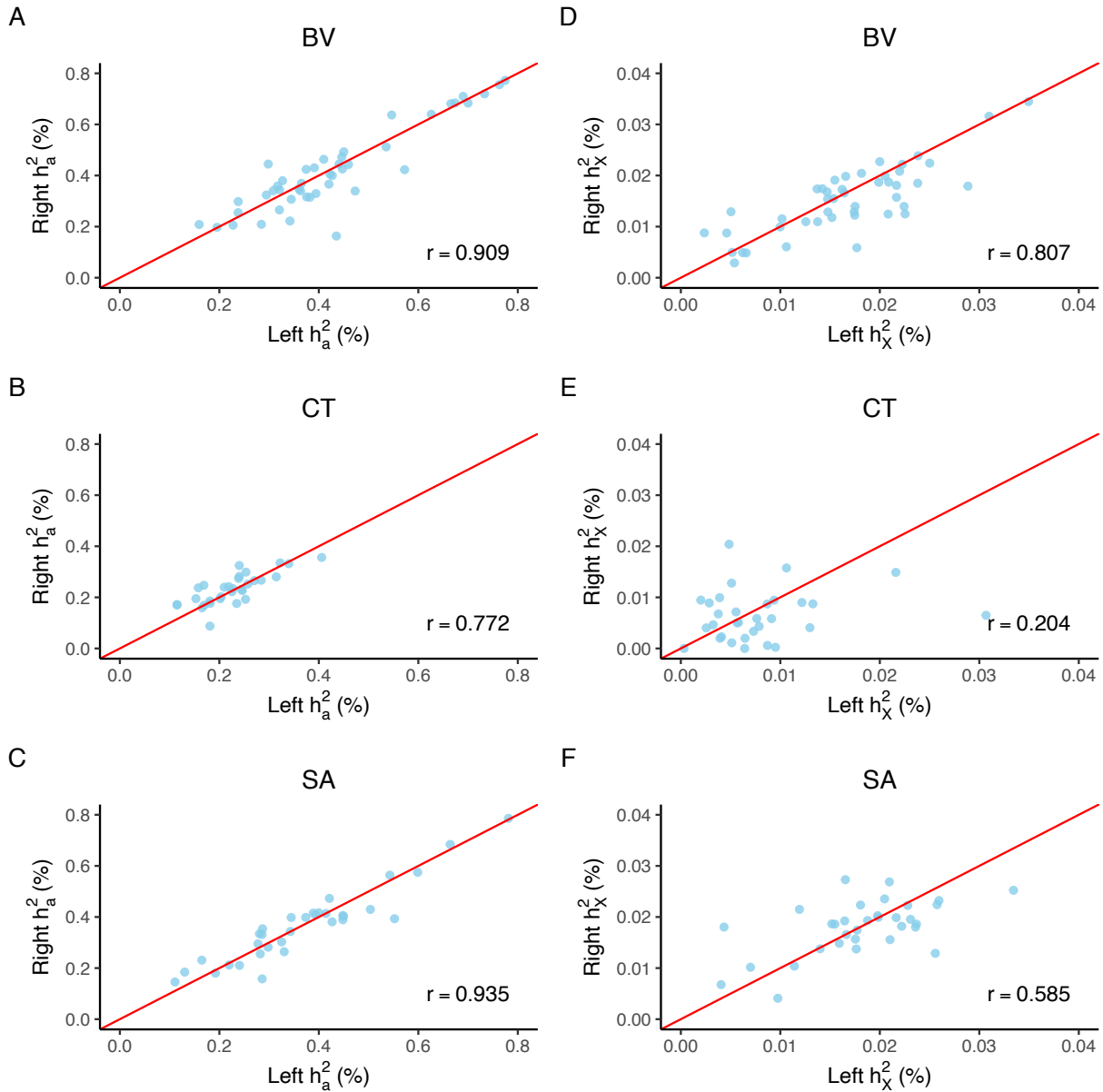
1  
2 **Fig. S5: Atlases of DC of DTI PC traits evaluated by different metrics. A) AD; B) FA; C) MD; D)**  
3 **MO; and E) RD.**



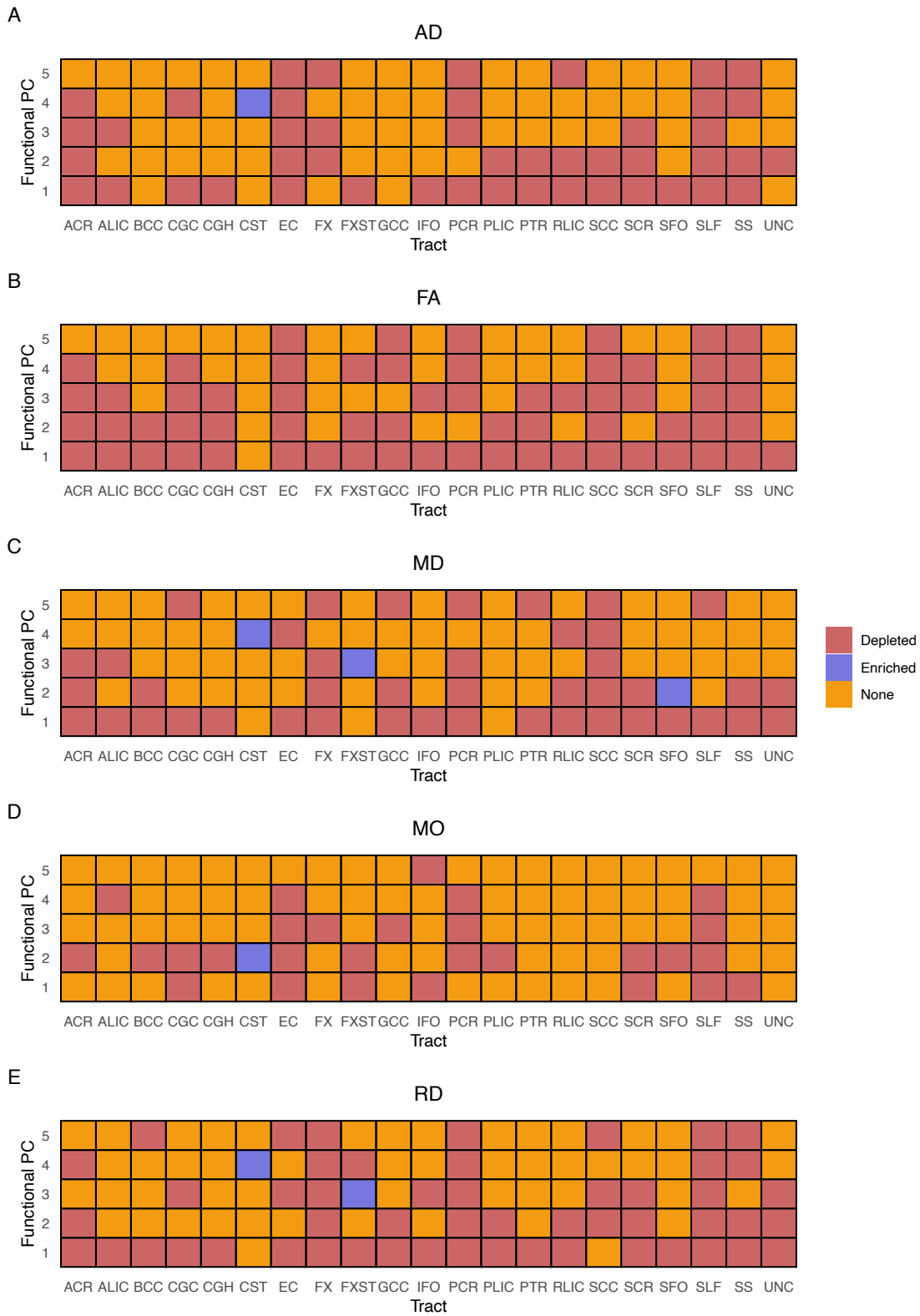
1  
2 **Fig. S6: Enrichment analysis of X-linked heritability ( $h_X^2$ ) for complex brain imaging traits.** A)–C)  
3 Scatter plots of enrichment ratios for individual traits in sMRI, DTI, and fMRI, respectively. The different  
4 trait categories are distinguished by color. Traits with significant enrichment are shown in solid circles,  
5 while traits with significant depletion are shown in crossed circles. Traits with no significant signals are  
6 shown in inverted triangles. Some traits with exceptionally high enrichment ratios are annotated.  
7 “CST\_4” means the fourth PC of corticospinal tract evaluated by a particular metric represented by color.  
8 D) Pairwise comparisons of  $h_X^2$  estimates between enrichment groups. The p-values are the results of the  
9 Wilcoxon rank sum test.



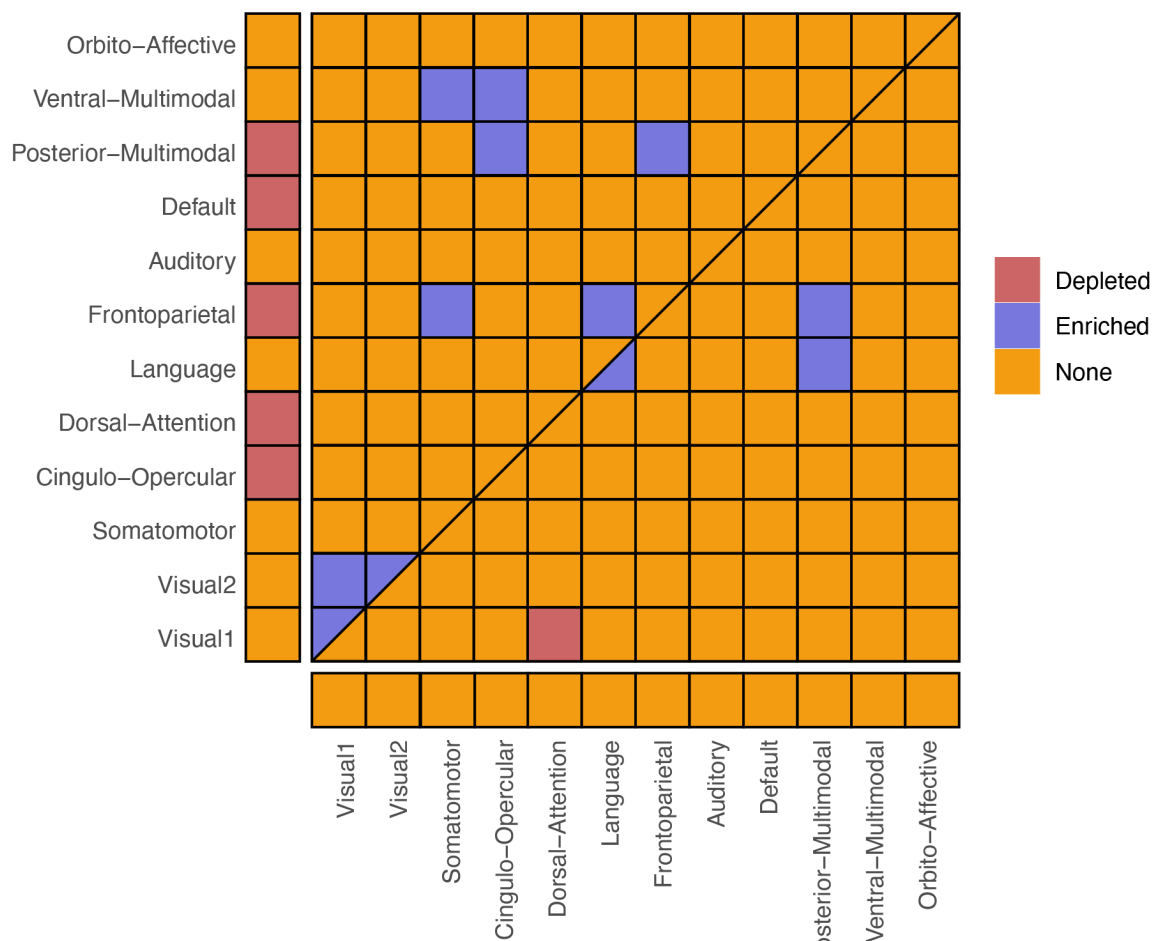
1  
 2 **Fig. S7: Regional patterning of X-linked heritability enrichment for sMRI traits.** Significantly  
 3 enriched traits are annotated. A) regional BV; B) CT; and C) SA.



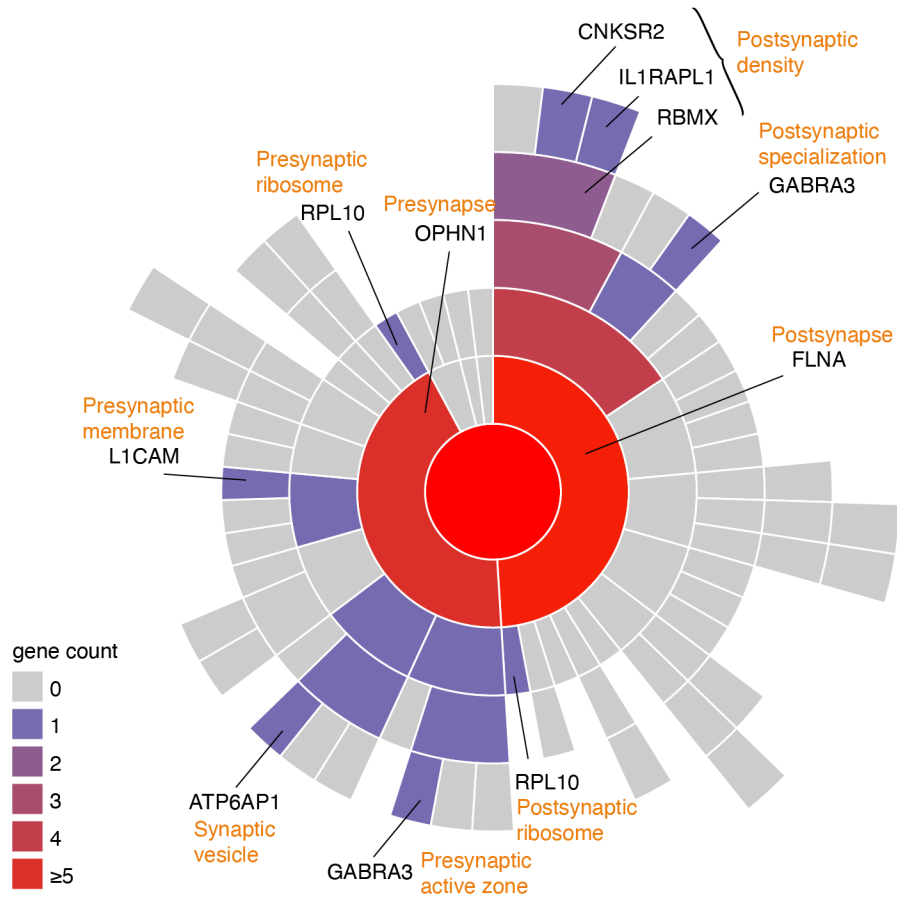
1  
2 **Fig. S8: Symmetricity of X-linked heritability between left and right hemispheres for brain**  
3 **anatomy.** The Pearson correlation coefficient is annotated on each plot. A)-C) Autosomal heritability  
4 ( $h_a^2$ ). D)-F) X-linked heritability ( $h_x^2$ ). Only the coefficient in panel E is not significant at 0.05 level (p-  
5 value = 0.20).



1  
2 **Fig. S9: Atlases of enrichment of DTI PC traits evaluated by different metrics. A) AD; B) FA; C)**  
3 **MD; D) MO; and E) RD.**



1  
 2 **Fig. S10: An atlas of enrichment for fMRI traits.** The upper triangle represents rfMRI, while the lower  
 3 triangle represents tfMRI. The diagonal extending from the bottom left to the top right, is split into two  
 4 triangles: the upper one represents enrichment for intra-connectivity of rfMRI and the lower for tfMRI.  
 5 Cells on the left margin represent enrichment for amplitude traits for rfMRI, while those at the bottom  
 6 indicate enrichment for amplitude traits for tfMRI.

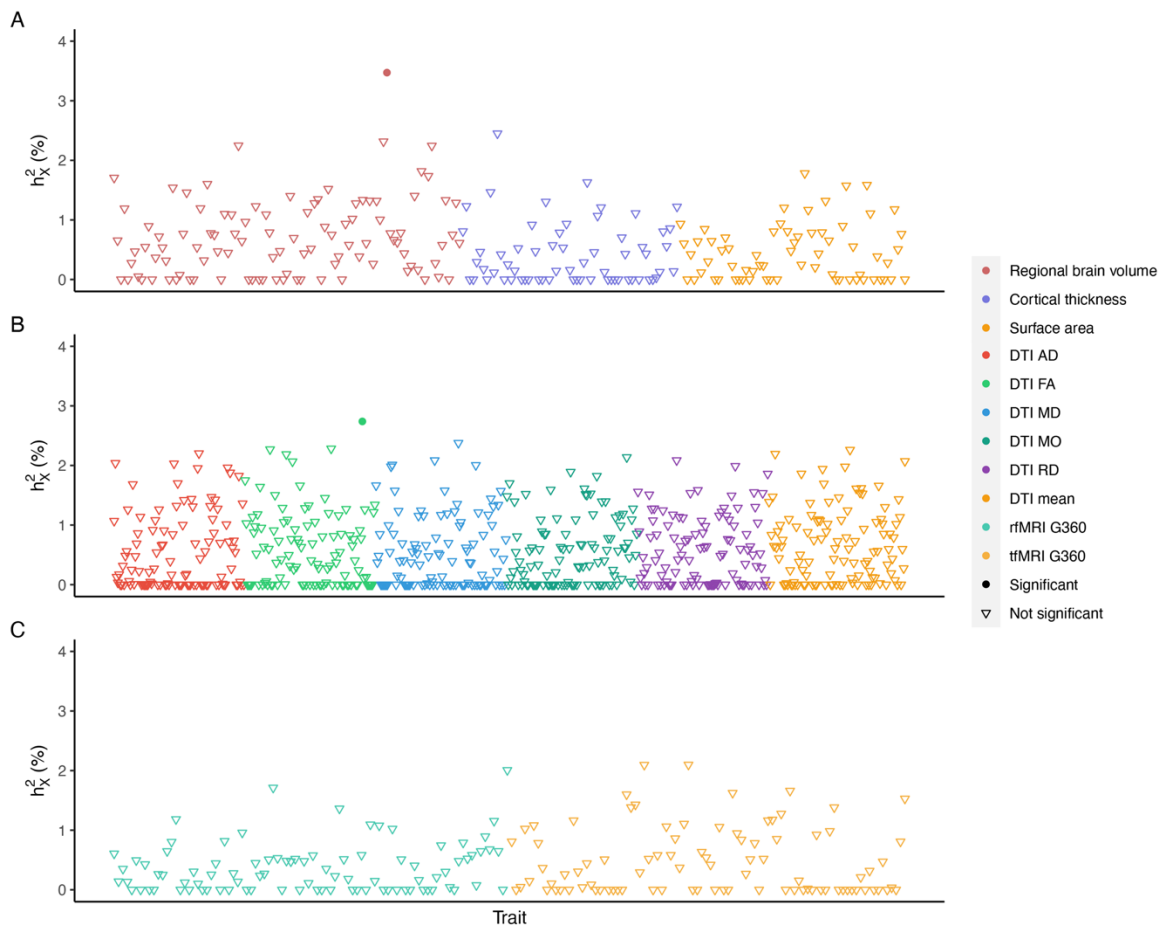


1  
2 **Fig. S11: Mapping of all functionally mapped genes in FUMA to synaptic locations using SynGO.**  
3 Higher-level terms are in the center and child terms are in subsequent rings. The name of the term (in  
4 orange) and genes mapped to the term (in black) are annotated.

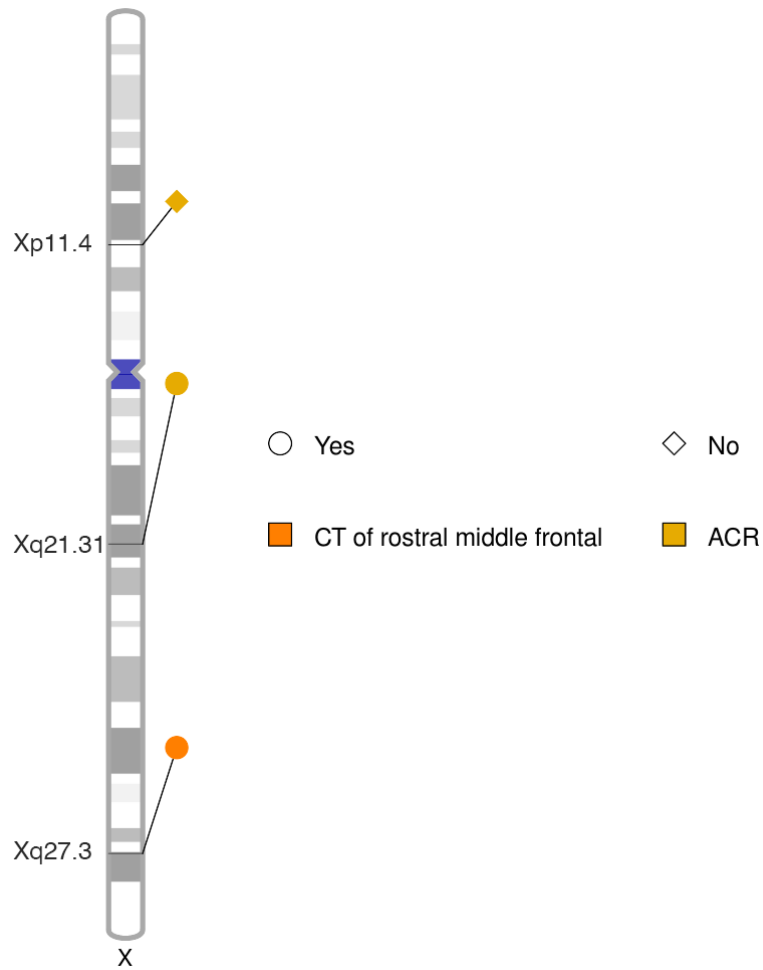




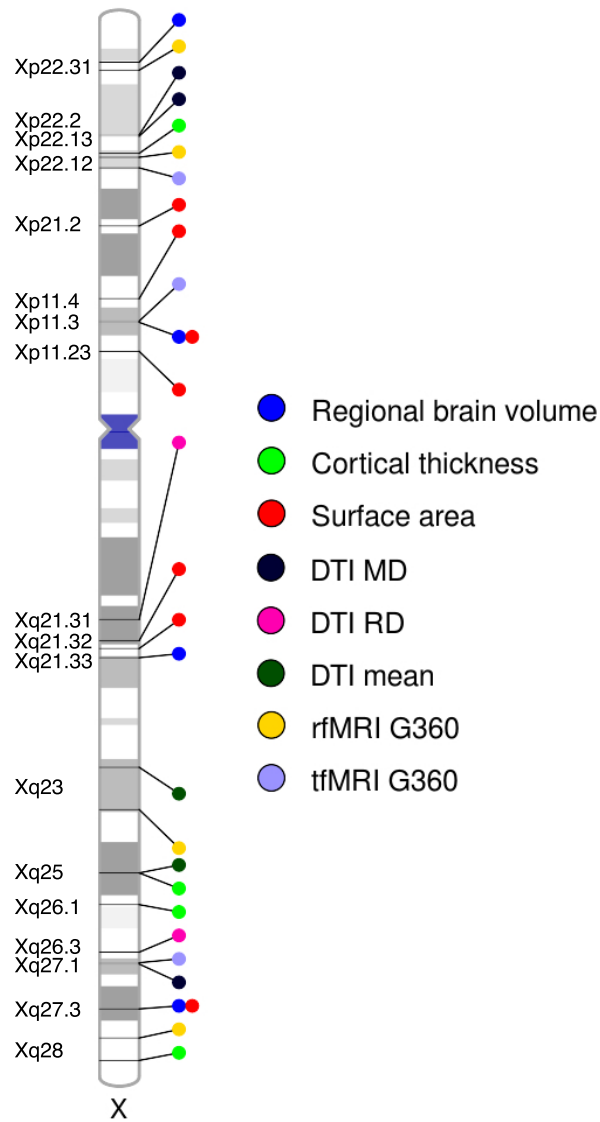
1  
 2 **Fig. S12: Estimated X-linked heritability ( $h_X^2$ ) in males.** The different categories of traits are  
 3 distinguished by color and traits with significant  $h_X^2$  estimates are shown in solid circles. Insignificant  
 4 traits are shown in inverted triangles. In total, 155 traits have significant signals. A) sMRI; B) DTI; and  
 5 C) fMRI.



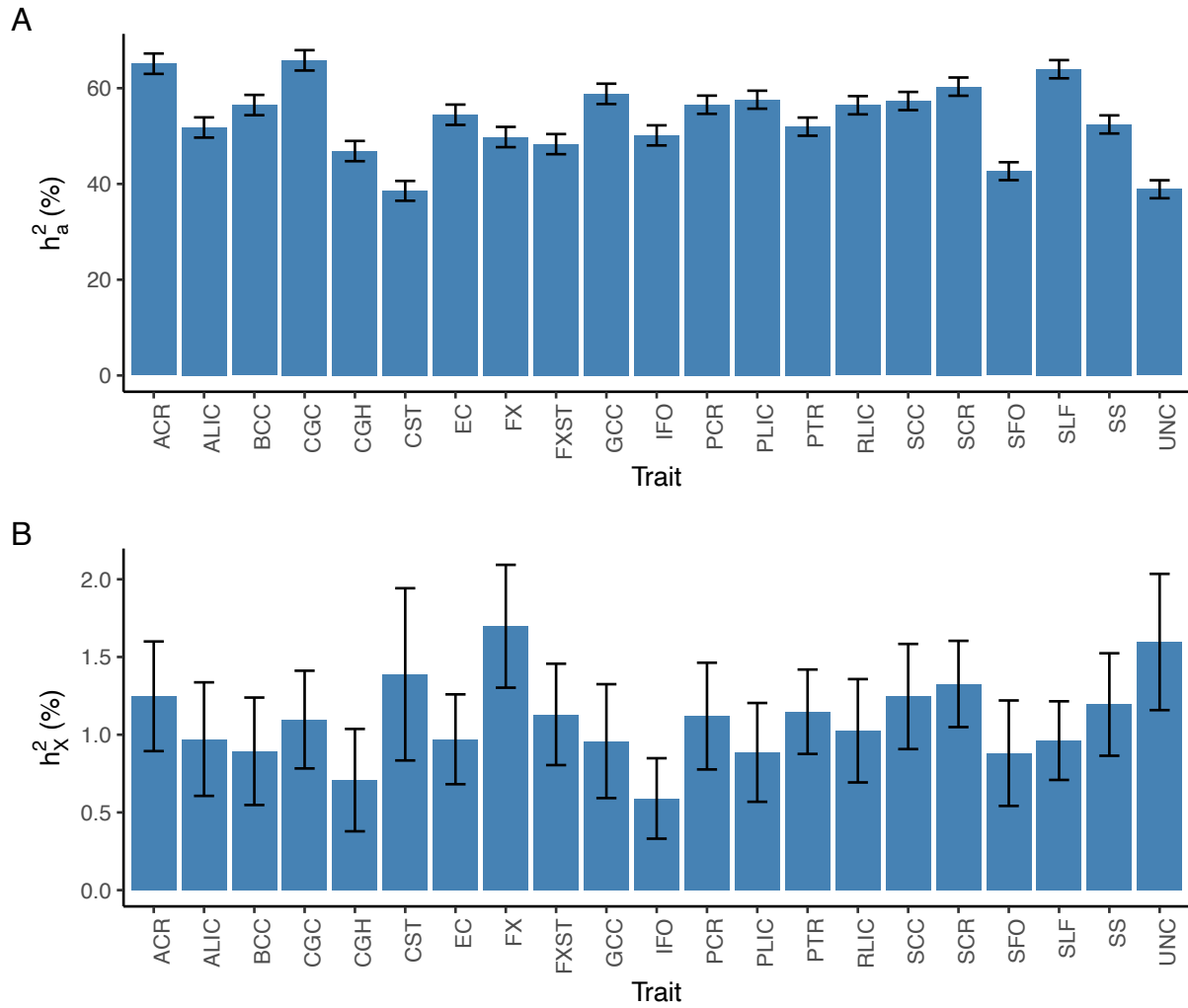
1  
 2 **Fig. S13: Estimated X-linked heritability ( $h_X^2$ ) in females.** The different categories of traits are  
 3 distinguished by color and traits with significant  $h_X^2$  estimates are shown in solid circles. Insignificant  
 4 traits are shown in inverted triangles. In total, only two traits have significant signals. A) sMRI; B) DTI;  
 5 and C) fMRI.



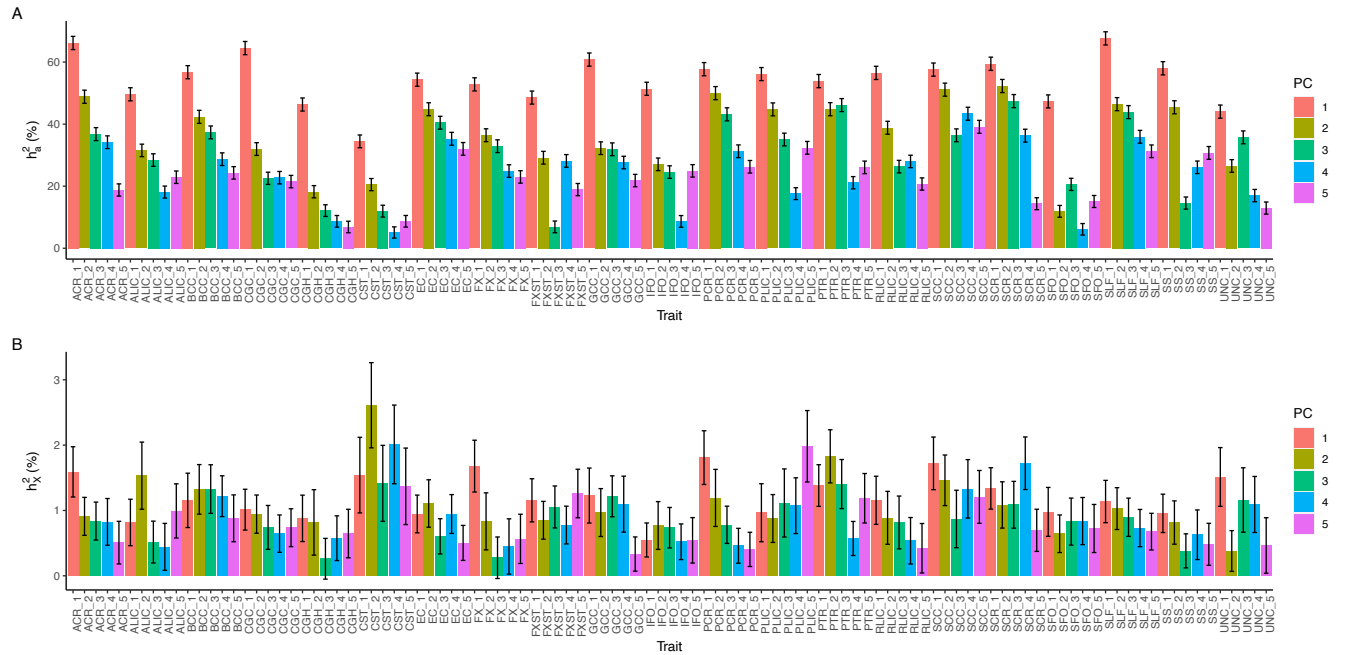
1  
 2 **Fig. S14: Genomic loci displaying significantly different effects between sexes.** Each trait is labeled  
 3 with a different color. The name of each genomic region is annotated on the ideogram. “Yes” and “No”  
 4 mean whether the signal can be recovered or not by a meta-analysis of sex-stratified results using  
 5 Stouffer’s method (Methods).



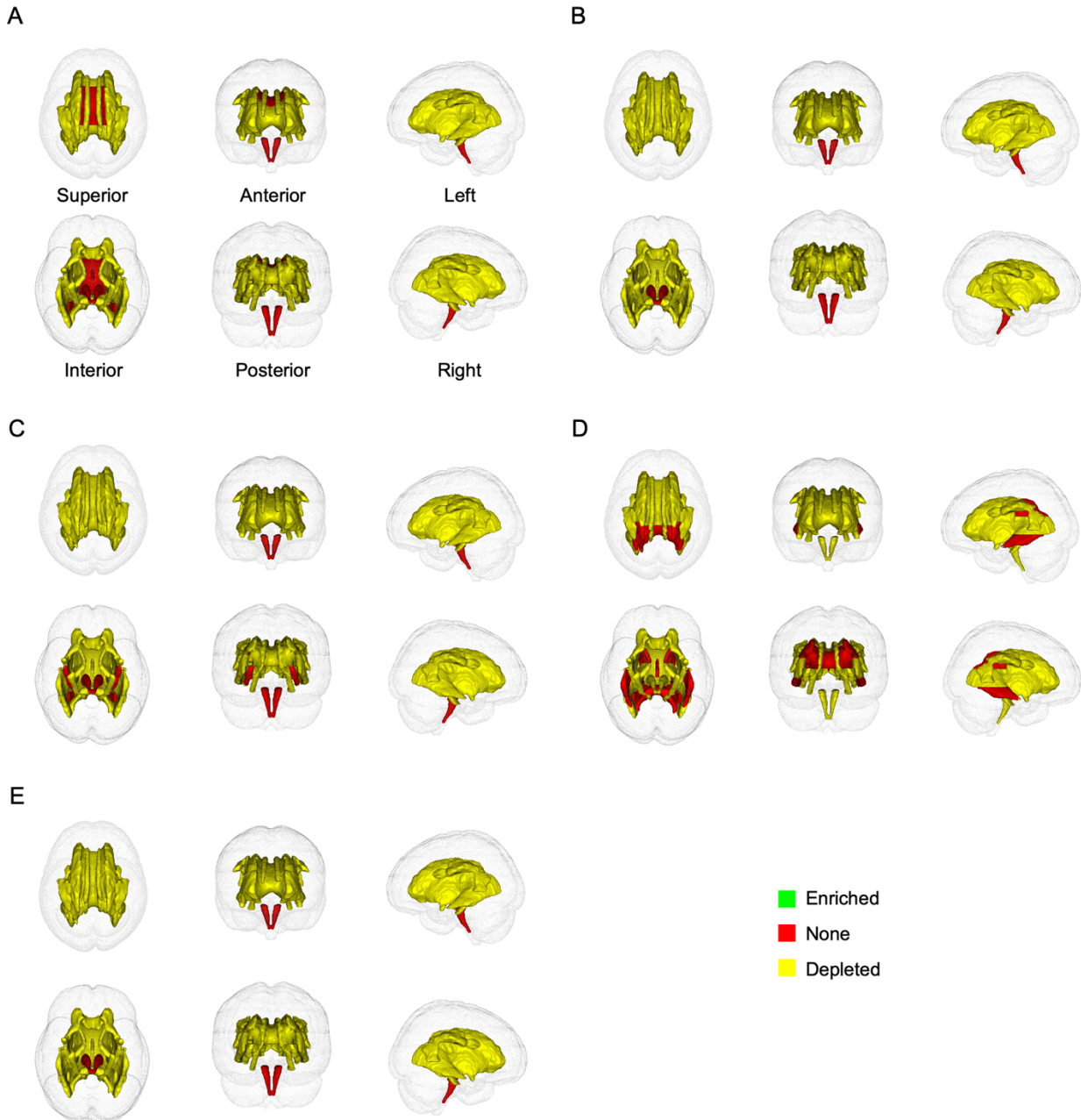
1  
 2 **Fig. S15: Genomic loci displaying significantly different profiles between males classified by**  
 3 **phenotypic quantiles at distribution tails.** Each category of traits is labeled with a different color. The  
 4 name of each genomic region is annotated on the ideogram.



1  
 2 **Fig. S16: Heritability estimates of DTI tract-mean traits.** A) Autosomal heritability ( $h_a^2$ ). B) X-linked  
 3 heritability ( $h_x^2$ ).

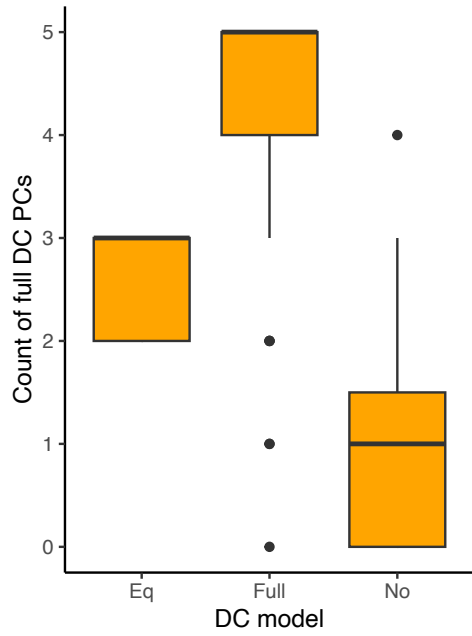


1  
 2 **Fig. S17: Heritability estimates of DTI PC traits.** A) Autosomal heritability ( $h_a^2$ ). B) X-linked  
 3 heritability ( $h_X^2$ ).

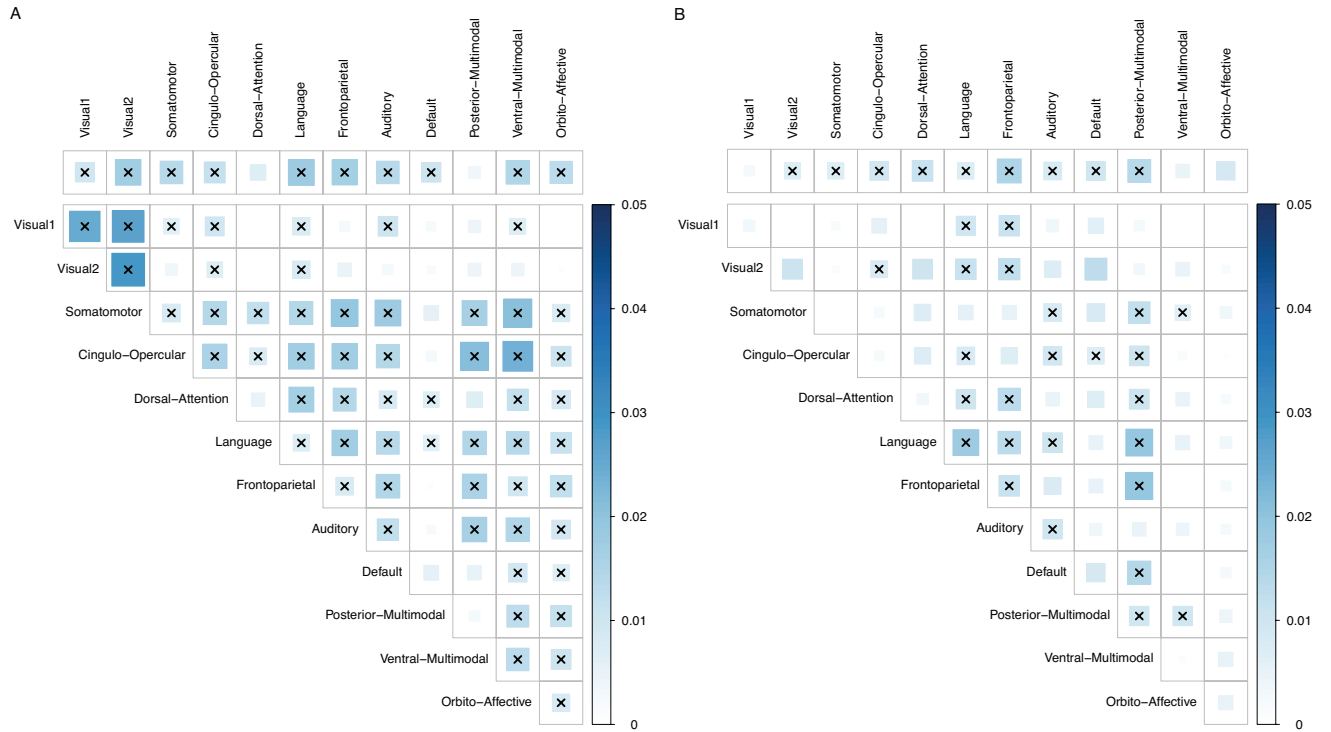


1  
 2 **Fig. S18: Atlases of enrichment of DTI tract-mean traits evaluated by different metrics. A) AD; B)**  
 3 **FA; C) MD; D) MO; and E) RD.**

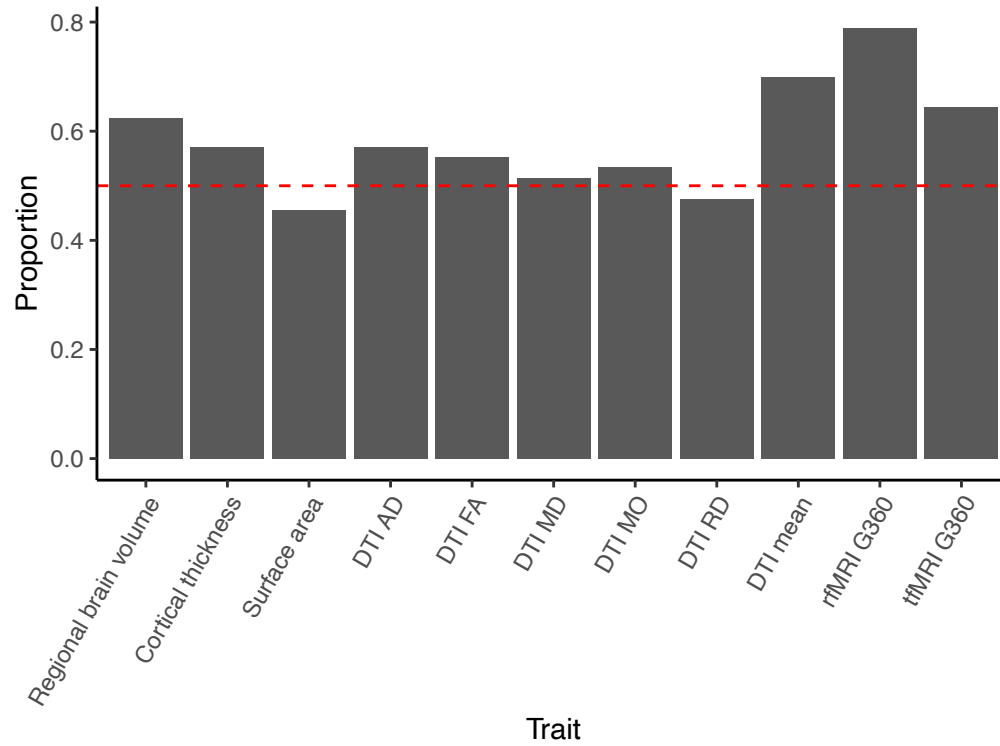




1  
2 **Fig. S19: Consistency between the number of full DC PCs and the DC of tract-mean traits.**



1  
 2 **Fig. S20: Heritability estimates ( $h_X^2$ ) of fMRI traits.** The cells in upper triangle matrix are  $h_X^2$  estimates  
 3 of mean functional connectivity traits. The cells above the top margin are  $h_X^2$  estimates of mean amplitude  
 4 traits. Significant traits are annotated by asterisk. A) rfMRI. B) tfMRI.



1  
2 **Fig. S21: Proportion of traits where males have a larger phenotype than females.** The red dash line  
3 represents 0.5.

- 1 **Table S1.**  
2 **Details for complex brain imaging traits.**  
3 Sheet 1: This sheet shows trait categories and numbers of traits.  
4 Sheet 2: This sheet shows abbreviations and full names for white matter tracts.  
5 Sheet 3: This sheet shows DTI metrics and descriptions.  
6 < in a separate .xlsx file >
- 7 **Table S2.**  
8 **Results of dosage compensation and enrichment analysis for X-linked heritability.**  
9 < in a separate .xlsx file >
- 10 **Table S3.**  
11 **Basic information of XWAS.**  
12 < in a separate .xlsx file >
- 13 **Table S4.**  
14 **Independent significant SNPs identified by LD pruning.**  
15 Sheet 1: Nominal significant SNPs ( $p$ -value  $< 5e-08$ ) in  $LD > 0.6$  were grouped under an independent  
16 significant SNP.  
17 Sheet 2: We did a further round of pruning by restricting the adjusted  $p$ -values of identified independent  
18 significant SNPs (in Sheet 1) less than  $5e-08$ , where the raw  $p$ -values were adjusted by wild bootstrap.  
19 < in a separate .xlsx file >
- 20 **Table S5.**  
21 **Significant loci identified by LD pruning.**  
22 Sheet 1: Nominal significant SNPs ( $p$ -value  $< 5e-08$ ) in  $LD > 0.6$  were grouped under an independent  
23 significant SNP. The LD blocks defined by independent significant SNPs with distances less than 250kb  
24 with each other were merged into a single genomic locus.  
25 Sheet 2: Based on results in Sheet 1, we further restricted that the top SNP for each locus was significant  
26 after adjusting for multiple hypothesis testing for all the traits using wild bootstrap. Whether the trait-  
27 locus pairs were newly identified and whether they were replicated by UKBE subjects only and by meta-  
28 analysis for UKBE, UKBA and UKBSAC subjects were also marked.  
29 < in a separate .xlsx file >
- 30 **Table S6.**  
31 **Significant variants of meta-analysis for sex-stratified XWAS results.**  
32 Sheet 1: The meta-analysis was performed by METAL  
33 ([https://genome.sph.umich.edu/wiki/METAL\\_Documentation](https://genome.sph.umich.edu/wiki/METAL_Documentation)). The raw  $p$ -values were adjusted by wild  
34 bootstrap considering the number of traits. The significant variants were those with adjusted  $p$ -values less  
35 than  $5e-08$ .  
36 Sheet 2: The bolded results are new SNPs that were not identified in the sex-agnostic XWAS.  
37 < in a separate .xlsx file >
- 38 **Table S7.**  
39 **XWAS replication results.** The top SNP for each locus was presented. A locus was replicated if the top  
40 SNP in the replication study was lower than the Bonferroni threshold ( $0.05/72 = 0.00069$ ).  
41 Sheet 1: The XWAS replication results using phase 4 European-ancestry subjects (UKBE,  $n = 4,181$ ).  
42 Replicated associations are bolded.  
43 Sheet 2: The XWAS replication results using phase 1-4 African-ancestry subjects (UKBA,  $n = 295$ ).

1 Sheet 3: The XWAS replication results using phase 1-4 South Asian-ancestry and Chinese-ancestry  
2 subjects (UKBSAC, n = 462).

3 Sheet 4: The individual XWAS replication results were meta-analyzed using METAL  
4 (<https://genome.sph.umich.edu/wiki/METAL>). Replicated associations are bolded.

5 < in a separate .xlsx file >

#### 6 **Table S8.**

7 **Significant variants of meta-analysis for XWAS discovery and white replication results.** The meta-  
8 analysis was performed by METAL ([https://genome.sph.umich.edu/wiki/METAL\\_Documentation](https://genome.sph.umich.edu/wiki/METAL_Documentation)). The  
9 raw p-values were adjusted by wild bootstrap considering the number of traits. The significant variants  
10 were those with adjusted p-values less than 5e-08.

11 < in a separate .xlsx file >

#### 12 **Table S9.**

13 **NHGRI-EBI GWAS catalog (2023.06) lookup for nominal significant SNPs (p-value < 5e-08) and  
14 significant SNPs after further adjustment.**

15 Sheet 1: The lookups in this sheet are based on nominal significant SNPs (p-value < 5e-08).

16 Sheet 2: The lookups in this sheet are based on significant SNPs after adjusting for multiple hypothesis  
17 testing for all the traits by wild bootstrap (adjusted p-value < 5e-08), as well as other SNPs that are in  
18 LD > 0.6 with these significant SNPs.

19 < in a separate .xlsx file >

#### 20 **Table S10.**

21 **Independent significant SNPs mapped to genes based on expression quantitative trait locus (eQTL)  
22 mapping.** PAR SNPs had no results.

23 < in a separate .xlsx file >

#### 24 **Table S11.**

25 **Summary data-based Mendelian randomization results.** The summary statistics of CAGE whole-  
26 blood eQTL analysis were provided by Sidorenko et al. (44). The analysis was conducted by SMR  
27 module in GCTA (<https://yanglab.westlake.edu.cn/software/gcta/#Overview>). We set all parameters as  
28 default.

29 < in a separate .xlsx file >

#### 30 **Table S12.**

31 **Biological annotation for prioritized genes using DAVID Bioinformatics Database  
32 (<https://david.ncifcrf.gov/home.jsp>).**

33 < in a separate .xlsx file >

#### 34 **Table S13.**

35 **Biological annotation for prioritized genes using SynGO (<https://syngoportal.org/>).**

36 < in a separate .xlsx file >

#### 37 **Table S14.**

38 **Sex phenotypic differences in complex brain imaging traits.**

39 < in a separate .xlsx file >

1 **Table S15.**  
2 **X-linked heritability estimated in sex-stratified analysis.** The heritability was computed by GCTA  
3 (<https://yanglab.westlake.edu.cn/software/gcta/#Overview>). The likelihood-ratio test (LRT) was  
4 conducted by comparing the full model including both the genetic relatedness matrices (GRMs) of  
5 autosomes and of the X-chromosome, with the reduced model including the GRM of autosomes only. The  
6 raw p-values were adjusted for the number of traits by Benjamini-Hochberg procedure to control the false  
7 discovery rate (FDR) at 0.05 level.  
8 < in a separate .xlsx file >

9 **Table S16.**  
10 **Significant variants in sex-stratified XWAS.**  
11 < in a separate .xlsx file >

12 **Table S17.**  
13 **Significant variants in meta-analysis for sex-stratified XWAS using UKB discovery data and UKB**  
14 **replication data of white subjects.**  
15 < in a separate .xlsx file >

16 **Table S18.**  
17 **Variants that had significantly different effect sizes between sexes.** The significant variants were then  
18 meta-analyzed across sexes using the Stouffer's method.  
19 < in a separate .xlsx file >

20 **Table S19.**  
21 **Differences in genetic profiles between subjects classified by phenotypic quantiles.**  
22 < in a separate .xlsx file >

23 **Table S20.**  
24 **MAGMA gene-based analysis results.** Summary statistics of nominal significant variants (p-value < 5e-  
25 08) were utilized in the analysis. GRCh37 was used to map SNPs to genes by physical location.  
26 < in a separate .xlsx file >

27 **Table S21.**  
28 **Independent significant SNPs mapped to genes based on functional consequence.** The independent  
29 significant SNPs were in LD > 0.6 with the top SNP of each locus.  
30 < in a separate .xlsx file >

31 **Table S22.**  
32 **Mapped genes identified in functional mapping using nominal significant SNPs at 5e-08 level.**  
33 < in a separate .xlsx file >

34 **Table S23.**  
35 **Gene mapping using chromatin interaction profiles with H-MAGMA.**  
36 < in a separate .xlsx file >  
37