The pivotal role of the X-chromosome in the genetic architecture of the human brain

Running title: XWAS for the human brain

Zhiwen Jiang[1], Patrick F. Sullivan[2], Tengfei Li[3,4], Bingxin Zhao[5,6], Xifeng Wang[1], Tianyou Luo[1], Shuai Huang[1], Peter Y. Guan[1], Jie Chen[1], Yue Yang[1], Jason L. Stein[2,7], Yun Li[1,2,8], Dajiang Liu[9,10], Lei Sun[11,12], and Hongtu Zhu[1,2,4,8,13]*

[1]Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
[2]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
[3]Department of Radiology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
[4]Biomedical Research Imaging Center, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
[5]Department of Statistics and Data Science, University of Pennsylvania, Philadelphia, PA 19104, USA.
[6]Department of Statistics, Purdue University, West Lafayette, IN 47907, USA.
[7]UNC Neuroscience Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
[8]Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
[9]Department of Public Health Sciences, Penn State University, Hershey, PA 17033, USA.
[10]Department of Biochemistry and Molecular Biology, Penn State University, Hershey, PA 17033, USA.
[11]Department of Statistical Sciences, University of Toronto, Toronto, ON M5G 1Z5, Canada.
[12]Division of Biostatistics, Dalla Lana School of Public Health, University of Toronto, Toronto, ON M5T 3M7, Canada.
[13]Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.

**Corresponding author:*
Hongtu Zhu
3105C McGavran-Greenberg Hall, 135 Dauer Drive, Chapel Hill, NC 27599.
E-mail address: htzhu@email.unc.edu Phone: (919) 966-7250

ABSTRACT

Genes on the X-chromosome are extensively expressed in the human brain, resulting in substantial influences on brain development, intellectual disability, and other brain-related disorders. To comprehensively investigate the X-chromosome's impact on the cerebral cortex, white matter tract microstructures, and intrinsic and extrinsic brain functions, we examined 2,822 complex brain imaging traits obtained from n = 34,000 subjects in the UK Biobank. We unveiled potential autosome-X-chromosome interaction, while proposing an atlas of dosage compensation (DC) for each set of traits. We observed a pronounced X-chromosome impact on the corticospinal tract and the functional amplitude and connectivity of visual networks. In association studies, we identified 50 genome-wide significant trait-locus pairs enriched in Xq28, 22 of which replicated in independent datasets (n = 4,900). Notably, 13 newly identified pairs were in the X-chromosome's non-pseudo-autosomal regions (NPR). The volume of the right ventral diencephalon shared genetic architecture with schizophrenia and educational attainment in a locus indexed by rs2361468 (located ~3kb upstream of *PJA1*, a conserved and ubiquitously expressed gene implicated in multiple psychiatric disorders). No significant associations were identified in the pseudo-autosomal regions (PAR) or the Y-chromosome. Finally, we explored sex-specific associations on the X-chromosome and compared differing genetic effects between sexes. We found much more associations can be identified in males (33 versus 9) given a similar sample size. In conclusion, our research provides invaluable insights into the X-chromosome's role in the human brain, contributing to the observed sex differences in brain structure and function.

KEYWORDS: X-chromosome; UK Biobank; Dosage compensation; Brain disorders; Sex differences; Genetic architecture

The genetic foundations governing gene regulation on the X-chromosome are inherently complex due to the XY sex-determination system[1]. In genetic females, who possess two X-chromosomes compared to males' one, an X-chromosome is silenced, either at random or with a bias toward a particular parental copy, to equalize the transcriptional dosages of X-linked genes between the sexes[2,3]. This process, termed dosage compensation (DC) or X-chromosome inactivation (XCI), ensures that every female becomes a cellular mosaic, with each cell containing one of the two possible active X-chromosomes[4]. However, DC is not absolute. Only about 60-75% of X-linked genes undergo full silencing[5], and the process can be tissue-specific[6,7]. Consequently, several factors, such as DC itself and individual variances in the expression of parent-of-origin genes, can influence female X-linked gene expression[8]. While most somatic cells in both sexes have only one active X-chromosome, the expression levels of genes on the X-chromosome can be amplified to match the expression levels of autosomal genes[9] (Ohno's hypothesis). This compensatory elevation in the expression of X-linked genes has been noted across several species, including human[10] and mouse[10-12], and other species[10].

The X-chromosome is home to a wealth of genes predominantly expressed in brain tissues[10,13] as the chromosome-X to autosome expression ratio exceeds one[10], underscoring the pronounced influence of the X-chromosome on brain anatomy, connection, and functions[14-17]. Mounting evidence has shown that the X-chromosome significantly impacts a myriad of neurological diseases and psychiatric disorders from both genetic and epigenetic viewpoints[5,18-20]. For instance, X-linked intellectual disability is a well-studied brain disorder and by 2022, 162 associated genes have been identified[21]. Remarkably, the concentration of genes on the X-chromosome is double that of intellectual disability-associated genes on the autosomes[21]. Chromosome mutations, such as sex chromosome aneuploidy, can have profound implications on brain structures[14,22], cognitive capacities[23,24], behaviors[24], and both neurological[24,25] and psychiatric disorders[26,27]. A combination of factors – including the dense presence of associated genes, varying gene expression between sexes, and the influence of epigenetic processes steered by sex steroid hormones – establishes the X-chromosome as a nexus for sex differences in the human brain across various age groups[4,7,8,28-33]. However, in spite of its pivotal role, the X-chromosome frequently goes unnoticed in genome-wide association studies (GWAS), as discussed by Wise et al.[34] and Sun. et al.[35].

In prior research, Smith et al.[17] conducted a GWAS for 3,144 complex brain imaging traits (n = 22,138) and identified four genome-wide significant loci on the X-chromosome with the top SNP from each locus registering a p-value $< 7.94 \times 10^{-12}$. They also characterized the genetic loci using eQTLs and examined the genetic co-architectures between brain traits and health-related disorders. Mallard et al.[16] focused on investigating DC for regional brain measurements, including cortical volumes (CVs), cortical thickness

(CT), and surface area (SA). They first observed a notable enrichment of X-chromosome influences on several regions of interest (ROIs) related to SA and then explored the genetic underpinnings of these ROIs through an X-chromosome-wide association study (XWAS[36]). However, the study of Smith et al.[17] used a smaller discovery sample size (n = 22,138) than is currently available and the work of Mallard et al.[16] was limited to brain anatomy. What is more crucial is that there remains a significant knowledge gap regarding the DC profile for brain imaging traits, as well as the extent to which sex differences in the human brain are influenced by the X-chromosome.

An overview of the present study is depicted in **Fig. 1**. We amassed 2,822 complex brain imaging traits derived from structural magnetic resonance imaging (sMRI) for cortical and subcortical structures of gray matter, diffusion MRI (dMRI) for microstructures of white matter tracts, resting-state functional MRI (rfMRI) and task-evoked functional MRI (tfMRI) for intrinsic and extrinsic brain functions, respectively (**Table S1**). Specifically, sMRI traits include BV, CT, and SA; dMRIs traits include axial diffusivity (AD), fractional anisotropy (FA), mean diffusivity (MD), mode of anisotropy (MO), and radial diffusivity (RD) along tracts and their tract-mean and functional principal component (PC) traits (hereafter DTI traits); and rfMRI and tfMRI traits generated using the Glasser360[37] atlas (hereafter G360 traits). Additionally, rfMRI images were subjected to whole-brain spatial independent component analysis (ICA, hereafter ICA traits). More details are in the Methods. A brief introduction and annotation of the DTI and ICA traits can be found in **Table S1**. The robustness and reproducibility of all trait measurements have been established in our previous studies[38-41].

We first meticulously assessed the DC status of each trait using model selection, while gauging the narrow-sense heritability tied to the NPR SNPs of the X-chromosome by using a sex-agnostic analysis (**Fig. 1**). We then created a detailed DC atlas for every trait cluster. In association analysis, we tested the NPR and PAR SNPs on the X-chromosome, in addition to the SNPs on the Y-chromosome. By implementing the optimal DC model for the NPR SNPs, we aimed to enhance our statistical accuracy. Interestingly, our brain imaging traits showed genetic overlaps with both brain-related disorders and educational benchmarks. Our further analyses probed into sex differences, focusing on facets like heritability, variance in phenotype, genetic associations, and genetic impact. These findings offer invaluable insights into the genetic factors influencing sex-based differences in the human brain's structure and functions. With this research, our goal is to deepen the understanding of the X-chromosome's role in human brain function and development. Our findings will pave the way for enriching future endeavors in biology, clinical sciences, and psychiatry.

RESULTS

Dosage compensation in the X-chromosome

We inferred a global DC status for each trait through model selection. Specifically, we employed three different model specifications: full DC, no DC, and equal variance to jointly estimate the narrow-sense heritability ascribed to autosomes (h_a^2) and the heritability credited to the NPR on the X-chromosome (h_X^2) using GCTA[42] (Methods). In these models, females were always coded as {0, 1, 2}, but males were coded as {0, 2} for full DC (known as random X-inactivation[43]), {0, 1} for no DC, {0, $\sqrt{2}$} for equal variance. These models illustrate the differing genetic variance between sexes, as males have twice the genetic variance as much as females in full DC, half in no DC, and identical in the equal variance model. For each trait, the model exhibiting the lowest Akaike information criterion (AIC) was selected as the best model. The DC status inferred in this way is not gene-specific but theoretically matches the DC based on the aggregated association statistics on the X-chromosome. Since the heritability attributable to the X-chromosome hinges on the inactivation status of the X-linked loci affecting a trait[42,44], we can estimate the X-linked heritability more accurately and enhance the statistical power of our association analysis by designating the DC status for each trait. For clarity in subsequent discussions, we did not distinguish h_X^2 (h_a^2) from their sample estimates $\widehat{h_X^2}$ ($\widehat{h_a^2}$) if there was no confusion.

Out of the 2,822 traits analyzed, 2,810 exhibited a total heritability ($h_a^2 + h_X^2$) that was positive, revealing that 69.4% favored full DC, 22.6% favored no DC, and 7.97% favored equal variance. **Fig. 2A** shows the distribution of DC status for each trait set (see **Table S2** for detail). The traits MO, CT[16], and ICA had the highest percentage of no DC preference. For the traits leaning towards full DC, no DC, and equal variance, the average h_X^2 estimates of 0.672% (se = 0.315%), 0.938% (se = 0.581%) and 0.250% (se = 0.417%), respectively (**Fig. 2B**). The differences in average h_X^2 across the three DC groups were statistically significant (Wilcoxon rank sum test, pairwise p-values < 5.5×10^{-14}). Higher h_X^2 estimates are reasonable for traits favoring the no DC because when the cumulative gene expression dosage amplifies due to the evasion from XCI, the X-chromosome assumes a more dominant role in genetic regulation.

Our analysis revealed that traits favoring full DC on the X-chromosome exhibited a stronger autosomal heritability compared to those favoring the alternative DC models (Wilcoxon rank sum test, pairwise p-values < 0.005, **Fig. 2C**). Specifically, traits with full DC exhibited an average heritability $h_a^2 = 20.0\%$. In comparison, no DC traits showed $h_a^2 = 15.4\%$, and equal variance traits showed $h_a^2 = 12.6\%$. This distinction was not caused by estimation bias, since for any given trait, h_a^2 estimated in different DC

models were consistent (**Fig. S1**). The observed differentiation suggests a potential autosome-X-chromosome interaction, such as DNA methylation on autosomes which is *trans* regulated by the X-chromosome[45]. Concurrently, we theorize that traits aligned with no DC and equal variance models might be more influenced by non-genetic contributors, such as environmental factors.

Under the equal variance model, the ratio of h_X^2 for a trait between males and females, denoted as $\rho_X = \frac{h_{X,male}^2}{h_{X,female}^2}$, should be two for traits favoring full DC and 0.5 for those favoring no DC. To verify whether the inferred DC aligned with this assumption, we estimated sex-specific heritability for each trait, then calculated ρ_X (Methods). We observed that, ρ_X=1.94, 0.874, and 1.16 for traits favoring full DC, no DC, and equal variance, respectively (**Fig. S2A**). We further compared the heritability ratios stratified by sex for autosomes, denoted as $\rho_a = \frac{h_{a,male}^2}{h_{a,female}^2}$, across traits based on their respective DC models for the X-chromosome. The observed differences among the three DC models were minor: ρ_a=0.983 for full DC, ρ_a=1.01 for no DC and ρ_a=1.01 for equal variance, all close to the expected value of one (**Fig. S2B**). Among these, only the X estimate for no DC notably strayed from the expected value (z-test p-values > 0.05/6). This implies that our model selection strategy is generally effective in identifying the appropriate model. The higher-than-expected value of X in the no DC category (when compared to 0.5) could potentially arise from classification errors and biological heterogeneity, such as variations in sex hormone levels[4,44].

Patterns of dosage compensation and X-linked heritability for complex brain imaging traits

We jointly estimated h_a^2 and h_X^2 via the GREML analysis of GCTA[42] based on the optimal DC model for each trait. Of the 2,810 traits with non-zero total heritability ($h_a^2 + h_X^2 > 0$), 1,118 (39.8%) traits displayed significant heritability with an average of $h_X^2 = 1.27\%$ (se = 0.406%) after adjusting for multiple comparisons by controlling false discovery rate (FDR) at the 0.05 level (**Fig. 2D-F and Table S2**). The majority of BV and SA traits were significant, succeeded by DTI tract-mean traits and rfMRI G360 traits. In contrast, CT and rfMRI ICA showcased the fewest traits with notable heritability estimates.

In the sex-stratified analysis, a larger number of traits demonstrated significant h_X^2 for males compared to females (**Fig. S3-4 and Table S3**). We then assessed the relative contribution of the X-chromosome to the total heritability, denoted as $\frac{h_X^2}{h_a^2+h_X^2}$ (**Fig. S5A**). Generally, the X-chromosome was pivotal in accounting for genetic variance, especially for fMRI traits, with values ranging between 4.68% and

5.99%. Additionally, we evaluated the enrichment of h_X^2 by contrasting the observed h_X^2 with the expected ones The expected h_X^2 represents heritability attributed to a genome segment of a comparable size (Methods). Out of all the traits, 45 (1.60%) displayed enriched h_X^2, while a substantial number of traits (476, 16.9%) exhibited depletion in h_X^2 (**Fig. S5C-E and Table S2**). The remaining traits did not demonstrate any significant deviation. Groups with different enrichment levels displayed varied mean h_X^2 (**Fig. S5B**), with Wilcoxon rank sum test yielding pairwise p-values < 1.68×10^{-11}. The limited presence of traits with enriched h_X^2 can be traced back to male haploidy and random XCI observed in females[44], as both factors reduce the genetic variance attributed to the X-chromosome.

Differences were observed in patterns of DC across measures of brain anatomy. All SA traits favored full DC, while all but one BV trait (specifically the optic chiasm) also favored full DC. Contrastingly, about half of CT traits leaned towards either no DC or exhibited equal variance. While the mean CT favored full DC, those favoring no DC were spread across numerous cortical regions (**Fig. 3A**). Previous research has suggested that the CT of the motor cortex may be influenced by genes that escape typical regulations[46]. Our findings support this, as the CT for both the precentral and postcentral cortex (located near the motor cortex) showed patterns consistent with no DC.

The X-chromosome significantly impacts variations in brain anatomy. Of the 230 traits examined, 179 showed notable h_X^2 estimates with average $h_X^2 = 1.71\%$ (se = 0.379%). Our h_X^2 and h_a^2 estimates for global brain measures echoed findings from previous studies[16,38,47,48]. We observed no significantly enriched traits for BV, but we identified one for CT (left caudal middle frontal) and three for SA (left banks of superior temporal sulcus, left rostral middle frontal, and right pars triangularis) (**Fig. S6**). Notably, the left caudal middle frontal showcased the highest X-chromosome heritability in CT ($h_X^2 = 3.07\%$), where the X-chromosome accounted for over 14% of total heritability. There are distinct differences in h_X^2 between CT traits and those of BV and SA. Firstly, the mean h_X^2 for BV and SA are double that of CT. Secondly, although the correlations of h_a^2 for traits in left and right hemispheres were comparable among BV, CT, and SA, h_X^2 showed the highest correlation for BV ($r = 0.807$, p-value = 2.05×10^{-11}), then SA ($r = 0.585$, p-value = 0.0003) and lastly CT ($r = 0.204$, p-value = 0.20) (**Fig. S7**), given the standard errors of h_X^2 estimates were similar among brain anatomy ($se = 0.362\% \sim 0.403\%$). Overall, the X-chromosome has a pronounced effect on BV and SA, but its influence on CT varies more between hemispheres, and the DC pattern of CT is more intricate. This suggests that CT undergoes unique biological processes during brain development compared to BV and SA[30,49].

White matter tracts evaluated by MO predominantly favored no DC, whereas RD had the highest proportion of full DC traits (**Fig. 3B** for the DC alas of MO and **Fig. S8** for other metrics). The DC patterns in the PCs of MO differed significantly from other metrics: MO displayed 51 no DC traits, whereas all other metrics exhibited fewer than 25 such traits (**Fig. S9**). Notably, all five PCs of MO associated with the anterior limb of the internal capsule, corticospinal tract, and splenium of corpus callosum favored no DC (**Fig. S9D**).

The corticospinal tract stood out with its distinct DC and heritability enrichment patterns. It was unique in consistently favoring no DC across all DTI metrics and PCs. Remarkably, of the seven functional PC traits enriched in h_X^2, four were linked to the corticospinal tract (**Fig. S10**). The h_X^2 of the fourth PC of the corticospinal tract, when assessed by AD, MD and RD, accounted for over 80% of total heritability, resulting in an enrichment ratio close to 20 (**Fig. S5E**). This heightened influence of the X-chromosome-coupled with its comparatively lower total heritability to other tracts-complements existing knowledge about significant sex differences in the development of the corticospinal tract[50,51]. Further evidence suggests that increased X-chromosome dosage can notably decrease the white matter volume of the corticospinal tract[14]. This unique heritability of the corticospinal tract might stem from the additional dosage of X-escapee genes. Comprehensive details about DC and heritability of DTI traits are available in the **Supplementary Results**.

G360 traits that favored no DC were predominantly associated with the two visual networks and the default network (**Fig. 3C**). For example, the mean amplitude of visual networks, as well as the mean connectivity both within and between primary visual and secondary visual networks, strongly favored no DC. Two of the mean connectivity traits showcased enriched heritability on the X-chromosome ($h_X^2 =$ 2.45%~2.86%), where h_X^2 represented over 21% heritability. Yet, when examining tfMRI, the h_X^2 of average functional connectivity linked to primary or secondary visual networks was more restrained, peaking at 1.28%. Two functional connectivity traits within the posterior-multimodal networks were significantly enriched (**Fig. S11**).

For the 1,777 rfMRI ICA traits generated by the whole-brain spatial independent component analysis, we classified the 76 amplitude traits into two groups[40]: within and outside the triple networks (central executive, default mode, and salience networks). Similarly, we classified the 1,701 functional connectivity traits into three groups: within, partially within, and outside the triple networks. We found that DC status was disproportionately distributed between within and outside-triple-network amplitude traits (Fisher's exact test, p-value = 3.43×10^{-5}). Specifically, 77.8% of the traits within the triple

network were in favor of full DC (**Fig. S12A**), while the corresponding proportion was only 32.3% for traits outside the triple network (**Fig. S12B**). However, we did not observe a significant difference in DC among the three groups for the functional connectivity traits (Fisher's exact test, p-value = 0.062). h_X^2 was neither significantly different between within and outside triple-network amplitude traits (Wilcoxon rank sum test, p-value = 0.252 > 0.05/4, **Fig S12C**), nor was it significantly different among within, partially within, and outside triple-network functional connectivity traits (Wilcoxon rank sum test, minimum p-value = 0.016 > 0.05/4, **Fig S12D**). More results on the heritability of fMRI traits are in **Supplementary Results**.

Genetic loci on the X-chromosome associated with complex brain imaging traits.

We carried out XWAS (n = 34,000) for 2,822 complex brain imaging traits on 300,000 SNPs on the X-chromosome (including 289,000 NPR SNPs and 11,000 PAR SNPs after imputation) using PLINK2[52] (www.cog-genomics.org/plink/2.0/). **Table S4** shows a breakdown of the trait-specific sample size and number of SNPs. For NPR SNPs, we determined the best DC model for each trait following the model selection results in the dosage compensation section. Since the equal variance model (the two male hemizygous genotypes coded as 0 and $\sqrt{2}$) is not available for PLINK2, we set the full DC model (males coded as 0 and 2) for traits favoring equal variance. On average, 1.37 more significant SNPs at the genome-wide significant level (5.0×10^{-8}) can be identified for no DC traits than if the full DC model would have been used. For completeness, we also conducted association analyses on the Y-chromosome for all the traits. No significant SNPs were identified at the genome-wide threshold, so we focused on the X-chromosome thereafter.

At the 5×10^{-8} threshold after using wild bootstrap[53] to control for multiple comparisons due to analyzing 2,822 traits (**Supplementary Methods**), we identified 21 unique top NPR SNPs (linkage disequilibrium (LD) $r^2 < 0.1$) in five genomic regions (Xp11.4, Xp21.3, Xq13.1, Xq26.3 and Xq28), associated with 50 different traits (6 SA, 4 BV, 2 AD, 4 FA, 8 RD, 3 MO, 4 MD, and 19 tract-mean), resulting in 50 trait-locus association pairs (**Table S5-6, Fig 4**). Three out of 50 traits favored no DC. We did not identify any significant locus in PARs (including the small ~300kb PAR2 region within Xq28; the rest of Xq28 are in NPR). In general, traits of different modalities did not share the same loci.

Genetic loci were disproportionately enriched in Xq28 (hypergeometric test, p-value = 2.79×10^{-18}) (**Fig. 4**). These loci were associated with the tract-mean traits of the anterior corona radiata, posterior corona radiata, superior longitudinal fasciculus, sagittal stratum, uncinate fasciculus, posterior thalamic radiation, and superior corona radiata. Splenium of the corpus callosum and posterior limb of the internal

capsule were also associated with Xq28, but they were identified through functional PCs of RD and MO, respectively. Unlike other DTI traits that were only pointed to Xq28, the second PC of the body of corpus callosum evaluated by RD was associated with a locus in Xp11.4. Three SA traits: left and right total SA, and left lateral occipital were also associated with loci at the NPR-PAR boundary in Xq28. Moreover, the SA of left and right supramarginal and the SA of rostral middle frontal were linked to variants in Xq26.3. Loci associated with BV traits spanned a broader range, including Xp21.3 (left and right thalamus proper), Xp11.4 (cerebrospinal fluid, CSF), and Xq13.1 (right ventral diencephalon).

By performing association lookups in NHGRI-EBI GWAS catalog[54], 13 trait-locus association pairs were newly identified (**Table 1**). We found evidence that some top SNPs were related to brain structures in previous studies. For instance, rs2272737 was associated with a variety of DTI measures[17]; rs12843772 was associated with BV, cortical areas[17], and brain shape[55].

We verified the XWAS results from two aspects considering the unique features of the X-chromosome. First, we estimated genetic effects separately for males and females in a sex-stratified analysis and then conducted a meta-analysis via Metal (https://genome.sph.umich.edu/wiki/METAL) (Methods). Almost all significant trait-locus pairs in the sex-agnostic analysis were also discovered in the meta-analysis, and three new trait-locus pairs could only be identified through meta-analysis (**Table S7**). Second, we separately performed XWAS on UKB phase 4 subjects with European ancestry (UKBE, n = 4,181), UKB phase 1-4 subjects with South Asian ancestry and Chinese ancestry (UKBSAC, n = 462), and UKB phase 1-4 subjects with African ancestry (UKBA, n = 295) to replicate our discoveries (**Table S8**). Of the 50 top SNPs in discovery and replication, we observed that 100% had effect sizes with the same sign (proportion test against 0.5, p-value = 4.22×10^{-12}) by using UKBE. Out of 50 trait-locus pairs, 11 (22.0%) could be replicated after multiple hypothesis correction at a conservative Bonferroni threshold (p-value < 0.05/50 = 0.001). None could be replicated by using UKBSAC and UKBA at the same threshold, possibly due to the limited sample size. We then conducted a meta-analysis (total n = 4,938) for the separate XWAS results on different ethnic groups, and 22 (44.0%) trait-locus pairs could be replicated. The effect direction of each replicated SNP was consistent across the ethnic groups. Finally, we meta-analyzed the XWAS results of UKB phase 1 to 3 (discovery) and UKB phase 4 (replication) in white subjects (n = 37,772), and 36 more trait-locus pairs were identified, including from new genomic regions of Xp22.2 and Xp11.23 (**Table S9**).

Shared genetic architectures with brain-related disorders and other phenotypes

The identified independent significant SNPs (along with SNPs in LD ($r^2 \geq 0.6$)) suggest genetic connections between the complex brain imaging traits and a spectrum of neurological diseases, psychiatric disorders, and cognitive functions (**Table S10**). The volume of the right ventral diencephalon (known as hypothalamus) was significantly associated with rs2361468 in the genomic region Xq13.1, where it tagged many variants in LD deeply linked to schizophrenia[56-59]. Additionally, rs62606709 in the same locus was related to educational attainment[60] (**Fig. 5A**). Xq13.1 was enriched with variants related to neuroticism[18] as well as to neuropsychiatric disorders, such as schizophrenia[61,62], autism spectrum disorder (ASD)[62,63], bipolar disorder (BD)[62], major depressive disorder (MDD)[62] and Parkinson's disease (PD)[64]. Through further literature review, we found more evidence of connections between the volume of the ventral diencephalon and neuropsychiatric and other disorders. For instance, ventral diencephalon is involved in many pathways found disrupted in schizophrenia[65] and enlargement of ventral diencephalon was observed in patients with schizophrenia[66] and MDD[67]. Ventral diencephalon atrophy was linked to AD[68,69], MDD[70-72], late-life depression[73], PD[74], and spinocerebellar ataxias \ type 3 (SCA3)[75].

We observed colocalization of a locus in Xq28 for tract-mean trait of superior corona radiata (SCR) for MO (termed by "SCR-MO") and various types of testosterone levels (**Fig. 5B**). The top SNP rs67596711 exhibited a positive effect for both SCR-MO and total testosterone level[76]. SCR-MO favored full DC and had relatively small but significant X-linked heritability ($h_X^2 = 0.91\%$) and mild male: female variance ratio (1.04). Strikingly, the phenotypic difference between sexes was pronounced: males registered a 0.44 standard deviation increase in SCR-MO compared to females (p-value = 7.23×10^{-210}), which can likely be attributed to the impact of testosterone. Our findings echo previous research suggesting that the increasing sex disparities in SCR during puberty are influenced by gonadal hormones[51,77,78]. Additionally, BV of CSF shared genetic architecture with sex hormone-binding globulin (SHBG) at Xp11.4 (**Fig. 5C**), indexed by rs35318931, and the effect directions were concordant. Although the gene expression of sex hormone-associated loci were not enriched in any brain cell type[76] in GTEx[79] (v7), it is worth mentioning that the brain tissue samples were sourced exclusively from gray matter. These insights underscore the crucial role of sex hormones in influencing brain functions and structrues.

Apart from brain-related disorders, we also identified colocalization between complex brain imaging traits and various conditions and health-related traits (**Table S10**). In Xq26.3, the genetic loci associated with SA of supramarginal also influenced the onset of myopia[80] and refractive error[81]. In contrast, the genetic loci associated with DTI traits in Xq28 were mainly connected to type 2 diabetes[82-86], blood-related traits including hemoglobin[87], hematocrit[87], red blood cell count[87], serum uric acid levels[88], serum

creatinine levels[88], blood urea nitrogen levels[88], and factor VIII levels[89], and cardiovascular disorders, such as venous thromboembolism[90].

Gene-based analysis and functional mapping

MAGMA[91] (v1.08) was applied to XWAS summary statistics for gene-based analysis. We identified 29 genes in NPR located in eight genomic regions (Xp11.21, Xp22.12, Xq13.1, Xq21,1, Xq21.2, Xq24, Xq26.3 and Xq28), associated with 78 traits using Bonferroni correction considering the effective number of independent traits[92] and the number of protein-coding genes on the X-chromosome ($p < 0.05/230/747 = 2.91 \times 10^{-7}$, Methods, **Table S11**). *CLIC2*, *DUSP9*, *RAB39B*, *TMLHE*, *VBP1*, and *PJA1* were previously seen being associated with regional SA phenotypes[16]. *ZNF275*, *DACH2*, and *VMA21* were identified as being associated with brain connectivity measurement[93] and brain shape[55]. *FAAH2* was associated with neuroticism measurement[18]. Many detected genes were involved in intellectual disability, educational attainment, and neuropsychiatric disorders. For instance, *RENBP*[94], *TKTL1*[58], and *MAP7D2*[58] were for schizophrenia; *USP51* was for educational attainment[5,60]; *DACH2* was for AD[95]. *IRAK1* and *TMEM187* were for Internet addiction disorder[96]. Some genes were linked to subtypes of autism. For instance, *TMLHE* deficiency, leading to a defect in carnitine biosynthesis, was a risk factor for nondysmorphic autism[97,98]. *DACH2* was associated with ASD[99].

We conducted eQTL mapping using significant independent SNPs based on datasets Genotype-Tissue Expression[79] (GTEx v8) and CommonMind Consortium[20] (CMC) (**Table S12**). We identified eQTLs in Xq28 associated with schizophrenia. For example, rs4370701, the eQTL of *FAM3A*, was associated with the left total SA. Mutations in *FAM3A* led to schizophrenia[58], and its gene expression exerting an effect on the cerebellum, cortex, and hypothalamus was regulated by rs4370701. As discussed before, *RENBP*, a risk gene for schizophrenia among Han Chinese[94], is regulated by rs12840700. The variant was proposed to be associated with FA of posterior corona radiata, superior longitudinal fasciculus, and mean FA of all tracts. Moreover, we identified the relationship between rs2361468 in Xq13.1 and the volume of the right ventral diencephalon. This variant also regulated gene expression of *PJA1*, for which the mutations increased the risk of schizophrenia[59] and other neuropsychiatric disorders[62].

We mapped the significant independent SNPs to genes based on functional sequences, resulting in 35 unique genes associated with 50 traits of brain anatomy and white matter microstructures. These genes exhibited diverse association patterns with SNPs that influenced brain measurements, neuropsychiatric disorders as well as cognitive abilities (**Table S13**). For instance, *DCAF8L1* was linked to educational attainment[60] and neuroticism[18], *EFNB1* was linked to schizophrenia[61], Parkinson's disease[64] and

educational attainment[60], while both *RN7SKP31* and *RNU6-985P* were linked to neuroticism[100], cognitive function[101] and educational attainment[60], through SNPs in nearby intergenic regions. Moreover, *SRPX* was tagged by rs35318931 in its exonic region, correlating with CSF volume, which was a risk factor for anorexia nervosa[102].

X-chromosome inactivation (XCI) status[6] was determined for all protein-coding genes identified by FUMA (**Table S14**). Of the 105 genes, 71 (67.6%) were previously reported inactive, 13 (12.4%) were escaping, and 21 (20.0%) were variable. When compared with the reference provided by Tukiainen et al[6], the identified genes were not enriched in any XCI category (hypergeometric test, p-value > 0.05/3).

Long non-coding RNAs (lncRNAs) have merged pivotal regulators. For example, *XIST* and *TIST* not only initiate XCI but also participate in subsequent complex processes[103,104]. We utilized two approaches – eQTL mapping and H-MAGMA[105] to map significant SNPs to lncRNAs. H-MAGMA incorporates a chromatin interaction profile to aggregate SNPs to the nearest genes (Methods, **Table S15**). At a Bonferroni threshold considering the effective number of independent traits[92] and the number of lncRNAs on the X-chromosome ($p < 0.05/230/107 = 2.03 \times 10^{-6}$), SNPs associated with 27 traits were mapped to three lncRNAs in regions Xq26.3, Xp11.21 and Xq28.

Finally, we conducted biological annotation via the DAVID Bioinformatics Database[106] (https://david.ncifcrf.gov/home.jsp) and SynGO (https://syngoportal.org/) on all the prioritized genes identified through functional mapping (**Table S16**). For DAVID, 182 of these genes were cataloged and subsequently analyzed. At the FDR threshold of 0.05, the genes were enriched in transcription elongation (IPR021156 and PIRSF008633) and were associated with diseases such as autism (KW-1269) and intellectual disability (KW-0991). At the nominal significance level (p-value < 0.05), these genes were enriched in biological pathways such as "axon development" (GO:0061564), "neurogenesis" (UP_KW-0964), and "nervous system development" (GO:0007399). For SynGO, 14 genes were uniquely mapped to SynGO annotated genes, and 11 genes have a cellular component annotation (**Fig. S13** and **Table S17**): four postsynaptic, two both pre- and postsynaptic, three presynaptic, and two could not be mapped to any specific compartment.

Mendelian randomization analysis for a causal effect of gene expression on the human brain

Through eQTL mapping, we have linked the significant variants to gene expression levels, but the causal influence of gene expression on the brain traits remained ambiguous. We employed summary statistics-based Mendelian randomization (SMR)[107] to investigate whether XWAS traits could be modulated by

gene expression. We also used the HEIDI test[107] to distinguish pleiotropy of causal variants from linkage (Methods). The eQTL data containing NPR SNPs across 1,639 probes were derived from Sidorenko et al.'s[44] CAGE whole-blood analysis. After controlling for FDR at 0.05 level, 11 genes showed evidence of the causal effect that can control the alteration of regional BV, SA, and white matter tracts, and some of them were also linked to neuropsychiatric diseases and neurodevelopmental disorders (**Table S18**). For instance, Parkinson's disease-linked gene *DNASE1L1*[64] and schizophrenia-linked gene *FAM50A*[58] exerted a causal effect on the SA of the left lateral occipital; nondysmorphic autism-linked gene *TMLHE* exerted a causal effect on the total SA of both hemispheres, the SA of left lateral occipital, and the third PC of AD of the posterior thalamic radiation. In contrast, the expression of *ZNF275* affected the microstructure of the superior corona radiata evaluated by AD and MO. After applying the threshold p-HEIDI > 0.05[107] to screen out linkage from pleiotropy of causal variants, the associations remained the same, which means the tagged SNPs can simultaneously affect gene expression and brain traits. This finding reinforces the genetic interplay between brain imaging traits and neurological disorders.

Disparity of genetic associations between sexes

We observed substantial phenotypic differences between sexes in complex brain imaging traits (**Supplementary results and Table S19**). Specifically, 80.1% of traits showed significant phenotypic differences sexes after controlling an FDR level of 0.05. These gender disparities in the human brain may be influenced by X-chromosome genetic regions uniquely associated with either sex. To verify this hypothesis, we conducted sex-stratified XWAS on males (n = 16,094) and females (n = 17,558) separately. Due to reduced sample sizes post-split, we used a significance threshold of 1.0×10^{-8} for all analyses here, aligning with Bernabeu's[108] approach in a related study. Within the NPR for males, we identified 31 trait-locus pairs from 12 genomic regions spanning 31 traits (**Fig. 6A** and **Table S20**). Notably, 25 out of these 31 traits displayed significant intergender differences. However, for females in NPR, only nine trait-locus pairs from six genomic regions spanning nine traits were recognized (**Fig. 6B**), with each trait manifesting pronounced gender disparities. In the PAR for males, we found two trait-locus pairs in p22.33 related to the third PC of fornix and stria terminalis (FXST) as gauged by FA and an ICA functional connectivity trait tied to default mode and central executive networks. However, no findings merged from the female data. To fully utilize the available data, we further meta-analyzed the sex-stratified XWAS using UKB phase 1-3 and phase 4 subjects with European ancestry (**Table S21**). The final samples consisted of 18,025 for males and 20,054 for females. Male-specific trait-locus pairs in NPR doubled from 31 to 68 at the set threshold, whereas the female dataset only added two more pairs in NPR. In PAR, the male-specific pairs became insignificant, but one new significant pair emerged for females in Xp22.33, linked to an ICA functional connectivity trait concerning the triple networks.

Substantial discrepancies were evident in the association patterns across sexes. Only two genomic regions (Xq26.3 and Xq28) were simultaneously tagged by male and female-specific associations, but the related traits differed (**Fig. 6**). The only shared associated traits between genders in Xq28 were the first PC and mean RD of the anterior corona radiata. Notably, 19 of the 31 male NPR trait-locus pairs (61.3%) overlapped with findings from the sex-agnostic XWAS, but there were only three out of nine (33.3%) trait-locus pairs for females (**Fig. 6**). This suggests that the significant loci in the sex-agnostic XWAS predominantly driven by males. Most traits with significant loci (regardless of sex) favored full DC. The paucity of significant loci in females aligns with the observation that, for full DC-favoring traits, more genetic variance manifests in males than females. XWAS statistical methodologies should also account for heteroskedasticity between genders.

According to NHGRI-EBI GWAS catalog[54], the third PC of posterior thalamic radiation evaluated by MO in males shared genetic underpinnings with cognitive performance in Xp22.2 (rs5934953). Additionally, the mean amplitude of language network (G360) in males had a genetic overlap with SHBG in Xp22.12 (rs7883287, **Fig. S14**).

Variants can have starkly different genetic effects between sexes, and the extent of these differences can vary based on traits and typically, tissues[108,109]. These variations might illuminate the observed disparities in human brain structures across sexes[4,8,32,110] (Methods). We pinpointed nine trait-locus pairs across seven genomic regions (**Fig. S15** and **Table S22**). All the traits displayed significant sex differences, including CT of left rostral middle frontal, PCs of anterior corona radiata evaluated by FA and RD, functional connectivity involving default mode, motor, subcortical cerebellum, central executive, and limbic networks. Notably, all the variants had moderate effect sizes but distinct directions between sexes, and therefore none of the trait-locus pairs overlapped with sex-specific associations. For example, rs12387759 in Xq27.3 revealed marked effect differences between sexes for CT of the left rostral middle frontal. Similarly, rs62589244 in Xp11.4 had divergent effects on the third PC of anterior corona radiata assessed by RD. Other genomic regions included Xp22.33, Xq21.31, Xp22.11, Xq22.1, Xp22.33 and Xq23.

Differences in genetic effects suggest an interaction between the variant and sex. Traditional XWAS which tests for the primary additive effect, can overlook these variants as effects with varying directions might negate each other in a linear model, leading to the "masking of genetic effect"[108]. Through meta-analysis using Stouffer's method[111], we identified variants overlooked in the sex-agnostic XWAS[112] (**Fig.**

S15). Four of the nine trait-locus pairs with varying effects were significant. This highlights the need for XWAS to account for interactions between variant and sex, reflecting XCI uncertainty[43].

Finally, we hypothesized that there might be distinct genetic profiles for subjects that were consistently located at two tails of phenotypic distributions. Separating these subjects by gender, we compared their genetic profiles using Fisher's exact test (Methods). For males, we identified 29 significant genetic loci (p-value $< 1.0 \times 10^{-8}$), but surprisingly, none for females (**Fig. S16** and **Table S23**). These loci spanned the entire X-chromosome, rather than clustering in specific regions. Some coincided with loci from our sex-stratified analysis, such as an Xq28 locus related to CT. This suggests male brain phenotypic variations can be reflected by diverse X-chromosome genetic profiles, whereas other factors might drive variations in females.

DISCUSSION

Genes on the X-chromosome are extensively expressed in the human brain[32]. Evidence shows that mutations in X-linked genes resulting in intellectual disability are about 3.5-fold more than those in autosomal genes[113]. To comprehensively investigate the X-chromosome's influence on brain anatomy, microstructure, and function, we conducted DC and heritability analysis, as well as XWAS for 2,822 complex brain imaging traits. By comparing the three DC models: full DC, no DC, and equal variance in GCTA-GREML[42], we determined the DC status for each trait, which may be linked to the DC behaviors of effective genes in the early development. We proposed that the DC status can be adopted in the model for XWAS, and more significant variants can be identified for no DC traits compared with if a full DC model were used. Our research offers a detailed atlas of DC and an atlas of enrichment of X-linked heritability. And we identified 13 new trait-locus pairs in the NPR at a more reasonable genome-wide threshold accounting for the number of traits relative to the Bonferroni threshold. By further investigation, we found that the human brain measurements shared genetic co-architectures with educational attainment and various brain-related disorders, including autism spectrum disorder, bipolar disorder, major depressive disorder, and Parkinson's disease. Notably, our data reveal sex-specific genetic association patterns, where each gender exhibits unique association signals, and for males, the NPR SNPs on the X-chromosome account for more phenotypic variance.

We discovered potential robust interactions between the X-chromosome and autosomes within the human brain. Traits favoring no DC displayed significantly larger X-linked heritability but diminished heritability from autosomes, compared to full DC traits. Consequently, the X-chromosome accounted for 11.5% of the total heritability for no DC traits, contrasting with only 6.57% for full DC traits (although it

is still greater than the proportion of genomic base pairs located on the NPR of the X-chromosome, ~5%). We postulate that the special feature of no DC traits is attributable to transcriptional, regulatory, and epigenetic processes in brain development[31,33]. For example, sex steroid receptors often signal through epigenetic actions[114]. Several epigenetic mechanisms, such as the levels of DNA methylation and acetylation, are sex-specific in the brain[115]. Recent studies have demonstrated that sex chromosomes can also induce sex differences in somatic gene expression in the absence of hormonal differences[29]. Further evidence indicates that some genes escaping XCI produce proteins that regulate chromatin structures, potentially influencing autosomal gene expression differences[29]. This includes the histone demethylases *UTX* and *KDM5C*[116,117], the histone deacetylase 8[118-120], and the histone acetyltransferase complex subunits male-specific lethal 3[121] and mortality factor 4-like 2[122]. Such *trans*-modifications, which do not alter the nucleotide sequence, might not be reflected in the narrow-sense heritability, accounting for the reduced heritability ascribed to autosomes. In essence, our analysis offers insights into potential XCI escape at the trait level within the human brain.

The Xq28 genomic region was overly represented in association signals as it contains more than 40% of the identified protein-coding genes. A third of white matter tracts were linked to a compact 40kb band segment in Xq28 (152,876,000~152,916,000). Moreover, many intellectual disabilities can be traced back to gene mutations in Xq28. For example, *CLIC2* and *VBP1* were related to the int22h-1/int22h-2-mediated duplication region in Xq28, believed to be a potential contributor to intellectual and developmental disability[123-125]. Loss-of-function mutations in *MECP2* were associated with Rett syndrome[126], which mainly affects brain development in girls. Conversely, in males, these mutations present a spectrum of clinical outcomes, from mild intellectual challenges to severe neonatal encephalopathy, and in some cases, premature death[127]. It is advisable to investigate the genetic co-architectures between the tracts as well as emphasize the exploration of the genetic origins of brain disorders in Xq28 in future studies.

In addition to sex-agnostic analyses, we systematically analyzed sex disparities in human brain characteristics, including phenotype, phenotypic variance, X-linked heritability, and genetic associations. We found that for most traits, males had both greater phenotypic variance and X-linked heritability than females, and more sex-specific associations can be identified in males. These results were concordant with the fact that one of the X-chromosomes is randomly silenced in females in most tissues and cells. We observed that genetic effects of some variants significantly differed between sexes, showing sex and variant interactions. Combined with the colocalization between brain measurements and sex-hormone related traits, such as testosterone levels and SHBG, we postulate that sex hormone plays an essential role

in regulating human brain development and sexual dimorphism. However, verification of the hypothesis needs more data on transcriptome, proteome, and metabolome (termed as "multi-omics"[128]), which is currently barren for the X-chromosome. We earmark this line of inquiry for future research, emphasizing the necessity for both innovative methodologies and robust multi-omics data dedicated to the X-chromosome.

DATA AVAILABILITY

All UK Biobank data utilized in the study was acquired under application 22783. The eQTL summary statistics from the CAGE whole-blood study can be downloaded at https://cnsgenomics.com/content/data. All files to generate annotation used in H-MAGMA can be accessed at https://doi.org/10.5281/zenodo.5503876. The summary statistics generated by association analyses in the current study can be accessed at https://bigkp.org/.

CODE AVAILABILITY

All software and packages used in this study are publicly available. See the URLs and references cited. The code for generating specific results can be provided upon request.

WEB RESOURCES

GCTA (v1.93.2 beta, https://yanglab.westlake.edu.cn/software/gcta/#Overview); PLINK2 (v2.00a3LM, https://www.cog-genomics.org/plink/2.0/); METAL (v2020-05-05, https://genome.sph.umich.edu/wiki/METAL_Documentation); FUMA (v1.4.1, https://fuma.ctglab.nl); MAGMA (v1.08), and ANNOVAR (v2017-07-17) are embedded in FUMA; NHGRI-EBI GWAS Catalog (2023.06, https://www.ebi.ac.uk/gwas/); SMR (https://yanglab.westlake.edu.cn/software/smr/); DAVID Bioinformatics Database (https://david.ncifcrf.gov); SynGO (https://www.syngoportal.org); H-MAGMA repository (https://doi.org/10.5281/zenodo.5503876).

ACKNOWLEDGEMENTS

Research reported in this publication was partially supported by the National Institute On Aging (NIA) of the National Institutes of Health (NIH) under Award Number RF1AG082938 (H.Z. and B.Z) and NIH MH116527 (TF.L. and H.Z.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank the individuals represented in the UKB study for their participation and the research teams for their work in collecting, processing and disseminating these datasets for analysis. We would like to thank University of North Carolina at Chapel Hill and the Research Computing groups for providing computational resources and

support that have contributed to the research results. This research has been conducted using the UK Biobank resource (application number 22783), subject to a data transfer agreement. The UKB has obtained ethics approval from the North West Multi-Centre Research Ethics Committee (MREC, approval number: 11/NW/0382), and obtained written informed consent from all participants prior to the study.

AUTHOR CONTRIBUTIONS

Z.J., B.Z ,and H.Z. designed the study. Z.J. analyzed the data. TF. L., X.W., TY.L., Y.Y., H.S., P.Y.G., J.C. processed the MRI data. L.S. carefully verified all statistical methods used in the study. P.F.S, Y.L., J.L.S., D.L., and H.Z. provided feedback on study design and results interpretations. Z.J. wrote the manuscript and made figures and tables with feedback from all authors.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

1. Uller, T., Pen, I., Wapstra, E., Beukeboom, L.W. & Komdeur, J. The evolution of sex ratios and sex-determining systems. *Trends in Ecology & Evolution* **22**, 292-297 (2007).
2. Lyon, M.F. Gene action in the X-chromosome of the mouse (Mus musculus L.). *nature* **190**, 372-373 (1961).
3. Posynick, B.J. & Brown, C.J. Escape From X-Chromosome Inactivation: An Evolutionary Perspective. *Front Cell Dev Biol* **7**, 241 (2019).
4. Gegenhuber, B. & Tollkuhn, J. Signatures of sex: Sex differences in gene expression in the vertebrate brain. *Wiley Interdisciplinary Reviews: Developmental Biology* **9**, e348 (2020).
5. Lee, J.J. *et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet* **50**, 1112-1121 (2018).
6. Tukiainen, T. *et al.* Landscape of X chromosome inactivation across human tissues. *Nature* **550**, 244-248 (2017).
7. Oliva, M. *et al.* The impact of sex on gene expression across human tissues. *Science* **369**, eaba3066 (2020).
8. Rubinow, D.R. & Schmidt, P.J. Sex differences and the neurobiology of affective disorders. *Neuropsychopharmacology* **44**, 111-128 (2019).
9. Ohno, S. *Sex chromosomes and sex-linked genes*, (Springer Science & Business Media, 2013).
10. Deng, X. *et al.* Evidence for compensatory upregulation of expressed X-linked genes in mammals, Caenorhabditis elegans and Drosophila melanogaster. *Nature genetics* **43**, 1179-1185 (2011).

11. Crowley, J.J. *et al.* Analyses of allele-specific gene expression in highly divergent mouse crosses identifies pervasive allelic imbalance. *Nature genetics* **47**, 353-360 (2015).
12. Lentini, A. *et al.* Elastic dosage compensation by X-chromosome upregulation. *Nature Communications* **13**, 1854 (2022).
13. Disteche, C.M. High expression of the mammalian X chromosome in brain. *Brain research* **1126**, 46-49 (2006).
14. Warling, A. *et al.* Sex Chromosome Dosage Effects on White Matter Structure in the Human Brain. *Cerebral Cortex* **31**, 5339-5353 (2021).
15. Sng, L.M.F., Thomson, P.C. & Trabzuni, D. Genome-wide human brain eQTLs: In-depth analysis and insights using the UKBEC dataset. *Sci Rep* **9**, 19201 (2019).
16. Mallard, T.T. *et al.* X-chromosome influences on neuroanatomical variation in humans. *Nat Neurosci* **24**, 1216-1224 (2021).
17. Smith, S.M. *et al.* An expanded set of genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature neuroscience* **24**, 737-745 (2021).
18. Luciano, M. *et al.* The influence of X chromosome variants on trait neuroticism. *Molecular psychiatry* **26**, 483-491 (2021).
19. West, A.E. & Orlando, V. Epigenetics in brain function. *Neuroscience* **264**, 1-3 (2014).
20. Fromer, M. *et al.* Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nature neuroscience* **19**, 1442-1453 (2016).
21. Schwartz, C.E. *et al.* X-Linked intellectual disability update 2022. *American Journal of Medical Genetics Part A* **191**, 144-159 (2023).
22. Lin, A. *et al.* Mapping the stability of human brain asymmetry across five sex-chromosome aneuploidies. *Journal of Neuroscience* **35**, 140-145 (2015).
23. Hong, D.S. & Reiss, A.L. Cognitive and neurological aspects of sex chromosome aneuploidies. *The Lancet Neurology* **13**, 306-318 (2014).
24. Printzlau, F., Wolstencroft, J. & Skuse, D.H. Cognitive, behavioral, and neural consequences of sex chromosome aneuploidy. *Journal of neuroscience research* **95**, 311-319 (2017).
25. Yurov, Y.B., Vorsanova, S.G., Liehr, T., Kolotii, A.D. & Iourov, I.Y. X chromosome aneuploidy in the Alzheimer's disease brain. *Molecular cytogenetics* **7**, 1-7 (2014).
26. Sánchez, X.C. *et al.* Associations of psychiatric disorders with sex chromosome aneuploidies in the Danish iPSYCH2015 dataset: a case-cohort study. *Lancet Psychiatry* **10**, 129-138 (2023).
27. Green, T., Flash, S. & Reiss, A.L. Sex differences in psychiatric disorders: what we can learn from sex chromosome aneuploidies. *Neuropsychopharmacology* **44**, 9-21 (2019).
28. Shanmugan, S. *et al.* Sex differences in the functional topography of association networks in youth. *Proc Natl Acad Sci U S A* **119**, e2110416119 (2022).
29. Wijchers, P.J. & Festenstein, R.J. Epigenetic regulation of autosomal gene expression by sex chromosomes. *Trends in Genetics* **27**, 132-140 (2011).
30. Kaczkurkin, A.N., Raznahan, A. & Satterthwaite, T.D. Sex differences in the developing brain: insights from multimodal neuroimaging. *Neuropsychopharmacology* **44**, 71-85 (2019).
31. Marrocco, J., Einhorn, N.R. & McEwen, B.S. Environmental epigenetics of sex differences in the brain. *Handbook of Clinical Neurology* **175**, 209-220 (2020).
32. Raznahan, A. & Disteche, C.M. X-chromosome regulation and sex differences in brain anatomy. *Neuroscience & Biobehavioral Reviews* **120**, 28-47 (2021).

33. Li, M. *et al.* Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362**, eaat7615 (2018).
34. Wise, A.L., Gyi, L. & Manolio, T.A. eXclusion: toward integrating the X chromosome in genome-wide association analyses. *The American Journal of Human Genetics* **92**, 643-647 (2013).
35. Sun, L., Wang, Z., Lu, T., Manolio, T.A. & Paterson, A.D. eXclusionarY: 10 years later, where are the sex chromosomes in GWASs? *The American Journal of Human Genetics* **110**, 903-912 (2023).
36. Gao, F. *et al.* XWAS: a software toolset for genetic data analysis and association studies of the X chromosome. *Journal of Heredity* **106**, 666-671 (2015).
37. Glasser, M.F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171-178 (2016).
38. Zhao, B. *et al.* Genome-wide association analysis of 19,629 individuals identifies variants influencing regional brain volumes and refines their genetic co-architecture with cognitive and mental health traits. *Nat Genet* **51**, 1637-1644 (2019).
39. Zhao, B. *et al.* Common genetic variation influencing human white matter microstructure. *Science* **372**(2021).
40. Zhao, B. *et al.* Common variants contribute to intrinsic human brain functional networks. *Nat Genet* **54**, 508-517 (2022).
41. Zhao, B. *et al.* Genetic influences on the intrinsic and extrinsic functional organizations of the cerebral cortex. *medRxiv*, 2021.07.27.21261187 (2021).
42. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76-82 (2011).
43. Chen, B., Craiu, R.V., Strug, L.J. & Sun, L. The X factor: A robust and powerful approach to X-chromosome-inclusive whole-genome association studies. *Genet Epidemiol* **45**, 694-709 (2021).
44. Sidorenko, J. *et al.* The effect of X-linked dosage compensation on complex trait variation. *Nat Commun* **10**, 3009 (2019).
45. Silkaitis, K. & Lemos, B. Sex-biased chromatin and regulatory cross-talk between sex chromosomes, autosomes, and mitochondria. *Biology of Sex Differences* **5**, 2 (2014).
46. Savic, I. & Arver, S. Sex Differences in Cortical Thickness and Their Possible Genetic and Sex Hormonal Underpinnings. *Cerebral Cortex* **24**, 3246-3257 (2013).
47. Zhao, B. *et al.* Heritability of Regional Brain Volumes in Large-Scale Neuroimaging and Genetic Studies. *Cereb Cortex* **29**, 2904-2914 (2019).
48. Hofer, E. *et al.* Genetic correlations and genome-wide associations of cortical structure in general population samples of 22,824 adults. *Nature communications* **11**, 4796 (2020).
49. Grasby, K.L. *et al.* The genetic architecture of the human cerebral cortex. *Science* **367**(2020).
50. Pangelinan, M.M. *et al.* Puberty and testosterone shape the corticospinal tract during male adolescence. *Brain Structure and Function* **221**, 1083-1094 (2016).
51. Ho, T.C. *et al.* Sex differences in the effects of gonadal hormones on white matter microstructure development in adolescence. *Developmental Cognitive Neuroscience* **42**, 100773 (2020).
52. Chang, C.C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, s13742-015-0047-8 (2015).

53. Zhu, H. *et al.* FRATS: Functional Regression Analysis of DTI Tract Statistics. *IEEE Transactions on Medical Imaging* **29**, 1039-1049 (2010).
54. Sollis, E. *et al.* The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res* **51**, D977-d985 (2023).
55. Naqvi, S. *et al.* Shared heritability of human face and brain shape. *Nature genetics* **53**, 830-839 (2021).
56. Pantelis, C. *et al.* Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-427 (2014).
57. Pardiñas, A.F. *et al.* Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature genetics* **50**, 381-389 (2018).
58. Trubetskoy, V. *et al.* Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502-508 (2022).
59. Lam, M. *et al.* Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nature genetics* **51**, 1670-1678 (2019).
60. Okbay, A. *et al.* Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature genetics* **54**, 437-449 (2022).
61. Periyasamy, S. *et al.* Association of schizophrenia risk with disordered niacin metabolism in an Indian genome-wide association study. *JAMA psychiatry* **76**, 1026-1034 (2019).
62. Yao, X. *et al.* Integrative analysis of genome-wide association studies identifies novel loci associated with neuropsychiatric disorders. *Translational psychiatry* **11**, 69 (2021).
63. Al-Sarraj, Y. *et al.* Family-based genome-wide association study of autism spectrum disorder in middle eastern families. *Genes* **12**, 761 (2021).
64. Le Guen, Y. *et al.* Common X-chromosome variants are associated with Parkinson disease risk. *Annals of neurology* **90**, 22-34 (2021).
65. Bernstein, H.G., Keilhoff, G. & Steiner, J. The implications of hypothalamic abnormalities for schizophrenia. *Handb Clin Neurol* **182**, 107-120 (2021).
66. Hodgins, S., Piatosa, M.J. & Schiffer, B. Violence among people with schizophrenia: phenotypes and neurobiology. *Neuroscience of aggression*, 329-368 (2014).
67. Sacchet, M.D., Livermore, E.E., Iglesias, J.E., Glover, G.H. & Gotlib, I.H. Subcortical volumes differentiate major depressive disorder, bipolar disorder, and remitted major depressive disorder. *Journal of psychiatric research* **68**, 91-98 (2015).
68. Loskutova, N., Honea, R.A., Brooks, W.M. & Burns, J.M. Reduced limbic and hypothalamic volumes correlate with bone density in early Alzheimer's disease. *Journal of Alzheimer's Disease* **20**, 313-322 (2010).
69. King, K.S. *et al.* Effect of Leukocyte Telomere Length on Total and Regional Brain Volumes in a Large Population-Based Cohort. *JAMA Neurology* **71**, 1247-1254 (2014).
70. Ancelin, M.-L. *et al.* Lifetime major depression and grey-matter volume. *Journal of psychiatry and neuroscience* **44**, 45-53 (2019).
71. Videbech, P. & Ravnkilde, B. Hippocampal volume and depression: a meta-analysis of MRI studies. *American Journal of Psychiatry* **161**, 1957-1966 (2004).
72. Stratmann, M. *et al.* Insular and hippocampal gray matter volume reductions in patients with major depressive disorder. *PloS one* **9**, e102692 (2014).

73. Geerlings, M.I. & Gerritsen, L. Late-life depression, hippocampal volumes, and hypothalamic-pituitary-adrenal axis regulation: a systematic review and meta-analysis. *Biological psychiatry* **82**, 339-350 (2017).
74. Breen, D.P. *et al.* Hypothalamic volume loss is associated with reduced melatonin output in Parkinson's disease. *Mov Disord* **31**, 1062-6 (2016).
75. Arruda, W.O. *et al.* Volumetric MRI Changes in Spinocerebellar Ataxia (SCA3 and SCA10) Patients. *The Cerebellum* **19**, 536-543 (2020).
76. Ruth, K.S. *et al.* Using human genetics to understand the disease impacts of testosterone in men and women. *Nature medicine* **26**, 252-258 (2020).
77. Bava, S. *et al.* Sex differences in adolescent white matter architecture. *Brain Research* **1375**, 41-48 (2011).
78. Tan, G.C.-Y. *et al.* The influence of microsatellite polymorphisms in sex steroid receptor genes ESR1, ESR2 and AR on sex differences in brain structure. *NeuroImage* **221**, 117087 (2020).
79. Consortium, G. *et al.* The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648-660 (2015).
80. Tedja, M.S. *et al.* Genome-wide association meta-analysis highlights light-induced signaling as a driver for refractive error. *Nature Genetics* **50**, 834-848 (2018).
81. Hysi, P.G. *et al.* Meta-analysis of 542,934 subjects of European ancestry identifies new genes and mechanisms predisposing to refractive error and myopia. *Nature Genetics* **52**, 401-407 (2020).
82. Voight, B.F. *et al.* Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nature Genetics* **42**, 579-589 (2010).
83. Li, H. *et al.* A Genome-Wide Association Study Identifies GRK5 and RASGRP1 as Type 2 Diabetes Loci in Chinese Hans. *Diabetes* **62**, 291-298 (2012).
84. Hara, K. *et al.* Genome-wide association study identifies three novel loci for type 2 diabetes. *Human Molecular Genetics* **23**, 239-246 (2013).
85. Ishigaki, K. *et al.* Large-scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. *Nature Genetics* **52**, 669-679 (2020).
86. Suzuki, K. *et al.* Identification of 28 new susceptibility loci for type 2 diabetes in the Japanese population. *Nature Genetics* **51**, 379-386 (2019).
87. Vuckovic, D. *et al.* The Polygenic and Monogenic Basis of Blood Traits and Diseases. *Cell* **182**, 1214-1231.e11 (2020).
88. Sakaue, S. *et al.* A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet* **53**, 1415-1424 (2021).
89. Sabater-Lleal, M. *et al.* Genome-Wide Association Transethnic Meta-Analyses Identifies Novel Associations Regulating Coagulation Factor VIII and von Willebrand Factor Plasma Levels. *Circulation* **139**, 620-635 (2019).
90. Lindström, S. *et al.* Genomic and transcriptomic association studies identify 16 novel susceptibility loci for venous thromboembolism. *Blood* **134**, 1645-1657 (2019).
91. 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).
92. Wang, H. *et al.* Genotype-by-environment interactions inferred from genetic effects on phenotypic variability in the UK Biobank. *Science advances* **5**, eaaw3538 (2019).

93. Sha, Z., Schijven, D., Fisher, S.E. & Francks, C. Genetic architecture of the white matter connectome of the human brain. *Science Advances* **9**, eadd2870 (2023).
94. Wong, E.H. *et al.* Common variants on Xq28 conferring risk of schizophrenia in Han Chinese. *Schizophrenia bulletin* **40**, 777-786 (2014).
95. Yu, W., Wang, M. & Zhang, Y. Construction of lncRNA-ceRNA Networks to Reveal the Potential Role of Lfng/Notch1 Signaling Pathway in Alzheimer's Disease. *Current Alzheimer Research* (2022).
96. Haghighatfard, A. *et al.* The first genome-wide association study of internet addiction; Revealed substantial shared risk factors with neurodevelopmental psychiatric disorders. *Research in Developmental Disabilities* **133**, 104393 (2023).
97. Celestino-Soper, P.B. *et al.* A common X-linked inborn error of carnitine biosynthesis may be a risk factor for nondysmorphic autism. *Proceedings of the National Academy of Sciences* **109**, 7974-7981 (2012).
98. Beaudet, A.L. Brain carnitine deficiency causes nonsyndromic autism with an extreme male bias: A hypothesis. *Bioessays* **39**, 1700012 (2017).
99. Bralten, J. *et al.* Autism spectrum disorders and autistic traits share genetics and biology. *Molecular Psychiatry* **23**, 1205-1212 (2018).
100. Luciano, M. *et al.* Association analysis in over 329,000 individuals identifies 116 independent variants influencing neuroticism. *Nat Genet* **50**, 6-11 (2018).
101. Homann, J. *et al.* Genome-Wide Association Study of Alzheimer's Disease Brain Imaging Biomarkers and Neuropsychological Phenotypes in the European Medical Information Framework for Alzheimer's Disease Multimodal Biomarker Discovery Dataset. *Frontiers in Aging Neuroscience* **14**(2022).
102. Wade, T.D. *et al.* Genetic variants associated with disordered eating. *International Journal of Eating Disorders* **46**, 594-608 (2013).
103. Engreitz, J.M. *et al.* The Xist lncRNA Exploits Three-Dimensional Genome Architecture to Spread Across the X Chromosome. *Science* **341**, 1237973 (2013).
104. Gayen, S., Maclary, E., Buttigieg, E., Hinten, M. & Kalantry, S. A Primary Role for the Tsix lncRNA in Maintaining Random X-Chromosome Inactivation. *Cell Reports* **11**, 1251-1265 (2015).
105. Sey, N.Y. *et al.* A computational tool (H-MAGMA) for improved prediction of brain-disorder risk genes by incorporating brain chromatin interaction profiles. *Nature Neuroscience* **23**, 583-593 (2020).
106. Sherman, B.T. *et al.* DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res* **50**, W216-w221 (2022).
107. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nature genetics* **48**, 481-487 (2016).
108. Bernabeu, E. *et al.* Sex differences in genetic architecture in the UK Biobank. *Nature genetics* **53**, 1283-1289 (2021).
109. Khramtsova, E.A., Davis, L.K. & Stranger, B.E. The role of sex in the genomics of human complex traits. *Nature Reviews Genetics* **20**, 173-190 (2019).
110. Ritchie, S.J. *et al.* Sex Differences in the Adult Human Brain: Evidence from 5216 UK Biobank Participants. *Cereb Cortex* **28**, 2959-2975 (2018).

111. Stouffer, S.A., Suchman, E.A., DeVinney, L.C., Star, S.A. & Williams Jr, R.M. The american soldier: Adjustment during army life.(studies in social psychology in world war ii), vol. 1. (1949).
112. Lin, B. & Sun, L. Better together against genetic effect heterogeneity: a sex-combined interaction analysis of testosterone levels in the UK Biobank data. *bioRxiv*, 2022.06.08.495202 (2022).
113. Neri, G., Schwartz, C.E., Lubs, H.A. & Stevenson, R.E. X-linked intellectual disability update 2017. *Am J Med Genet A* **176**, 1375-1388 (2018).
114. Tetel, M.J. Nuclear receptor coactivators: essential players for steroid hormone action in the brain and in behaviour. *Journal of neuroendocrinology* **21**, 229-237 (2009).
115. Labonté, B. *et al.* Sex-specific transcriptional signatures in human depression. *Nature medicine* **23**, 1102-1111 (2017).
116. Christensen, J. *et al.* RBP2 belongs to a family of demethylases, specific for tri-and dimethylated lysine 4 on histone 3. *Cell* **128**, 1063-1076 (2007).
117. Iwase, S. *et al.* The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. *Cell* **128**, 1077-1088 (2007).
118. Buggy, J.J. *et al.* Cloning and characterization of a novel human histone deacetylase, HDAC8. *Biochemical Journal* **350**, 199-205 (2000).
119. Hu, E. *et al.* Cloning and characterization of a novel human class I histone deacetylase that functions as a transcription repressor. *Journal of Biological Chemistry* **275**, 15254-15264 (2000).
120. Van den Wyngaert, I. *et al.* Cloning and characterization of human histone deacetylase 8. *FEBS letters* **478**, 77-83 (2000).
121. Smith, E.R. *et al.* A human protein complex homologous to the Drosophila MSL complex is responsible for the majority of histone H4 acetylation at lysine 16. *Molecular and cellular biology* **25**, 9175-9188 (2005).
122. Cai, Y. *et al.* Identification of new subunits of the multiprotein mammalian TRRAP/TIP60-containing histone acetyltransferase complex. *Journal of Biological Chemistry* **278**, 42733-42736 (2003).
123. Andersen, E.F., Baldwin, E.E., Ellingwood, S., Smith, R. & Lamb, A.N. Xq28 duplication overlapping the int22h-1/int22h-2 region and including RAB39B and CLIC2 in a family with intellectual and developmental disability. *American Journal of Medical Genetics Part A* **164**, 1795-1801 (2014).
124. El-Hattab, A.W. *et al.* Clinical characterization of int22h1/int22h2-mediated Xq28 duplication/deletion: new cases and literature review. *BMC medical genetics* **16**, 1-12 (2015).
125. Ballout, R.A. *et al.* Int22h1/Int22h2-mediated Xq28 duplication syndrome: de novo duplications, prenatal diagnoses, and additional phenotypic features. *Human mutation* **41**, 1238-1249 (2020).
126. Van Esch, H. MECP2 Duplication Syndrome. in *GeneReviews(®)* (eds. Adam, M.P. *et al.*) (University of Washington, Seattle
Copyright © 1993-2023, University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle. All rights reserved., Seattle (WA), 1993).
127. Pascual-Alonso, A., Martínez-Monseny, A.F., Xiol, C. & Armstrong, J. MECP2-Related Disorders in Males. *Int J Mol Sci* **22**(2021).

128. Xu, Y. *et al.* An atlas of genetic scores to predict multi-omic traits. *Nature* **616**, 123-131 (2023).
129. Klein, A. & Tourville, J. 101 labeled brain images and a consistent human cortical labeling protocol. *Frontiers in neuroscience* **6**, 171 (2012).
130. Desikan, R.S. *et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968-980 (2006).
131. Jahanshad, N. *et al.* Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: A pilot project of the ENIGMA–DTI working group. *Neuroimage* **81**, 455-469 (2013).
132. Kochunov, P. *et al.* Multi-site study of additive genetic effects on fractional anisotropy of cerebral white matter: comparing meta and megaanalytical approaches for data pooling. *Neuroimage* **95**, 136-150 (2014).
133. Ji, J.L. *et al.* Mapping the human brain's cortical-subcortical functional network organization. *Neuroimage* **185**, 35-57 (2019).
134. Alfaro-Almagro, F. *et al.* Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage* **166**, 400-424 (2018).
135. Elliott, L.T. *et al.* Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature* **562**, 210-216 (2018).
136. Rolls, E.T., Huang, C.-C., Lin, C.-P., Feng, J. & Joliot, M. Automated anatomical labelling atlas 3. *Neuroimage* **206**, 116189 (2020).
137. Finn, E.S. *et al.* Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature neuroscience* **18**, 1664-1671 (2015).
138. Yeo, B.T. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology* (2011).
139. Alfaro-Almagro, F. *et al.* Confound modelling in UK Biobank brain imaging. *NeuroImage* **224**, 117002 (2021).
140. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic acids research* **38**, e164-e164 (2010).
141. Ramasamy, A. *et al.* Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat Neurosci* **17**, 1418-1428 (2014).
142. Schmitt, A.D. *et al.* A compendium of chromatin contact maps reveals spatially active regions in the human genome. *Cell reports* **17**, 2042-2059 (2016).
143. Sey, N.Y., Pratt, B.M. & Won, H. Annotating genetic variants to target genes using H-MAGMA. *Nature Protocols* **18**, 22-35 (2023).

METHODS

Image acquisition and processing

The raw structural MRI (sMRI), diffusion MRI (dMRI), resting-state functional MRI (rfMRI), and task-evoked functional MRI (tfMRI) raw images were acquired from the UK Biobank (http://www.ukbiobank.ac.uk/resources/) with application 22783. Detailed information for the image acquisition is available at https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf. After processing the raw images, 2,822 imaging-derived traits for the human brain were utilized in the study. That is, 230 sMRI traits for cortical structures, 635 diffusion tensor imaging (DTI) traits from dMRI for microstructures of white matter tracts; and 1,957 rfMRI and tfMRI traits for intrinsic and extrinsic brain functions, respectively. For each trait and continuous covariate variable (discussed later), we removed values greater than five times the median absolute deviation from the median value.

We processed the sMRI locally using consistent procedures via advanced normalization tools (ANTs, http://stnava.github.io/ANTs) and conducted multi-atlas cortical parcellation based on the manually edited labels of the publicly available MindBoggle-101 dataset[129]. We removed three ROIs (5th ventricle, left, and right lesion) due to a high missing rate. More details can be found in ref[38]. There were 101 traits for regional BVs, including three global traits – gray matter volume (GMV), white matter volume (WMV), and total BV. We picked up 62 cortical ROIs and generated 62 regional traits for CV as well as the global mean CV. The 66 SA traits used in our analysis were directly downloaded from UKB Category 193. The traits were generated with Freesurfer (https://surfer.nmr.mgh.harvard.edu) by parcellation of the pial surface using Desikan-Killiany[130] parcellation.

DTI evaluated dMRI in a tensor model and analyzed water molecular diffusions in all directions. Five metrics of DTI: Axial diffusivity (AD), fractional anisotropy (FA), mean diffusivity (MD), mode of anisotropy (MO), and radial diffusivity (RD) were applied to each voxel of the image. Given a metric, a tract-mean trait was generated by taking the average of all voxels in a tract. In total, we got 110 tract-mean traits, including 105 tract-mean traits for all tract-metric pairs and 5 overall-mean traits across all tracts (21*5 + 5 = 110). We also applied functional principal component analysis (FPCA) to the voxels in a tract and picked up the top five functional PCs for the tract-metric pair. We generated 525 (= 5*5*21) functional PC traits for all tracts and metrics. The five DTI metrics can reflect different patterns of water diffusion in white matter tracts. For example, AD is the eigenvalue of the principal direction; FA is related to directionality; MD quantifies the magnitude of absolute directionality; MO is the third moment of a tensor; and RD is the average of the eigenvalues of secondary diffusion directions. The tracts were

labeled by the ENIGMA-DTI pipeline[131,132]. Check **Table S1** for the full names of the 21 tracts, and ref[39] for details of trait generation based on ENIGMA-DTI pipeline and FPCA.

We applied parcellation-based methods with Glasser360[37] atlas to generate 90 mean amplitude and functional connectivity traits for both rfMRI and tfMRI. We first projected the rfMRI and tfMRI data on the Glasser360 atlas and generated 360×360 functional connectivity matrices. The 360 functional areas were grouped into 12 functional networks[133]. Then 12 mean amplitude traits and 78 (=12+11×122) mean pairwise functional connectivity traits were extracted from the 12 functional networks. Refer to the Supplementary Note in ref[41] for detailed steps of the parcellation-based dimension reduction procedure. In addition, we locally used the whole brain spatial independent component analysis (ICA) approach to estimate functional brain regions for rfMRI. The detailed procedures were documented in UKB imaging pipeline[134]. We generated 76 node amplitude traits for spontaneous neuronal activity, 1,695 pairwise functional connectivity traits for coactivity for node pairs, and six global connectivity measures for all pairwise functional connectivity[135]. We also manually labeled the 76 node amplitude traits using the automated anatomical labelling atlas[136] (refer to Table S24 in ref[40]) and then mapped them onto major functional networks[137,138]. The assigned location and functional networks are available in **Table S1**.

Discovery and replication data processing for association analyses

In X-chromosome association analysis (XWAS), we analyzed the UKB phase 1-3 imaging data (up to February 2020) for discovery purposes, encompassing 36,000 samples. For replication, we employed the phase 4 imaging data, which included 3,100 samples. We downloaded the version 3 of imputed genetic data from UKB. Details regarding genotyping and imputation are available in the UKB documentation. For data processing, we utilized PLINK2[52] (v2.00a3LM, https://www.cog-genomics.org/plink/2.0/) and treated each imaging trait set individually. Blow is a brief introduction to data preprocessing, for more details please refer to **Supplementary Methods**.

In our discovery data, we prioritized subjects of non-Hispanic white ancestries (Field 21000). Based on UKB-provided quality control details, we filtered out subjects based on specific criteria: excessive heterozygosity (Field ID 22027), inconsistencies between reported and genetic gender (Field ID 22001), potential sex chromosome anomalies (Field ID 22019), and a missing genotype rate exceeding 5%. We further refined our SNP data based on imputation score, minor allele frequency (MAF), and a Hardy-Weinberg equilibrium test. In addition, multiallelic sites were excluded. To address potential relatedness among subjects, we employed GCTA[42] (v1.93.2 beta, https://yanglab.westlake.edu.cn/software/gcta/#Overview). We calculated the genetic relationship matrix

(GRM) for each autosome and merged them. Subjects with a high degree of relatedness (--grm-cutoff 0.05) were pruned; about 1,800 subjects were excluded in this step.

The resulting discovery dataset consisted of 33,591 subjects, with a range of 29,078 to 35,793 across various imaging traits. This included 15,939 males and 17,652 females. For our analyses, 289,866 NPR SNPs and 11,508 PAR SNPs on the X-chromosome were considered. We did not impute the Y-chromosome genetic data, but after the above filters, 140 SNPs remained for analysis.

For replication, the genetic data encompassed UKB phase 4 non-Hispanic white subjects (UKBE, n = 4,181), phases 1-4 South Asian and Chinese subjects (UKBSAC, n = 462), and phases 1-4 African subjects (UKBA, n = 295). To optimize our sample size, we combined white subjects previously excluded from the discovery phase due to relatedness with phase 4 white subjects. Subsequently, we executed another round of relatedness pruning (--grm-cutoff 0.05). Both Asian and African subjects underwent relatedness pruning at the same threshold. All other quality control measures remained consistent with prior steps.

Dosage compensation and heritability analysis

We employed the GREML analysis tool from GCTA[42] for heritability analysis on NPR SNPs. Three distinct model assumptions for GRMs on the X-chromosome are acknowledged: full DC (--dc 1), no DC (--dc 0), and equal variance (without specifying --dc)[42]. These models differ in their coding schemes between sexes, affecting genetic relationships. For instance, females are consistently coded as {0, 1, 2}. The full DC model codes males as {0, 2}, leading to double the genetic variance of females. In contrast, the no DC model codes males as {0, 1}, halving their genetic variance. The equal variance model codes males similarly to females in terms of genetic variance.

We produced GRMs for the X-chromosome under various assumptions to determine the optimal model for each trait. By default, we presumed a consistent allele frequency distribution between causal and genotyped SNPs, adjusting for imperfect LD (--grm-adj 0). Both GRMs for autosomes and the X-chromosome were included in one model to jointly estimate their respective heritabilities, as this method can identify more total heritability compared to analyzing them separately[16].

For sMRI and dMRI traits, we accounted for various predictors like the indicator of phase 3 data (1 if the subject was released in phase3 and 0 otherwise), the UK Biobank assessment center (Field 54), genotype measurement batch (Field 22000), top 40 genetic principal components (Field 22009), age at imaging,

age-squared, sex, age-sex interaction, and age-squared-sex-interaction. Additionally, for sMRI's non-global traits, we adjusted for measurements like total BV, mean CT, and left/right total SA. For fMRI traits, we adjusted for head size (Field ID 25000), scan position X (Field ID 25756), scan position Y (Field ID 25757), scan position Z (Field ID 25758), scan table position (Field ID 25759), mean rfMRI head motion (Field ID 25741), and mean tfMRI head motion (Field ID 25742), as well as scan position X squared, scan position Z squared, mean rfMRI head motion squared and mean tfMRI head motion squared by following Alfaro-Almagro et al.[139].

We employed the likelihood-ratio-test (LRT) with a specific null distribution $0.5\chi_1^2 + 0.5\chi_0^2$ to assess the X-chromosome's variance component. The significance of h_X^2 across traits was determined after adjusting for the false discovery rate (FDR) using the Benjamini-Hochberg procedure at an $\alpha = 0.05$ level. The Akaike information criterion (AIC) was computed to compare the three model assumptions. We selected the model with the smallest AIC as the optimal DC model, subsequently creating an atlas of DC for complex brain imaging traits.

Enrichment analysis for heritability

Building on the methodology of Mallard et al.[16], we characterized enrichment as the ratio of h_X^2 to the proportion of genetic variants on the X-chromosome. Variant counts for each chromosome were sourced from the Genome Reference Consortium Human Build 37 (GRCh37 release 13), accessible at https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.13. Specifically, the measurement for h_X^2 enrichment is $\frac{h_X^2/h_{all}^2}{length_x/length_{all}}$. We then tested if h_X^2 was enriched or depleted by using a two-sided Z-test with statistic $\frac{observed\ h_X^2 - expected\ h_X^2}{standard\ error\ of\ h_X^2}$, where $expected\ h_X^2 = \frac{h_{all}^2 \times length_x}{length_{all}}$, and standard error of h_X^2 was computed by using GCTA. After calculating the p-value from the Z-test, it was adjusted using FDR at the level $\alpha = 0.05$. If the resulting Z-statistic for a trait exceeded 0, it indicated h_X^2 enrichment for that trait; if not, h_X^2 depletion.

Sex-stratified heritability and phenotypic variance analysis

We conducted a sex-stratified heritability analysis to directly contrast the h_X^2 differences between males and females. Within the GREML framework, we accounted for the same covariates as in the sex-agnostic analysis, excluding sex and its related interactions. We adopted the equal variance model in GCTA for all traits. This model presumes males and females to have equivalent heritability. Any deviation of the ratio $\rho_X = \frac{h_{X,male}^2}{h_{X,female}^2}$ from 1 indicates evidence of full DC or no DC. For each DC group, we calculated average

heritability estimates for both sexes and then derived the ratio $\widehat{\rho_X} = \frac{\sum_i \widehat{h^2_{X,i,male}}/p}{\sum_i \widehat{h^2_{X,i,female}}/p}$, where p represents the number of traits with non-zero total heritability in sex-agnostic analysis (a total of 2,810). We used bootstrap methods (employing R's "boot" function over 5,000 iterations) to determine the standard error of $\widehat{\rho_X}$. The GCTA output provides the phenotypic variance, denoted as V_p for both sexes. We subsequently computed the phenotypic variance ratio for each trait as $\frac{V_{p,male}}{V_{p,female}}$.

XWAS, sex-stratified XWAS, and meta-analysis

We conducted linear association tests for NPR, PAR, and Y-chromosome SNPs using PLINK2 (v2.00a3LM). For traits favoring full DC or equal variance, we adopted the full DC model (--xchr-model 2); otherwise, we used the no DC model (--xchr-model 1). The direction of the effect size corresponded to the minor allele in the input data. However, the minor allele for a particular variant might differ between datasets. In the discovery analysis, we adjusted for the same covariates as in the heritability analysis. For replication, adjustments were made for phase 3 and phase 4 indicators, the UK Biobank assessment center, the top 10 genetic PC, and all other imaging-related covariates. We then adjusted the raw p-values using wild bootstrap across all traits (**Supplementary Methods**). SNPs achieving a genome-wide threshold of 5×10^{-8} were considered significant. For LD pruning, we used FUMA (v1.4.1, https://fuma.ctglab.nl/), incorporating both sexes for LD computation. SNPs in LD > 0.6 were grouped under one independent significant SNP. Those in LD > 0.1 were consolidated under a single top SNP. LD blocks defined by adjacent independent significant SNPs within 250kb of each other were merged into one genetic locus. For sex-stratified XWAS, we partitioned males and females in the discovery cohort and carried out separate XWAS for each. The covariate adjustments remained consistent, excluding sex and its interactions.

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

Gene-level analysis and biological annotation

We executed a gene-based association analysis on 747 protein-coding genes on the X-chromosome using MAGMA (v1.08) within FUMA (v1.4.1). We employed GRCh37 to map SNPs to genes by their physical locations, excluding upstream and downstream regions from our consideration. A Bonferroni correction was applied for significance, factoring in both the number of genes and the effective number of independent traits[92], resulting in a threshold of p-value < 0.05/747/230. For the phenotype matrix comprising 2,822 traits, we utilized singular value decomposition (SVD). The effective number was determined using the squared sum of the singular values (s_i) relative to the fourth power of their sum, $\frac{(\sum_{i=1}^{p} s_i^2)^2}{\sum_{i=1}^{p} s_i^4}$. Independent significant SNPs, along with SNPs in LD > 0.6 (including some not in XWAS but from the 1000G dataset), were passed to positional mapping (ANNOVAR[140], version 2017-01- 11), eQTL mapping (with reference database: CommonMind Consortium[20], GTEx v8 brain[79], BRAINEAC[141]), and 3D chromatin interaction mapping (built-in chromatin interaction data: adult cortex, fetal cortex, dorsolateral prefrontal cortex, and hippocampus[142]; annotate enhancer/promoter regions: E053-E082 (brain)). For biological annotation, we utilized the DAVID Bioinformatics Database[106] (https://david.ncifcrf.gov/home.jsp) and SynGO (https://syngoportal.org/). Inputs for this annotation consisted of significant genes identified through the three functional mapping methods, with all other parameters retained as default.

H-MAGMA[105] uses chromatin interaction profiles to map SNPs to the closest genes. We created an annotation file for the X-chromosome according to the protocol[143] and applied MAGMA (v1.08) for H-MAGMA execution. While all necessary files were provided in the protocol, we made minor modifications to the R code to retain only the X-chromosome annotation.

Summary data-based Mendelian randomization (SMR)

The summary statistics of the XWAS and the summary statistics of eQTL analysis using CAGE whole-blood data provided by Sidorenko et al.[44] were utilized in the SMR analysis. The genetic data in XWAS was used as a reference for LD estimation. There were 1,639 probes for genes in the NPR of the X-chromosome, and significant trait-gene pairs were identified by controlling FDR at 0.05 level. The significance indicates that the gene expression level may have a causal effect on the trait. Then we did the HEterogeneity In Dependent Instrument (HEIDI) test to distinguish the pleiotropy of causal SNPs from linkage for the significant trait-gene pairs. A non-significant result at the nominal level (p-value > 0.05) corresponds to no linkage effect and, thus, pleiotropy.

Sex differences in genetic effect

The sex difference in genetic effect of each NPR SNP was tested by a two-sided z-test. The null hypothesis is H_0: $\beta_m = \beta_f$, where β_m and β_f are true per-allele genetic effects for males and females, respectively. We coded males {0, 2} for a full DC trait, and {0, 1} for a no DC trait while always coded females {0, 1, 2}. The test statistic was

$$z = \frac{b_m - b_f}{\sqrt{se_m^2 + se_f^2}},$$

regardless of assumption of DC, where b is a genetic effect estimate, se is the corresponding standard error. Both b and se were adjusted for each sex's trait standard deviation, which was determined by the square root of the phenotypic variance. This was achieved by dividing the original SNP effect size and its standard error by the trait's standard deviation.

Differences in genetic profiles between subjects classified by phenotypic quantiles

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

FIGURE LEGENDS

Figure 1: Overview of the study design. This study encompasses four main components: dosage compensation (DC), association analysis, sex-stratified analysis, and genetic and biological annotation. The unique challenge with X-chromosome analyses arises from the coding scheme uncertainty for males, rooted in the DC in females. We first identified the DC for each trait and established a comprehensive DC atlas for each trait set. Using this atlas, we undertook association analysis across 2,822 complex brain imaging traits. Through genetic and biological annotation, we bridged the association signals with other brain-related disorders and traits linked to sex hormones. In the sex-stratified analysis, we delved into the

disparities in phenotype, phenotypic variance, X-linked heritability, and genetic associations between sexes.

Figure 2: Distribution of DC and patterns of heritability in DC groups and trait sets. A) Distribution of the three DC groups across each trait set. B) Analysis of heritability attributed to the X-chromosome (h_X^2) among DC groups. The p-values are derived from pairwise comparisons through the Wilcoxon rank sum test. C) Examination of heritability contributed by autosomes (h_a^2) among DC groups. The p-values are sourced from pairwise comparisons via the Wilcoxon rank sum test. D)-E) Scatter plots display h_X^2 for individual traits in sMRI, DTI, and fMRI, respectively. Different trait sets are color-coded, with traits having notable h_X^2 represented as filled circles. The average h_X^2 for significant traits is highlighted in its respective color.

Figure 3: Atlases of DC for CT and MO (tract-mean) and fMRI G360 traits. A) DC atlas for CT: the left side showcases the left hemisphere while the right side displays the right hemisphere. Regions of Interest (ROIs) favoring no DC are annotated. B) DC Atlas for tract-mean characteristics assessed by MO, presented in six perspectives. The sequence from left to right, top to bottom includes superior, anterior, left, interior, posterior, and right views. Tracts that favor no DC are highlighted. "SFO" is in parenthesis since it is blocked by other tracts. C) DC Atlas for fMRI G360 traits: The upper triangle illustrates rfMRI, whereas the lower triangle depicts tfMRI. The diagonal, extending from the bottom left to the top right, divides into two sections—the upper triangle portrays DC for rfMRI intra-network connectivity and the lower for tfMRI. Cells on the left margin represent DC amplitude traits for rfMRI, while those at the bottom indicate DC amplitude traits for tfMRI.

Figure 4: Genomic loci associated with complex brain imaging traits identified in sex-agnostic association analysis. Ideogram illustrating the genomic regions affecting brain imaging traits. Each trait is represented by a unique color, with the corresponding genomic region labeled directly on the ideogram. DTI PC traits follow the naming convention "metrics_tract_PC", while DTI trait-mean traits are denoted as "metrics_tract". The term "metrics_Average" refers to the comprehensive average trait for a metric, which is derived from the average of all voxels across all tracts.

Figure 5: Selected genetic loci with important colocalizations in sex-agnostic association analysis. The top lead SNP, defined as the SNP with the lowest p-value in its locus, surpasses the 5×10^{-8} threshold after adjusting for multiple comparisons using wild bootstrap. However, the p-values shown in the figure remain unadjusted. The top lead SNP along with all the SNPs in LD ($r^2 > 0.6$) in the

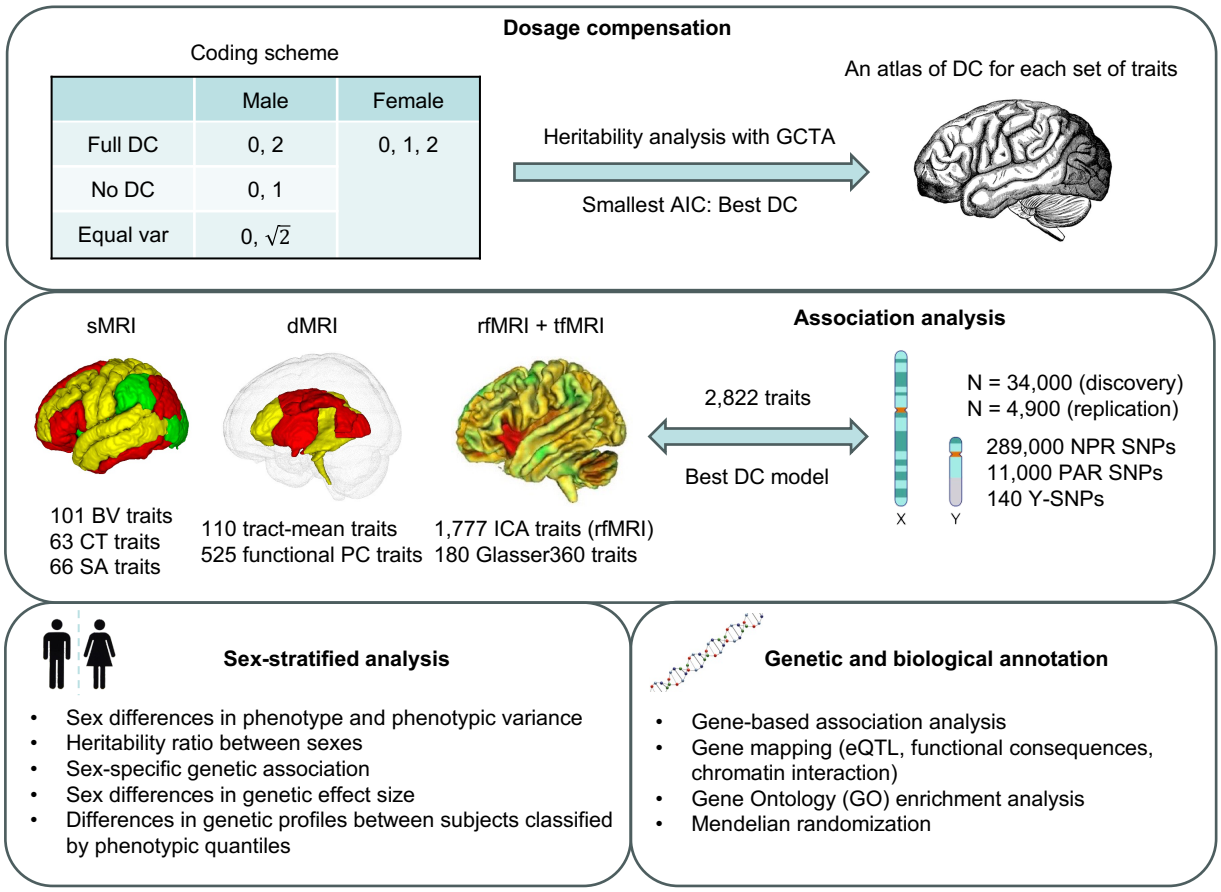
same locus are the targets for XWAS results lookup in the NHGRI-EBI GWAS catalog (2023.06). A) The volume of the right ventral diencephalon exhibits shared genetic links with educational attainment and schizophrenia, pinpointed at an Xq13.1 locus by rs2361468. B) The tract-mean trait of the Superior corona radiata, as assessed by MO, shares a genetic foundation with testosterone levels, identified at an Xq28 locus via rs67596711. C) A shared genetic influence was identified between the volume of cerebrospinal fluid and sex hormone-binding globulin (SHBG), situated at an Xp11.4 locus indexed by rs35318931.

Figure 6: Genomic loci associated with complex brain imaging traits identified in sex-stratified association analysis. Ideogram depicting genomic regions influencing brain imaging traits. Each trait is distinguished by a unique color, with the name of each genomic region clearly marked on the ideogram. Indicators "Yes" and "No" specify whether the same locus was pinpointed in sex-agnostic association analysis. DTI PC traits follow the naming format “metrics_tract_PC”, while DTI trait-mean traits adhere to “metrics_tract”. The term “metrics_Average” signifies an overall-average trait for a metric that encompasses the average across all voxels and tracts. For the naming convention of ICA traits, please refer to Table S1. The label “Language (amplitude)” denotes a rfMRI G360 trait measuring the mean amplitude of the language network. “CO-DA” stands for a rfMRI G360 trait that captures the functional connectivity between the cingulo-opercular and dorsal-attention networks. A) represents male-specific associations, while B) indicates female-specific associations.

Table 1: Newly identified trait-locus pairs for complex brain imaging traits in sex-agnostic analysis. The trait-locus pairs were compared with results on the NHGRI-EBI GWAS catalog (2023.06). The raw p-values are adjusted by using wild bootstrap. The physical locations of loci are according to GRCh37 release 13.

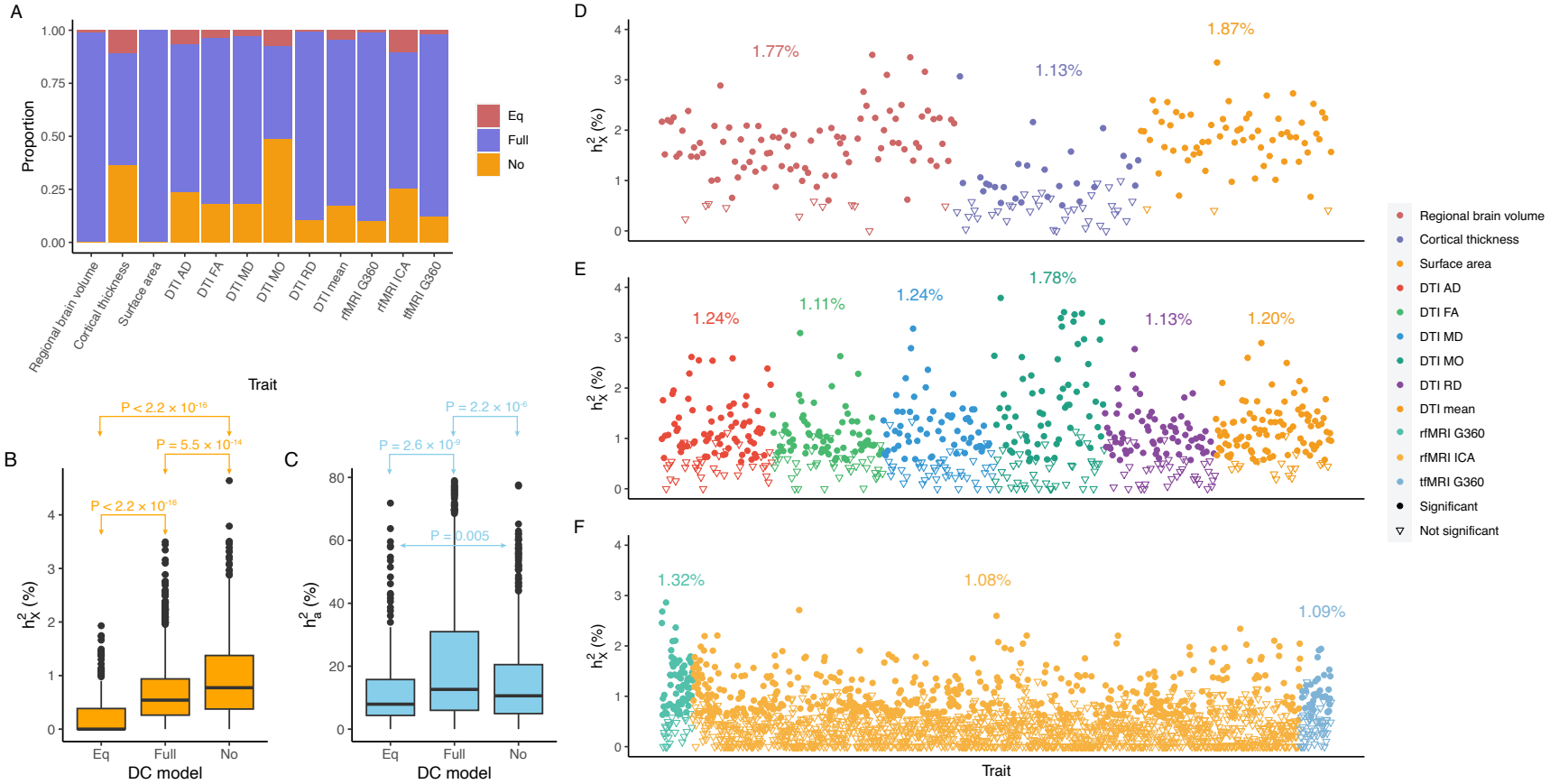
Trait	UniqID	RsID	p-Value	p-Value (adj)	Start	End	Region
BV of cerebrospinal fluid	23:38009121:A:G	rs35318931	6.71E-12	3.82E-10	37933437	38017972	Xp11.4
BV of thalamus proper (left)	23:28140326:A:G	rs2131673	4.69E-14	4.51E-12	28033828	28284470	Xp21.3
BV of thalamus proper (right)	23:28140326:A:G	rs2131673	1.99E-13	2.8E-11	28033828	28284470	Xp21.3
BV of ventral DC (right)	23:68377485:A:G	rs2361468	4.22E-11	4.74E-08	68354618	68384580	Xq13.1
Total SA (left)	23:154740180:A:G	rs76536573	3.09E-13	4.85E-11	153732991	154949953	Xq28
Total SA (right)	23:154678628:C:G	rs28880039	1.13E-11	4.32E-09	153732991	154949953	Xq28
SA of supramarginal (left)	23:136505579:C:T	rs2743914	8.02E-15	4.82E-13	136471150	137073809	Xq26.3
SA of supramarginal (right)	23:136500146:A:G	rs2840674	4.65E-12	1.43E-09	136450688	136841801	Xq26.3
PTR (AD PC3)	23:154740180:A:G	rs76536573	5.57E-13	7.18E-13	154379088	154972517	Xq28
SCR (AD PC1)	23:152590165:C:G	rs5970442	1.41E-12	3.23E-10	152576561	152656246	Xq28
BCC (RD PC2)	23:39464147:C:T	rs7888246	4.66E-12	1.43E-09	39383773	39494923	Xp11.4
PLIC (MO PC4)	23:150542930:G:T	rs201850695	1.1E-13	5.5E-14	150497978	150614462	Xq28
PLIC (MO PC5)	23:150514840:C:T	rs5924909	2.75E-10	1.07E-08	150514840	150614316	Xq28

1 Figure 1


Dosage compensation
Coding scheme
Male
Female
Full DC
0, 2
0, 1, 2
No DC
0, 1
Equal var
0, √2
Heritability analysis with GCTA
Smallest AIC: Best DC
An atlas of DC for each set of traits
sMRI
dMRI
rfMRI + tfMRI
Association analysis
101 BV traits
63 CT traits
66 SA traits
110 tract-mean traits
525 functional PC traits
1,777 ICA traits (rfMRI)
180 Glasser360 traits
2,822 traits
Best DC model
N = 34,000 (discovery)
N = 4,900 (replication)
289,000 NPR SNPs
11,000 PAR SNPs
140 Y-SNPs
X
Y
Sex-stratified analysis
Sex differences in phenotype and phenotypic variance
Heritability ratio between sexes
Sex-specific genetic association
Sex differences in genetic effect size
Differences in genetic profiles between subjects classified by phenotypic quantiles
Genetic and biological annotation
Gene-based association analysis
Gene mapping (eQTL, functional consequences, chromatin interaction)
Gene Ontology (GO) enrichment analysis
Mendelian randomization

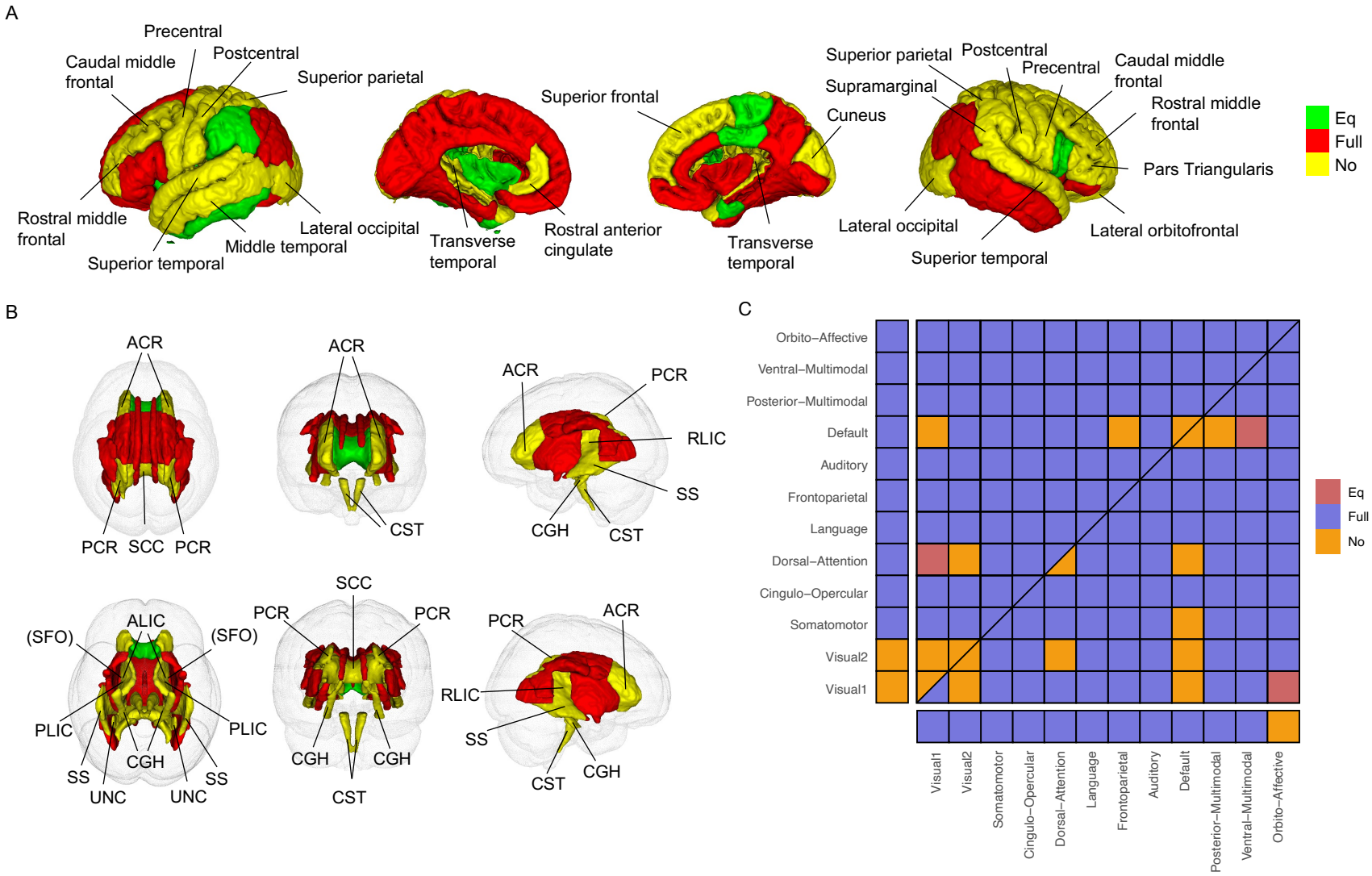
2

1 Figure 2

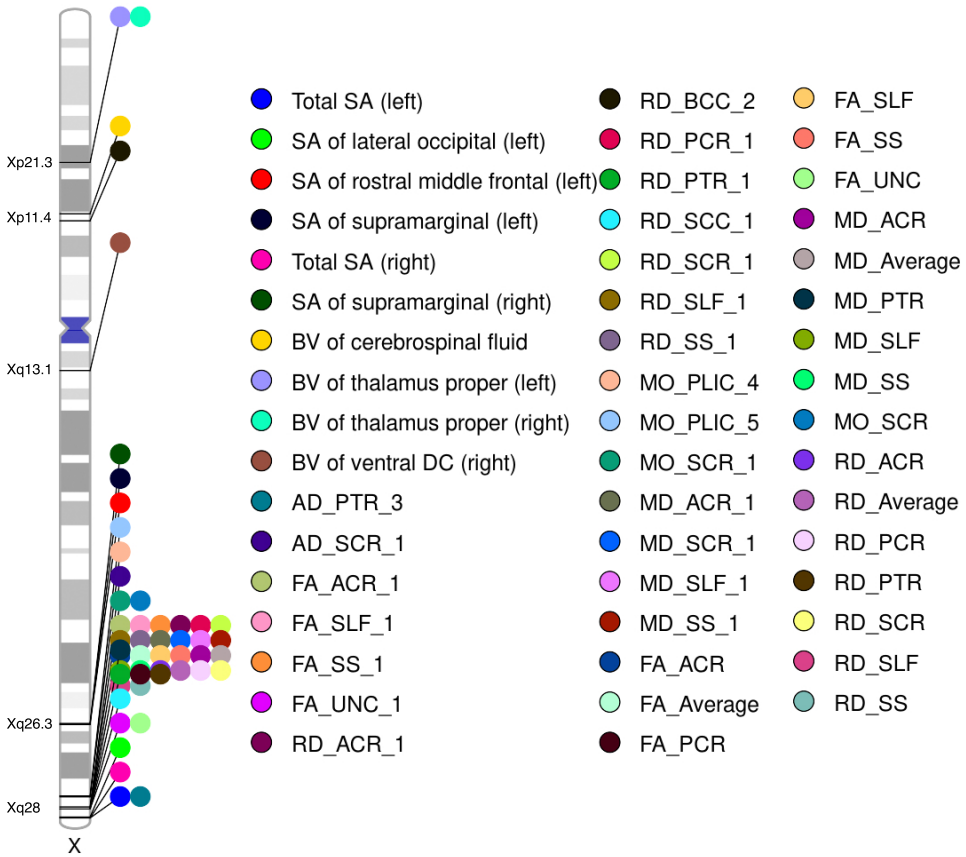

A
Proportion
1.00
0.75
0.50
0.25
0.00
Eq
Full
No
Regional brain volume
Cortical thickness
Surface area
DTI AD
DTI FA
DTI MD
DTI MO
DTI RD
DTI mean
rfMRI G360
rfMRI ICA
tfMRI G360
Trait
B
P < 2.2 × 10-16
P = 5.5 × 10-14
P < 2.2 × 10-16
h2X (%)
Eq
Full
No
DC model
C
P = 2.2 × 10-6
P = 2.6 × 10-9
P = 0.005
h2a (%)
Eq
Full
No
DC model
D
1.77%
1.13%
1.87%
h2X (%)
E
1.24%
1.11%
1.24%
1.78%
1.13%
1.20%
h2X (%)
F
1.32%
1.08%
1.09%
h2X (%)
Trait
Regional brain volume
Cortical thickness
Surface area
DTI AD
DTI FA
DTI MD
DTI MO
DTI RD
DTI mean
rfMRI G360
rfMRI ICA
tfMRI G360
Significant
Not significant

2

Figure 3

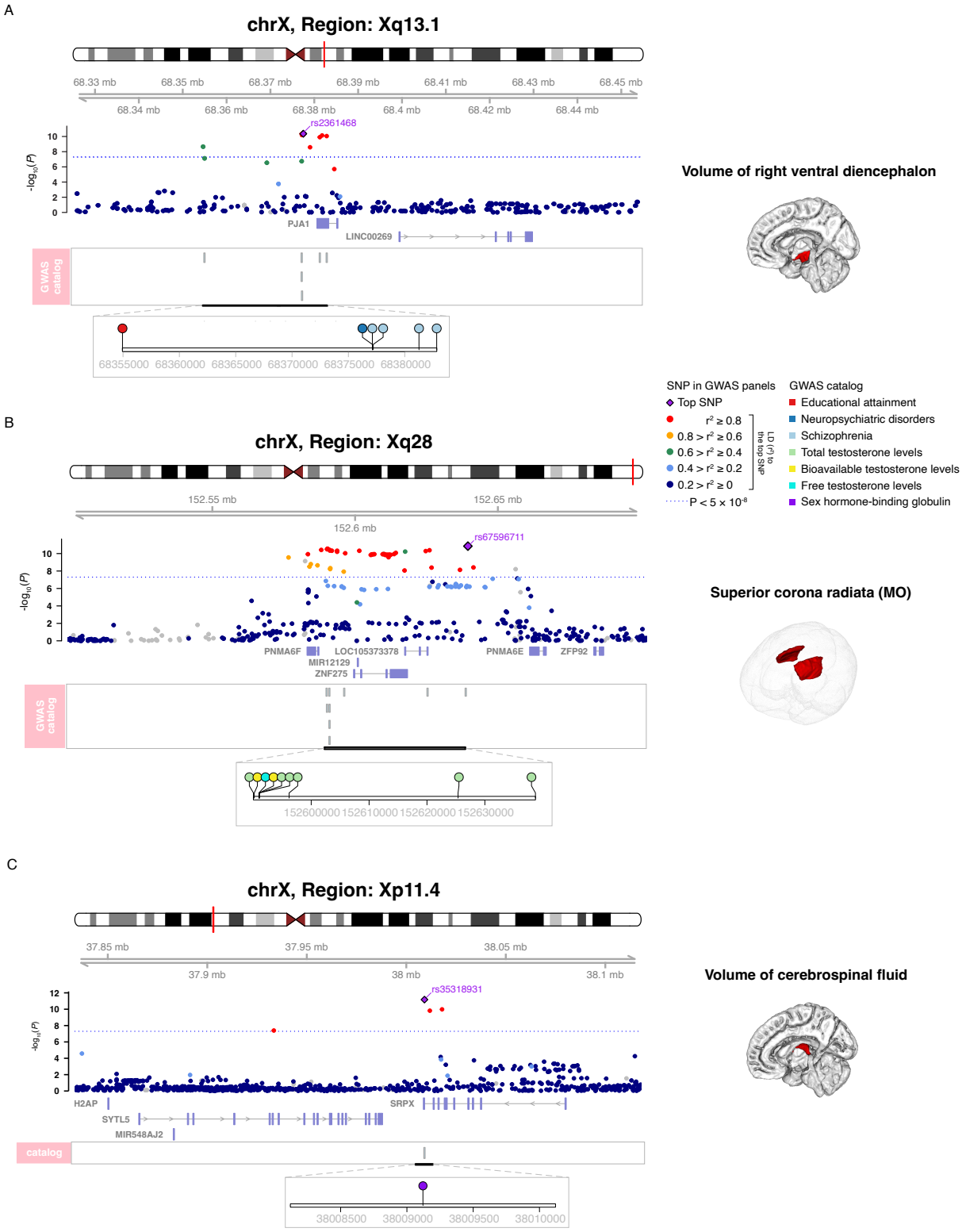

A
Precentral
Caudal middle frontal
Postcentral
Superior parietal
Rostral middle frontal
Superior temporal
Middle temporal
Lateral occipital
Transverse temporal
Rostral anterior cingulate
Superior frontal
Transverse temporal
Superior parietal
Postcentral
Precentral
Caudal middle frontal
Supramarginal
Cuneus
Rostral middle frontal
Pars Triangularis
Lateral occipital
Lateral orbitofrontal
Superior temporal
Eq
Full
No
B
ACR
PCR
SCC
CST
RLIC
SS
CGH
(SFO)
ALIC
PLIC
UNC
C
Orbito−Affective
Ventral−Multimodal
Posterior−Multimodal
Default
Auditory
Frontoparietal
Language
Dorsal−Attention
Cingulo−Opercular
Somatomotor
Visual2
Visual1
Eq
Full
No

1 Figure 4


Xp21.3
Xp11.4
Xq13.1
Xq26.3
Xq28
X
Total SA (left)
SA of lateral occipital (left)
SA of rostral middle frontal (left)
SA of supramarginal (left)
Total SA (right)
SA of supramarginal (right)
BV of cerebrospinal fluid
BV of thalamus proper (left)
BV of thalamus proper (right)
BV of ventral DC (right)
AD_PTR_3
AD_SCR_1
FA_ACR_1
FA_SLF_1
FA_SS_1
FA_UNC_1
RD_ACR_1
RD_BCC_2
RD_PCR_1
RD_PTR_1
RD_SCC_1
RD_SCR_1
RD_SLF_1
RD_SS_1
MO_PLIC_4
MO_PLIC_5
MO_SCR_1
MD_ACR_1
MD_SCR_1
MD_SLF_1
MD_SS_1
FA_ACR
FA_Average
FA_PCR
FA_SLF
FA_SS
FA_UNC
MD_ACR
MD_Average
MD_PTR
MD_SLF
MD_SS
MO_SCR
RD_ACR
RD_Average
RD_PCR
RD_PTR
RD_SCR
RD_SLF
RD_SS

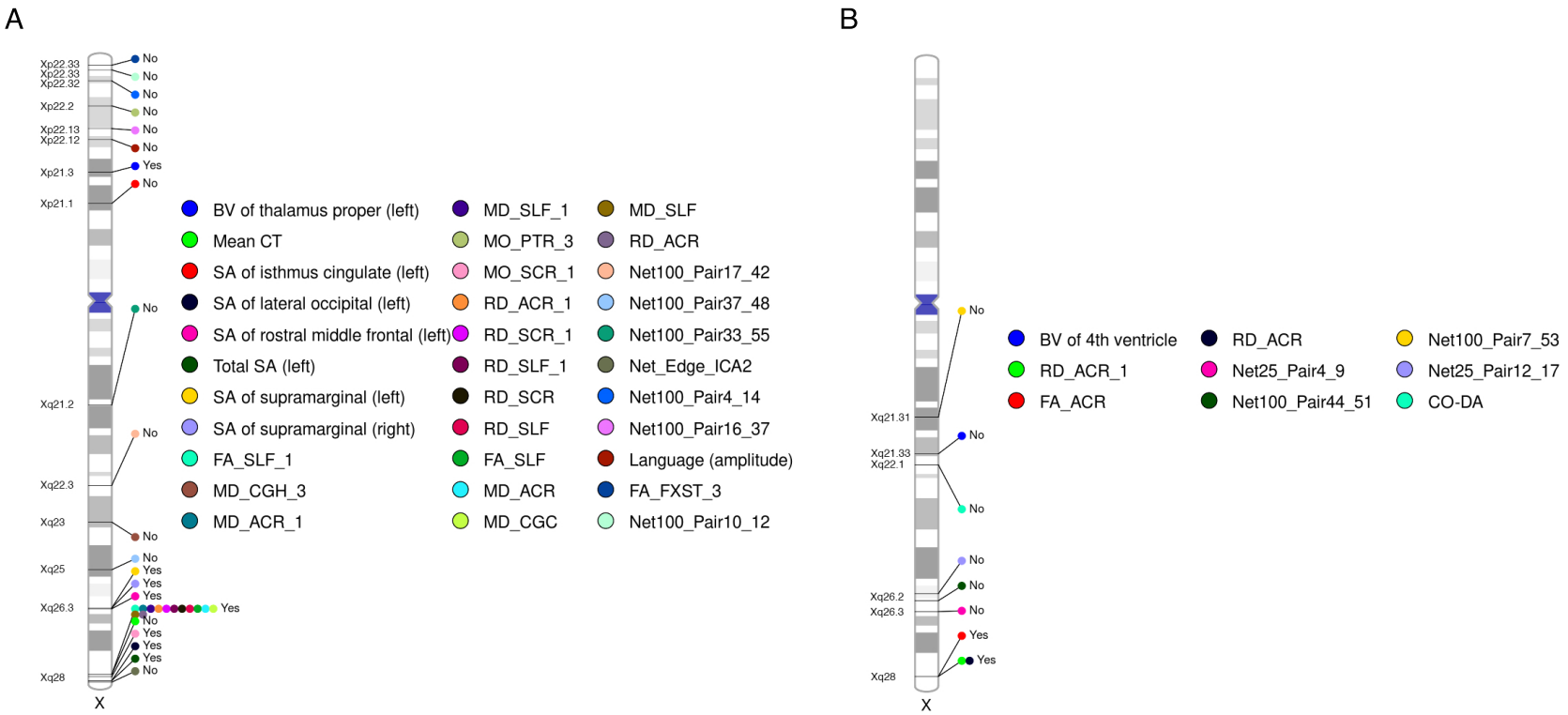
2

1 Figure 5


A
chrX, Region: Xq13.1
68.33 mb
68.35 mb
68.37 mb
68.39 mb
68.41 mb
68.43 mb
68.45 mb
68.34 mb
68.36 mb
68.38 mb
68.4 mb
68.42 mb
68.44 mb
rs2361468
-log10(P)
PJA1
LINC00269
GWAS catalog
68355000 68360000 68365000 68370000 68375000 68380000
Volume of right ventral diencephalon
B
chrX, Region: Xq28
152.55 mb
152.65 mb
152.6 mb
rs67596711
PNMA6F
LOC105373378
MIR12129
ZNF275
PNMA6E
ZFP92
152600000 152610000 152620000 152630000
Superior corona radiata (MO)
C
chrX, Region: Xp11.4
37.85 mb
37.95 mb
38.05 mb
37.9 mb
38 mb
38.1 mb
rs35318931
H2AP
SYTL5
MIR548AJ2
SRPX
catalog
38008500 38009000 38009500 38010000
Volume of cerebrospinal fluid
SNP in GWAS panels
Top SNP
r² ≥ 0.8
0.8 > r² ≥ 0.6
0.6 > r² ≥ 0.4
0.4 > r² ≥ 0.2
0.2 > r² ≥ 0
LD (r²) to the top SNP
P < 5 × 10⁻⁸
GWAS catalog
Educational attainment
Neuropsychiatric disorders
Schizophrenia
Total testosterone levels
Bioavailable testosterone levels
Free testosterone levels
Sex hormone-binding globulin

2

1 Figure 6


A
Xp22.33
Xp22.33
Xp22.32
Xp22.2
Xp22.13
Xp22.12
Xp21.3
Xp21.1
Xq21.2
Xq22.3
Xq23
Xq25
Xq26.3
Xq28
X
No
Yes
BV of thalamus proper (left)
Mean CT
SA of isthmus cingulate (left)
SA of lateral occipital (left)
SA of rostral middle frontal (left)
Total SA (left)
SA of supramarginal (left)
SA of supramarginal (right)
FA_SLF_1
MD_CGH_3
MD_ACR_1
MD_SLF_1
MO_PTR_3
MO_SCR_1
RD_ACR_1
RD_SCR_1
RD_SLF_1
RD_SCR
RD_SLF
FA_SLF
MD_ACR
MD_CGC
MD_SLF
RD_ACR
Net100_Pair17_42
Net100_Pair37_48
Net100_Pair33_55
Net_Edge_ICA2
Net100_Pair4_14
Net100_Pair16_37
Language (amplitude)
FA_FXST_3
Net100_Pair10_12
B
Xq21.31
Xq21.33
Xq22.1
Xq26.2
Xq26.3
Xq28
X
No
Yes
BV of 4th ventricle
RD_ACR_1
FA_ACR
RD_ACR
Net25_Pair4_9
Net100_Pair44_51
Net100_Pair7_53
Net25_Pair12_17
CO-DA

2

3