

Published in final edited form as:

*Nat Genet.* 2023 September 01; 55(9): 1483–1493. doi:10.1038/s41588-023-01475-y.

## Genetic insights into human cortical organization and development through genome-wide analyses of 2,347 neuroimaging phenotypes

Varun Warri<sup>1,2,\*</sup>, Eva-Maria Stauffer<sup>1</sup>, Qin Qin Huang<sup>3</sup>, Emilie M. Wigdor<sup>3</sup>, Eric A.W. Slob<sup>4,5,6</sup>, Jakob Seidlitz<sup>7,8,9</sup>, Lisa Ronan<sup>1</sup>, Sofie Valk<sup>10,11,12</sup>, Travis T. Mallard<sup>13,14</sup>, Andrew D. Grotzinger<sup>15,16</sup>, Rafael Romero-Garcia<sup>1,17</sup>, Simon Baron-Cohen<sup>1,2</sup>, Daniel H. Geschwind<sup>18,19,20,21</sup>, Madeline Lancaster<sup>22</sup>, Graham K. Murray<sup>1,23,24</sup>, Michael J. Gandal<sup>7,8</sup>, Aaron Alexander-Bloch<sup>7,8,9</sup>, Hyejung Won<sup>#25</sup>, Hilary C. Martin<sup>#3</sup>, Edward T. Bullmore<sup>#1</sup>, Richard A.I. Bethlehem<sup>1,2,\*</sup>

<sup>1</sup>Department of Psychiatry, University of Cambridge, Cambridge, UK

<sup>2</sup>Department of Psychology, University of Cambridge, Cambridge, UK

<sup>3</sup>Wellcome Trust Sanger Institute, Hinxton, UK

<sup>4</sup>Medical Research Council Biostatistics Unit, University of Cambridge, Cambridge, UK

<sup>5</sup>Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, the Netherlands

<sup>6</sup>Erasmus University Rotterdam Institute for Behavior and Biology, Erasmus University Rotterdam, Rotterdam, the Netherlands

<sup>7</sup>Department of Child and Adolescent Psychiatry and Behavioral Science, The Children's Hospital of Philadelphia, Philadelphia, PA, USA

This work is licensed under a [BY 4.0 International license](#).

\*Corresponding authors.

### Software used

1. FastGWA: [GCTA | Yang Lab \(westlake.edu.cn\)](#)
2. Plink 1.9 and Plink 1.2: [PLINK 1.9 \(cog-genomics.org\)](#)
3. genomic-SEM: [GitHub - GenomicSEM/GenomicSEM: R-package for structural equation modeling based on GWAS summary data](#)
4. LDSC v1.01: [bulik/ldsc: LD Score Regression \(LDSC\) \(github.com\)](#)
5. Two sample MR 0.4.26: [Two Sample MR functions and interface to MR Base database • TwoSampleMR \(mrcieu.github.io\)](#)
6. MR PRESSO: [GitHub - rondolab/MR-PRESSO: Performs the Mendelian Randomization Pleiotropy RESidual Sum and Outlier \(MR-PRESSO\) method.](#)
7. CAUSE: [Introduction to CAUSE \(jean997.github.io\)](#)
8. MR power calculator: [Online sample size and power calculator for Mendelian randomization with a binary outcome \(shinyapps.io\)](#)
9. MAGMA: <https://ctg.cncr.nl/software/magma>
10. HMAGMA: [GitHub - thewonlab/H-MAGMA](#)
11. GCTB: [GCTB \(cnsgenomics.com\)](#)
12. Polyfun (installed November 2021): [GitHub - omerwe/polyfun: PolyFun \(POLYgenic FUNctionally-informed fine-mapping\)](#)
13. SMR: [SMR | Yang Lab \(westlake.edu.cn\)](#)
14. Hyprcoloc: [GitHub - jrs95/hyprcoloc: Hypothesis prioritisation in multi-trait colocalization](#)
15. PRSes: [GitHub - getian107/PRSes: Polygenic prediction via continuous shrinkage priors](#)
16. R statistics packages: nlme, stats, Psychmeta, lmer, effsize, lsa, ape, factoExtra, igraph,
17. R plotting packages: ggseg, ggplot2, tidyverse, ggpubr, ggstatsplot, ComplexHeatmap, scico, karyoploteR, tidySEM, circlize

<sup>8</sup>Lifespan Brain Institute, The Children's Hospital of Philadelphia and Penn Medicine, Philadelphia, PA, USA

<sup>9</sup>Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA

<sup>10</sup>Institute of Neuroscience and Medicine, Brain & Behaviour (INM-7), Research Centre Jülich, FZ Jülich, Jülich'Germany

<sup>11</sup>Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

<sup>12</sup>Otto Hahn Group Cognitive Neurogenetics, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

<sup>13</sup>Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, USA

<sup>14</sup>Department of Psychiatry, Harvard Medical School, Boston, MA, USA

<sup>15</sup>Department of Psychology and Neuroscience, University of Colorado at Boulder, Boulder, CO, USA

<sup>16</sup>Institute for Behavioral Genetics, University of Colorado at Boulder, Boulder, CO, USA

<sup>17</sup>Instituto de Biomedicina de Sevilla (IBiS) HUVR/CSIC/Universidad de Sevilla/ CIBERSAM, ISCIII, Dpto. de Fisiología Médica y Biofísica

<sup>18</sup>Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA

<sup>19</sup>Program in Neurogenetics, Department of Neurology, University of California, Los Angeles, CA 90095, USA

<sup>20</sup>Center for Autism Research and Treatment, Jane and Terry Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, California, USA

<sup>21</sup>Institute of Precision Health, University of California, Los Angeles, California, USA

<sup>22</sup>MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus, Francis Crick Avenue, Cambridge CB2 0QH, UK

<sup>23</sup>Cambridgeshire and Peterborough NHS Trust, Cambridge, UK

<sup>24</sup>Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia

<sup>25</sup>Department of Genetics and the Neuroscience Center, University of North Carolina at Chapel Hill

# These authors contributed equally to this work.

## Abstract

Our understanding of the genetic architecture of the human cerebral cortex is limited both in terms of the diversity of brain structural phenotypes and the anatomical granularity of their associations with genetic variants. Here, we conducted genome-wide association meta-analysis of 13 structural and diffusion magnetic resonance imaging derived cortical phenotypes, measured

globally and at 180 bilaterally averaged regions in 36,843 individuals from the UK Biobank and the ABCD cohorts. These phenotypes include cortical thickness, surface area, grey matter volume, and measures of folding, neurite density, and water diffusion. We identified 4,349 experiment-wide significant loci associated with global and regional phenotypes. Multiple lines of analyses identified four genetic latent structures and causal relationships between surface area and some measures of cortical folding. These latent structures partly relate to different underlying gene expression trajectories during development and are enriched for different cell types. We also identified differential enrichment for neurodevelopmental and constrained genes and demonstrate that common genetic variants associated with surface area and volume specifically are associated with cephalic disorders. Finally, we identified complex inter-phenotype and inter-regional genetic relationships among the 13 phenotypes which reflect developmental differences among them. These analyses help refine the role of common genetic variants in human cortical development and organisation.

---

The human cerebral cortex is morphologically complex, with extensive inter-individual and inter-regional variation associated with cognition, behaviour, health, development and ageing<sup>1-4</sup>. This variation is partly genetic<sup>5-8</sup>, with several loci identified primarily (although not exclusively) with cortical thickness, surface area, and volume<sup>6,9-12</sup>. Less is known about the common variant genetics (including single nucleotide polymorphisms or SNPs) associated with more complex cortical morphometric phenotypes, such as gyrification or curvature or with microstructural MRI measures of cortical myelination and cyto-architecture. Consequently, the extent of shared genetics across surface area, cortical thickness, volume and the aforementioned MRI phenotypes is unknown. These relatively under-investigated cortical phenotypes may be as important as thickness, volume or other more traditional cortical phenotypes in determining individual differences in cognition, behaviour, and health<sup>13-15</sup>. Second, we still do not fully understand how complex cellular and molecular mechanisms of neurodevelopment give rise to these distinct cortical brain phenotypes. Mapping this may help us better pinpoint genetic mechanisms underlying structural cortical abnormalities. Third, linked to this, the degree to which genes that are constrained (i.e., genes which are intolerant to damaging variants) or genes linked to neurodevelopmental conditions are enriched for common genetic variants for cortical morphology is unknown. Fourth, it is also unclear if common genetic variants contribute to cephalic disorders, although the impact of *de novo* damaging variants has been well documented<sup>16</sup>. Finally, the role of common genetic variants in regional cortical phenotypes and organisation is also unclear. This is important as regional organisation may partly emerge from heterochronous regional differences in gene expression<sup>17</sup>.

To address these questions, we conducted 2,347 genome-wide association studies (GWAS) (for 13 global and 2,334 regional phenotypes that met quality control [see Methods]) of cortical brain morphology in 36,843 individuals from the UK Biobank (UKB)<sup>18</sup> and the Adolescent Brain Cognitive Development (ABCD)<sup>19</sup> cohorts. These included eight cortical macrostructural phenotypes extracted from high resolution anatomical MRI sequences, and five cortical microstructural phenotypes extracted from diffusion MRI sequences, which were estimated both globally and across 180 bilaterally averaged regions based on the Human Connectome Project parcellation scheme<sup>20</sup>. The 13 phenotypes included: cortical

thickness (CT), cortical surface area (SA), grey matter volume (volume), folding index (FI), intrinsic curvature index (ICI), local gyrification index (LGI), mean curvature (MC), gaussian curvature (GC), fractional anisotropy (FA), mean diffusivity (MD), intracellular volume fraction (ICVF, also called neurite density index or NDI), isotropic volume fraction (ISOVF), and orientation dispersion index (ODI) (Figure 1, Methods). We share this GWAS summary statistics openly.

## Genome-wide associations of global cortical phenotypes

We first focused on the common variant genetics of 13 global structural MRI cortical phenotypes, calculated from the total or average of all 180 cortical areas (henceforth “global phenotypes”, Figure 1, Methods). We conducted GWAS for the 13 global phenotypes in the UKB after rigorous quality control and restricting the sample to participants of predominantly European genetic ancestries<sup>1</sup> ( $N_{\max} = 31,977$ ). We identified 314 approximately independent ( $r^2 < 0.1$ , 1000 kb) genome-wide significant ( $p < 5 \times 10^{-8}$ ) loci for the global phenotypes. 81 of these were significant at the more stringent experiment-wide significance threshold (Supplementary Table (ST) 1, Methods). We additionally conducted GWAS for the same 13 global phenotypes in individuals of predominantly European genetic ancestries in ABCD ( $N_{\max} = 4,866$ ). For 237 GWAS loci in UKB for which data were available in ABCD, 204 SNPs (86%) had concordant sign of genetic association ( $p < 0.001$ , two-tailed binomial sign test), compared to 119 (~50%) under the null hypothesis that all effects are noise. Furthermore, 40 (16%) of these SNPs had concordant effect direction and had p-values ( $p$ )  $< 0.01$ , against an expectation of 1.18 (0.5%). 3 had concordant effect direction and  $p < 1E-5$  against an expectation of less than 1 (0.0005%) respectively under the null, and thereby rejecting the null hypothesis. In ABCD, 34 of these SNPs were significant after FDR correction, and 13 after Bonferroni correction. We identified a modest positive correlation of effect size ( $r=0.54$ , 95% confidence interval [CI] 0.45-0.63; Supplementary Figure (SF) 1). Additionally, Linkage Disequilibrium Score (LDSC) based <sup>21</sup> genetic correlations between UKB and ABCD were positive and high (genetic correlations ranged from 0.2 to 1, SF 1, ST 2) for all 13 phenotypes except MD, albeit with wide confidence intervals due to the relatively small size of the ABCD dataset. The robust replicability of the significant loci and high genetic correlation given the age difference between UKB (45-81 yrs at scanning) and ABCD (9-10 yrs at scanning) is notable as brain structure and its genetic influences change over time <sup>22,23</sup>.

Given the observed shared genetics between UKB and ABCD, we conducted inverse-variance weighted meta-analyses <sup>24</sup> to combine the GWAS results across both UKB and ABCD. These meta-analyses identified 367 genome-wide significant loci, of which 89 were significant at an experiment-wide threshold (ST 3). This ranged from 50 genome-wide significant loci (18 experiment-wide significant) for SA to 6 GWAS loci (with 0 experiment-wide significant) for FA (Figure 2), with some SNPs being associated with two or more phenotypes. In total there were 75 independent experiment-wide significant SNPs across all

<sup>1</sup>We use the term “predominantly European genetic ancestries” to reflect ongoing conversations about terms used to refer to genetically homogenous groups. We recognise that the genetically homogenous group that clusters with European samples need not be entirely European or that their genetic ancestries are fully captured.

phenotypes. For all GWAS, the attenuation ratio (Methods) was not statistically different from 0 (ST 4), indicating no inflation in test statistics due to uncontrolled population stratification. All phenotypes had significant SNP heritabilities (LDSC<sup>25</sup>: 0.06 for FA to 0.37 for SA), with higher SNP heritabilities for cortical macrostructural metrics (ST 4) compared to cortical microstructural phenotypes.

For SA and CT, we identified high genetic correlations with a larger GWAS based on a partly overlapping sample from the ENIGMA consortium (SA:  $r_g = 0.91 \pm 0.03$ ; CT:  $r_g = 0.83 \pm 0.04$ )<sup>6</sup>. Notably, despite the smaller sample size of the current GWAS meta-analyses, we identified a higher number of genome-wide significant loci for both SA (50 vs 19) and CT (31 vs 3), and had higher statistical power measured using mean  $\chi^2$  (SA: 1.30 for current GWAS vs 1.23 for ENIGMA; CT: 1.23 for current GWAS vs 1.18 for ENIGMA). The gain in power is likely due to reduced heterogeneity in imaging and genotyping in the current study compared to ENIGMA. All three significant loci for CT and 15 out of the 19 significant loci for SA from ENIGMA were significant in our GWAS with concordant effect directions.

With the exception of CT, SA and volume, GWAS have not been conducted for the other cortical phenotypes. We investigated if any of the 75 independent experiment-wide significant SNPs identified in our GWAS or their proxies ( $r^2 > 0.8$  in CEU or GBR populations) were associated (i.e.,  $p < 5 \times 10^{-8}$ ) with any other neuroimaging phenotype using four databases (Methods). We found that only 11 of these 75 SNPs were not associated with any other neuroimaging phenotypes indicating substantial pleiotropy (ST 5). However, we cannot exclude the possibility that these 11 SNPs are associated with other neuroimaging phenotypes that are not included in the four databases investigated, or that proxies with weaker linkage disequilibrium ( $r^2 < 0.8$ ) may be associated with neuroimaging phenotypes.

## Four clusters explain shared genetics among most of the cortical phenotypes

Previous research has identified both distinct (e.g., between SA and CT) and shared genetics between some (e.g., SA and volume) cortical phenotypes<sup>6,26</sup>. It seems likely that some genetic effects are pleiotropic for multiple cortical phenotypes, but it is unclear how these phenotypes cluster based on shared genetics. To address this question, we estimated bivariate genetic correlations and bivariate phenotypic correlations across all 13 global phenotypes (ST 6; Figure 3A). The overall pattern of correlations between measures was highly similar between the genotypic and phenotypic correlation matrices (Mantel's test,  $r=0.89$ ,  $P=1 \times 10^{-4}$ ), in line with Cheverud's conjecture that phenotypic correlations are representative of genetic correlations<sup>27</sup>. We specifically identified high positive (genetic and phenotypic) correlations between five macrostructural phenotypes - SA, volume, LGI, ICI, and FI. Clustering of the genetic correlation matrix using multiple different methods consistently found that 12 of the 13 phenotypes (excluding only CT) formed four clusters relating to cortical expansion, curvature, water diffusion, and neurite density and orientation (SF 2). For the phenotypic correlation matrix, 11 of the 13 phenotypes formed four clusters, with CT and ICVF clustering separately (SF 2). Overall we identified high similarity between

phenotypic and genetic cluster dendrograms ( $r_{\text{cophenetic}} = 0.76$  (Ward D linkage) to 0.86 (Complete linkage)).

To formalise this exploratory analysis of pleiotropic effects on distinct subsets of global brain phenotypes, we used genomic structural equation modelling (SEM) <sup>28</sup> to identify latent structures among the 13 global phenotypes. After excluding CT due to singleton-clustering and moderate genetic correlations ( $r_g$  between -0.3 and -0.7 with eight of the 12 cortical phenotypes (Figure 3A; SF 2)), exploratory followed by confirmatory factor analyses identified a correlated four-factor model with acceptable fit (CFI = 0.89, SRMR = 0.13; ST7, Figure 3B). The four factors were similar to the four clusters and relate to cortical expansion (factor 1), curvature (factor 2), neurite density and orientation (factor 3), and water diffusion (factor 4). Phenotypic factor analyses produced four similar factors, albeit only after the removal of CT which did not cluster with any phenotypes and ICI which exhibited high cross-loading onto two factors (Supplementary Note 1 [SN], SF 3).

Co-localisation analysis of the experiment-wide significant associations supported the clustering and genomic-SEM analyses. Co-localisation analysis, which tests if genetic variants at a genetic locus are shared between two or more phenotypes, identified 56 co-localised genetic clusters among the global phenotypes (posterior probability of co-localisation > 0.6). The highest number of co-localised loci was for cortical expansion phenotypes, followed by water diffusion, neurite density and orientation phenotypes, and then curvature (ST 8, Figure 3C).

Cluster analysis, genomic-SEM, and co-localisation analysis thus convergently indicate four latent factors, each phenotypically represented by two or more MRI phenotypes. However, there is considerable shared genetics between the latent factors. Furthermore, CT was not included in the latent trait analyses. Given this, we conduct all downstream analysis at the level of individual phenotypes, but for consistency, additionally interpret results through the lens of four latent phenotypes and using the terms macrostructural and microstructural phenotypes (Figure 3D).

## Causal relationships between cortical expansion phenotypes

The genetic correlations between the phenotypes can be due to pleiotropy or causality. Consequently, we employed Mendelian randomisation (MR) <sup>29</sup> to investigate whether the genetic relationships between phenotypes represent causal mechanisms, especially among the five cortical expansion phenotypes. We tested three theories of causation, and corrected for multiple testing. First, consistent with the Radial Unit Hypothesis <sup>30</sup> which suggests that surface area emerges from the number of cortical columns but thickness emerges from the number of cells within a cortical column, we would not expect causal effects between SA and CT. Indeed, we observed no significant evidence for causal association between SA and CT. Second, since volume is geometrically related to SA and estimated by the product of SA and CT, we expected to find a bidirectional causal relationship between SA and volume, and indeed, we found evidence for this. Third, previous research <sup>31–34</sup> suggests that sulco-gyral folding emerges from differential tangential expansion of the cortex, partly due to the heterogeneous cortical distribution of progenitor cells <sup>35,36</sup>, suggesting a causal



relationship of SA on folding (FI, LGI and ICI). Consistent with this, we found robust evidence that genetically predicted SA is associated with an increase in certain measures of folding (FI, LGI, and ICI); but no evidence for reciprocally robust causal effects of folding metrics on SA (ST 9-11, SN 2, SF 4-5). Together, these analyses suggest causal relationships between SA and some measures of folding.

## Global cortical phenotypes show differential developmental and cellular associations

The complex genetic architecture among the 13 global phenotypes likely represents shared and distinct developmental and cellular processes. To better understand this, we first aggregated SNP-based p-values to gene-based p-values using two methods (MAGMA<sup>37</sup> and H-MAGMA<sup>38</sup>), and then investigated if these genes exhibited specific developmental trajectories of gene expression using post-mortem brain tissue data from PsychEncode<sup>39</sup>. We excluded FA due to the small number of genes identified. Genes associated with six of the seven macrostructural phenotypes had high relative expression prenatally, a peak in the late mid-gestation period (~19-22 post-conception weeks), and a decline in gene expression postnatally (Figure 4A). In contrast, the four microstructural phenotypes were associated with genes that had peak expression at birth, followed by a less steep decline, or increased expression postnatally. Inspection of the phenotypes based on the four previous clusters did not identify any additional obvious trends in gene-expression trajectories. Linear mixed effects models confirmed significantly higher prenatal gene expression for most cortical macrostructural phenotypes, and higher postnatal gene expression for most microstructural phenotypes (ST 12).

The differences in developmental trajectories likely reflect different underlying cellular compositions for these phenotypes. We used data from both single-cell RNA seq (scRNA-seq, enrichment using HMAGMA and MAGMA genes) and epigenetic data (LDSC-based<sup>25</sup> enrichment) to identify candidate cell types. Focusing on the developing brain, using scRNA-seq data from psychENCODE<sup>17</sup>, we identified enrichment for intermediate progenitor cells for SA, volume, and FI. (ST 13). To provide further temporal resolution into the developing brain, we investigated enrichment using scRNA-seq data from the first trimester (6-10 post conception weeks [pcw])<sup>40</sup> and scRNA-seq<sup>41</sup> and scATAC-seq<sup>42</sup> data from mid-gestation (marked by neural progenitor expansion)<sup>43-45</sup>. We did not identify any enrichment with cell types in the first trimester (ST 14) but FI, volume, and SA (cortical expansion phenotypes) were enriched for progenitor cells during mid-gestation (ST 15-16, Figure 4B), specifically for progenitor cells in the S-phase and G2-M phases of mitosis. Additionally, CT and MC were enriched for multiple neuronal and glial cell types in both datasets, suggesting that these phenotypes are a composite of multiple cell types.

Considering the postnatal brain, there was no significant enrichment of genes in scRNA-seq data from psychENCODE (ST 17). However, analyses using epigenetic signatures of four broad cell types<sup>46</sup> (neurons, astrocytes, oligodendrocytes, microglia: ST 18) identified enrichment across multiple phenotypes (Figure 4C). For instance, cortical microstructural phenotypes were primarily enriched for epigenetic marks in oligodendrocytes and

astrocytes, but not neurons, consistent with the idea that these phenotypes primarily reflect myelination and related processes<sup>47</sup>. Taken together, these results demonstrate that genes underlying the 13 global phenotypes have different developmental trajectories reflecting specific cellular developmental dynamics.

## Cortical expansion phenotypes are associated with neurodevelopmental conditions

Given the enrichment of several of the global phenotypes with prenatal cellular and developmental processes, we hypothesised that these phenotypes are under negative selection pressure wherein damaging alleles are removed from the population by natural selection. Thus, we modelled the relationship between the minor allele frequency of the SNP and variance in effect size to quantify genome-wide signatures of selection using SBayesS<sup>48</sup>. This suggested that the majority of the cortical macrostructural phenotypes are under significant negative selection (Figure 5A). The selection pressure for cortical microstructural phenotypes was weaker, and with the exception of ISOVF, not statistically significant (ST 19).

We further reasoned that GWAS signals for some of these phenotypes would be enriched for constrained genes (i.e., genes from which damaging variants are removed by natural selection)<sup>49</sup>, genes associated with severe neurodevelopmental conditions<sup>50</sup>, or genes related to microcephaly. Indeed, cortical macrostructural phenotypes were significantly enriched for highly constrained genes ( $pLOUEF < 0.37$ ) and, SA was enriched for genes associated with neurodevelopmental conditions (ST 20, Figure 5B) using both MAGMA and HMAGMA based enrichment. However, we identified no enrichment for genes linked to microcephaly, possibly because (1) several genes associated with microcephaly and other relevant cephalic disorders (e.g., lissencephaly and holoprosencephaly) are yet to be discovered or properly documented, or (2) clinical microcephaly (and macrocephaly) might be genetically distinct from normative variation in brain size.

To investigate the second hypothesis, we tested the effect of polygenic scores for SA, volume and CT on risk of cephalic conditions in a sample of 6,916 individuals with severe neurodevelopmental disorders from the DDD study<sup>51,52</sup>. Polygenic scores for SA and volume, but not CT were associated with both macrocephaly and microcephaly in the expected directions (Figure 5C, ST 21). We obtained consistent results in another cohort (SPARK)<sup>53</sup> consisting of autistic individuals and their siblings, some of whom have macrocephaly or microcephaly ( $N=25,621$ ). Furthermore, in the DDD cohort, polygenic scores for both volume and SA were significantly associated with occipital-frontal circumference standardised for age and sex, in both individuals with and without a genetic diagnosis (Figure 5D). These results demonstrate that SNPs associated with cortical expansion phenotypes are enriched in constrained genes and genes linked to neurodevelopmental conditions, and contribute to cephalic disorders and quantitative variation in occipital-frontal circumference even among individuals with severe neurodevelopmental disorders.



Finally, we also conducted bivariate genetic correlations between the 13 phenotypes and 15 different neurodevelopmental, psychiatric, and cognition-related conditions. After multiple testing correction, we identified significant genetic correlations between several cortical expansion phenotypes and measures of cognition (cognitive aptitude and educational attainment). Additionally FI was negatively genetically correlated with ADHD, and GC was positively genetically correlated with insomnia (ST 22).

## Prioritising candidate genes supports progenitor proliferation in cortical expansion

Given the previous enrichment and polygenic association with neurodevelopmental and cephalic disorders, we were interested in identifying potential causal genes from the GWAS and investigate if these genes are associated with cephalic or neurodevelopmental conditions. Whilst HMAGMA and MAGMA are useful methods to aggregate SNP based information to provide gene-based p-values, they do not necessarily identify either the causal variant or the candidate gene. We thus conducted functionally informed fine-mapping of all experiment-wide significant loci using Polyfun<sup>54</sup> to identify causal variants. For 29 of these loci we were able to finemap to fewer than five credible variants, and for eight, a single credible variant (ST 23). We used nine overlapping methods to identify candidate genes (Methods) and identified 181 candidate genes associated with one of the 13 global phenotypes (ST 24) by at least one method. From this list, we defined prioritised candidate genes if they were supported by at least two experimental methods, leading to 40 different prioritised candidate genes, including 19 in the 17q21.31 region (ST 25). Of these, 29 were identified for cortical expansion phenotypes, four for curvature phenotypes, 13 for neurite density and orientation phenotypes, 14 for water diffusion phenotypes, and 12 for CT, with considerable overlap between the phenotypic domains.

Several genes identified for cortical expansion phenotypes are involved in mitosis, neural progenitor proliferation, and cephalic and neurodevelopmental conditions including *ATR*<sup>55</sup>, *CENPW*<sup>56</sup>, *KANSL1*<sup>57</sup> and *HMGA2*<sup>58–60</sup>. Mutations in *ATR* cause Seckel syndrome, characterised by dwarfism, severe microcephaly, and intellectual disability<sup>55</sup>. *KANSL1* is associated with Koolen-de Vries syndrome, characterised by global developmental delays, and with over 50% of published individuals having microcephaly<sup>57</sup>. Mutations in *HMGA2* lead to macrocephaly and Silver-Russell syndrome<sup>60</sup>. The overlap between fine-mapped genes from common variants and genes implicated through rare variants suggest convergence between rare and common variants. The genes identified for the cortical expansion phenotypes were enriched for the wnt signalling pathway (GO:1904953,  $q = 0.04$ ) which regulates progenitor proliferation and cortical size<sup>61</sup>.

Some genes implicated in CT and neurite density and orientation phenotypes involved in axogenesis and neuronal migration including *VCAN*<sup>62</sup> and *MACF1*, mutations in which cause lissencephaly and defects in neuronal migration and axon guidance<sup>63</sup>. Finally, genes associated with water diffusion phenotypes included *MOBP*, which encodes a structural component of the myelin sheath, the neuronal proline and glycine transporter gene *SLC6A20*, and the lipid gated potassium channel gene *KCNK2*.

## Genetic loci associated with regional cortical phenotypes

To identify genetic influences on regional measures of the 13 neuroimaging modalities, we conducted 2,338 GWAS using regional phenotypes measured for 180 bilaterally averaged regions of the cortex using the Human Connectome Parcellation scheme<sup>20</sup>. We excluded ICI and FI in three regions because of a lack of variance. We did not adjust for global phenotypes as we were interested in identifying genetic loci in the global phenotypes that co-localised with regional phenotypes and because controlling for heritable, highly correlated covariates ( $|r| > \sim 0.5$ ) biases the GWAS and increases the number of false positives<sup>64</sup> (SN 3 and SF 14-15). This additionally limits the utility of the GWAS summary statistics for downstream analyses, for example due to collider bias in MR<sup>64,65</sup>. In total we identified 4,260 experiment-wide significant loci. The highest number were associated with regional SA (1,033) (ST 26). These loci were more likely to contain constrained regions of the genome<sup>66</sup> ( $p = 3.97 \times 10^{-3}$ , one-sided Wilcoxon rank-sum test). This enrichment was driven by loci that were significant for regional cortical expansion phenotypes ( $p = 4.38 \times 10^{-4}$ , one-sided Wilcoxon rank-sum test). The 4,263 loci clustered into 456 semi-independent regions when accounting for LD ( $r^2 > 0.1$ , 1000 kb) agnostic of the neuroimaging phenotype, indicating widespread pleiotropy across the regional measures.

To understand the extent to which these signals reflect genetic influences on the global phenotypes, we used the “GWAS-by-subtraction” method to regress out a latent factor representing genetic variance<sup>67</sup> on global phenotypes for 3,216 of the experiment-wide significant loci (Methods, ST 27). 1,633 (50%) of these loci remained experiment-wide significant, and 3,049 (95%) remained genome-wide significant suggesting that the vast majority of these loci had statistically significant regional effects. In contrast the global genetic latent trait reached experiment-wide significance for 966 of these loci (30%), and 1,499 (46%) reached genome-wide significance, suggesting that as many as half of these loci are also associated with the global genetic latent trait. However, this could be partly by construct, as the global phenotypes in this study are simply the sum of the regional phenotypes. Genome-wide modelling of genetic influences on global and regional factors<sup>68</sup> is beyond the scope of the current study, but can be pursued using summary data made available here.

To further identify shared genetic loci across regional and global phenotypes, we conducted co-localisation analyses across all experiment-wide significant loci (regional and global) for each of the 13 phenotypes separately (ST 28). We identified between 409 (for SA) and 17 (for FA) co-localised clusters, where we use the term “cluster” to refer to a group of phenotypes within one of the 13 neuroimaging modalities that share causal variants in an LD-defined genomic region. The largest cluster was at chr12: 65559695 - 67181144 (12q14.3) comprising the global SA and 156 other regional SA GWAS. This region includes the aforementioned *HMGA2*, associated with Silver-Russell Syndrome<sup>58–60</sup>. For all phenotypes except FA and MD, larger clusters were more likely to include hits in the global GWAS ( $p < 0.05$ , one-sided Wilcoxon rank sum test). However, there were some large clusters that comprised only regional GWAS, suggesting more localised regional effects. Visual inspection of all clusters with cluster size of 30+ GWAS (i.e., clusters based on 30 or more regional GWAS) revealed that topologically closer regions were more likely to have

higher genetic co-localisation (SF 6 - 7). Furthermore, median geodesic distance between regions within a cluster was smaller than the median geodesic distance between regions within and outside a cluster ( $p < 2 \times 10^{-16}$ , wilcoxon rank sum test).

Clusters that included the global GWAS also exhibited broader regional patterns of co-localisation. For example, a locus at chr6:125424383-127540461 which includes *CENPW*, was associated with FI and ICI both globally and in over 30 regions in the superior (dorsal) cortex (SF 8). *CENPW* exhibits regional differences in gene expression in the developing cortex<sup>69</sup>. These analyses demonstrate that SNPs associated with global phenotypes may be associated with only some regional phenotypes.

As with the global features, regional cortical macrostructural phenotypes showed an on average higher heritability compared to regional cortical microstructural phenotypes (SF 9, ST 29) ( $t = -19.4$ ,  $p < 2 \times 10^{-16}$ ,  $F_{(12,2327)} = 420.7$ ). There was however widespread regional variability in SNP heritability of cortical morphology (SF 9). Given this variability, we evaluated if SNP heritability systematically varied across previously established functional (Yeo and Krienen communities)<sup>70</sup> and morphological (Mesulam classes) parcellations<sup>71</sup> of the cortex. Permutation analyses that account for spatial correlation between regions (spin-permutation)<sup>72</sup> revealed that only CT had relatively higher heritability in idiosyncratic sensory areas (Mesulam), and a similar profile was observed for the sensory-motor network (Yeo and Krienen)<sup>70</sup> (ST 30, SF 10). This may reflect better histological and functional demarcation of the sensory-motor regions relative to other regions. Overall, these results suggest limited evidence of SNP heritability for cortical morphology being preferentially larger or smaller in known functional and morphological organisational classes.

Previous research has indicated that asymmetry in some phenotypes across the cortex is modestly genetic<sup>73</sup>. To assess this, we calculated the Asymmetry Index for all regions for all 13 phenotypes (SF 11, ST 31). On average, we identified greater absolute asymmetry for cortical expansion and cortical microstructural phenotypes compared to curvature-related phenotypes or CT. However, SNP heritability of the Asymmetry Index was minimal (SF 11, ST 31), and reached statistical significance for only 21 phenotype-region combinations after multiple testing correction (FDR adjusted p-value  $< 0.05$ ). Taken together, this indicates minimal genetic contribution to asymmetry across the cortex, and is suggestive of high genetic correlation across the hemispheres.

## Differential regional and cross-morphological genetic organisation of the cortex

The high-resolution parcellation scheme used in this study also allowed us to evaluate the protomap hypothesis, which suggests that regional differentiation of the cortex is intrinsically (genetically) determined early in cortical development<sup>30,74</sup>. If this is true, then we would expect regions that are spatially closer to each other to be genetically more similar. Partly supporting this, genetic correlations were moderately correlated with geodesic distances among the 180 regions for each of the 13 phenotypes ( $r = 0.57$  for LGI to 0.13 for ICVF,  $p = 0.001$  for all tests, Mantel test, ST 32).

We further investigated if regional genetic correlations were higher within either functionally similar networks (Yeo and Krienen communities<sup>70</sup>) and morphologically similar classes of laminar differentiation (Mesulam classes<sup>71</sup>). Across multiple phenotypes we identified higher genetic correlations among Mesulam's heteromodal association cortical regions but not in any of the Yeo-Krienen communities<sup>70</sup> (ST 33, SF 12).

To better understand if the 13 phenotypes are similar in their pattern of regional genetic correlations, we calculated cophenetic correlation coefficients amongst all 13 neuroimaging modalities using the regional genetic correlation matrices. Grouping based on cophenetic correlations identified four clusters with similar regional genetic correlation patterns (Cluster 1: SA, volume, and LGI; Cluster 2: all folding measures and CT; Cluster 3: FA and OD; and Cluster 4: MD and ISOVF, Figure 6A). Similar clusters were also observed when using regional phenotypic correlation matrices. These clusters differed from the clusters identified from the global phenotypes in that FI and ICI clustered together with CT and other measures of curvature. This suggests that clusters based on shared genetics of global phenotype moderately overlap with clusters based on regional genetic organisation.

These four clusters were also distinguishable based on their correlation between regional geodesic distances and genetic correlation (Figure 6B). Cluster 1 phenotypes, which relate to progenitor proliferation, had the highest correlation between genetic correlation and geodesic distance between regions. This was followed by Cluster 4 (MD and ISOVF: the water diffusion phenotypes), which both increase with age in adults<sup>75,76</sup>. We speculate that this patterning might reflect the heterochronous cellular and developmental trajectories of these phenotypes: regional differences in gene expression in the cortex exhibit a cup-shaped pattern with high regional differences in midgestation that re-emerge during adolescence and increase in adulthood<sup>17,77</sup>.

To explore further how the 13 phenotypes differed in their regional organisation, we extracted the first principal component from their respective genetic correlation matrices. The first principal component explained between 25% (LGI) to 62% (MD) of the variance. The clustering of the neuroimaging modalities based on the similarity of the first principal component of the region-to-region similarity was similar to the clustering based on the cophenetic correlations of the same region-to-region similarity (Figure 6C), suggesting that the first principal component largely captures regional genetic organisation. Visual inspection of the first principal component identified four different axes of variation: anterior-posterior (SA, volume, and LGI: Cluster 1 phenotypes), inferior-superior (ISOVF, MD: Cluster 4 phenotypes), and a mix of primary-association and inferior superior (CT, GC, MC, FI, ICI: Cluster 2 phenotypes, and ICVF) (Figure 6D). For OD and FA (Cluster 3 phenotypes), we were unable to identify a clear topological axis of variation. These are in line with patterns of gene expression in the human cortex: anterior-posterior gradients during development, and primary-association gradients postnatally up until adolescence and early adulthood<sup>78</sup>, and in the inferior-superior direction for water diffusion phenotypes, which are late-emerging<sup>79,80</sup>. Using the first principal component derived from regional phenotypic correlation, we identified clear axes of variation for SA, Volume (anterior-posterior) and LGI, CT (inferior superior), but not for the other phenotypes (SF 13). This likely reflects

the additional influence of directionless non-genetic factors in the development of cortical microstructure and curvature.

## Discussion

Our results provide granular insights into the organisation and development of the human cortex and links to cephalic and neurodevelopmental conditions, after testing several different hypotheses regarding cortical development, organisation, and links to neurodevelopment (ST 34). We find that cortical macrostructural and microstructural phenotypes are genetically distinct, and are enriched for different cellular and developmental processes. These results suggest that SA, volume, and related measures of gyrification emerge from the earliest cellular processes in cortical development, related to progenitor proliferation, which was supported by fine-mapping analyses. Cortical thickness and some measures of curvature are influenced by multiple, subsequent and sequential cellular processes, as indicated by widespread enrichment for different cell types. For the first time, we investigate causal relationships between cortical macrostructural phenotypes using Mendelian randomisation, and provide support for the differential tangential expansion hypothesis<sup>31,34,35</sup>, suggesting that some measures of cortical folding emerge from differential expansion of the cortex.

The cortical expansion phenotypes are under purifying selection, and consequently, enriched for highly constrained genes and genes associated with neurodevelopmental conditions. Furthermore, even among individuals with severe developmental disorders<sup>51,52</sup>, common genetic variants for volume and surface area contribute to cephalic disorders, expanding our understanding of the role of common and rare genetic variants in developmental disorders. However, we did not observe significant genetic correlations between many other psychiatric conditions and global cortical phenotypes, suggesting a more complex relationship between cortical structure and psychiatric conditions. For microstructural phenotypes we did not observe any significant negative selection pressure. This weaker negative signatures for the microstructural phenotypes may be due to a combination of differences in underlying biology, lower statistical power, less well-defined phenotypes, and differences in genetic architecture between cortical microstructural measures and similar measures obtained from white matter fibre tracts.

Genome-wide association analyses of regional phenotypes identified 4,260 additional associations with regional phenotypes. Conditional analyses indicated that the majority of these associations were not driven by global phenotypes. However, topologically closer regions were likely to share genetic loci, suggesting that regional effects are not constrained to parcellation boundaries. Furthermore, in support of the protomap hypothesis<sup>30,74</sup>, topologically closer regions were more likely to be genetically similar. However, we observed no enrichment for either SNP heritability or genetic correlations in known functional and morphological organisational classes of the cortex.

Probing the regional genetic organisation of these phenotypes further, cortical expansion phenotypes that derive from early developmental processes show the greatest shared genetics between topologically nearer regions, in line with findings from cortical gene

expression<sup>78</sup>. However, whilst this is in the anterior-posterior direction for the early emerging phenotypes, it is in the inferior-superior direction for the late emerging phenotypes. Notably, the anterior-posterior gradient is also observed using cell lineage-tracing in the brain<sup>81</sup>. Taken together, these analyses suggest that cortical organisation is informed by distinct waves of molecular processes, some of which are highly directional.

Our analyses focused on individuals predominantly of European genetic ancestries, as we were limited by sample size, computational power, and methodology. With ongoing neuroimaging in the UK Biobank and other cohorts, these analyses should be expanded to include genetically diverse populations. Additionally, the current study investigates only the role of common genetic variants, and the extent to which rare genetic variants contribute to normative differences in cortical morphology is unclear. Somewhat analogously the current study focuses on imaging phenotypes that have been commonly used in literature but does not claim to provide an exhaustive overview. In light of the present finding on microstructure it might be specifically interesting for future studies to take a similar approach to more myelin specific phenotypes such as those that might be derived from T1/T2w imaging<sup>82</sup>. It should also be noted that there is considerable heterogeneity in MRI pre- and post-processing approaches, including the application of parcellation schemes<sup>83,84</sup>. We chose a commonly used approach to increase compatibility of our summary statistics with existing datasets and note that alternative approaches produced similar results, but do not provide an exhaustive evaluation of different strategies in this work. Finally, expanding the number of phenotypes such as functional MRI, white-matter, and subcortical measures will provide a more precise atlas of the genetics of structure and function of the human brain, and the genetic relationships between them.

In conclusion, by conducting and analysing GWAS of 13 different neuroimaging modalities both globally and across 180 cortical regions we provide unprecedented insights into the genetic organisation, and development of the human cortex. We make this resource freely available to researchers for further analyses.

## Methods

### Datasets

**UK BioBank**—The UK Biobank is a prospective cohort of 500,000 individuals from the UK. Of these individuals, 100,000 will undergo brain scanning<sup>5,18,85</sup>, with approximately 40,000 scans having been completed when the current study commenced. Participants were excluded from the MRI study on the basis of standard MRI safety criteria such as metal implants, recent surgery, or conditions problematic for scanning such as hearing problems, breathing problems, or claustrophobia. Ethical procedures for the UK Biobank are controlled by the Ethics and Guidance council (<http://www.ukbiobank.ac.uk/ethics>), and the study was conducted in accordance with the UK Biobank an Ethics and Governance Framework document (<http://www.ukbiobank.ac.uk/wp-content/uploads/2011/05/EGF20082.pdf>), with institutional review board approval by the North West Multi-center Research Ethics Committee.



**ABCD**—The Adolescent Brain Cognitive Development (ABCD) study is an ongoing study of childhood and adolescence<sup>86</sup>. Participants from the general population were recruited from all over the United States across 21 sites by providing select schools with information packets to all families with 8-10 year old students. Ethical approval was obtained from multiple institutional review boards.

## Image acquisition

Data was acquired as part of the UK Biobank and ABCD cohort studies with the following protocols.

### UK Biobank<sup>2</sup>

- T1-weighted structural imaging: 1.0mm isotropic resolution, TR = 2000ms, TE = 2.01ms, TI = 880ms Flip angle 8 degrees,
- T2-weighted FLAIR structural imaging: 1.0x1.0x1.1mm resolution, TR = 5000ms, TE = 395.0 ms, TI = 1800 ms
- Diffusion weighted imaging - 2.0mm isotropic resolution, MB = 3, R = 1, TE/TR = 92/3600 ms, PF 6/8, fat sat, b = 0 s/mm<sup>2</sup> (5x + 3x phase-encoding reversed), b = 1 000 s/mm<sup>2</sup> (50x), b = 2000 s/mm<sup>2</sup> (50x), 105 + 6 time-points (PA-AP)

### ABCD<sup>3</sup>

- T1-weighted - 1.0mm isotropic resolution, TR = 2500ms, TE = 2.88ms, TI = 1060ms, Flip angle 8 degrees
- T2-weighted - 1.0mm isotropic resolution, TR = 3200ms, TE = 565ms, Flip angle variable
- DWI - 1.7mm isotropic resolution, TR = 4100ms, TE = 88ms, Flip angle 90 degrees, 500 (6-dirs); 1000; (15-dirs) 2000; (15-dirs); 3000 (60-dirs)

While not processed as part of the present analysis we also obtained framewise displacement parameters from each individual's accompanying resting-state fMRI scan.

## Image processing

Structural minimally processed T1 and T2-FLAIR weighted data was obtained from UK BioBank (application 20904) and the ABCD study (via the NIH Data Archive Repository). These images were preprocessed with FreeSurfer (v6.0.1)<sup>87</sup> using the T2-FLAIR weighted image to improve pial surface reconstruction when available. Recon-all reconstruction included bias field correction, registration to stereotaxic space, intensity normalisation, skull-stripping, and white matter segmentation. A triangular surface tessellation fitted a deformable mesh model onto the white matter volume, providing grey-white and pial surfaces with >160,000 corresponding vertices registered to fsaverage standard space. When no T2-FLAIR image was available FreeSurfer reconstruction was done using the

<sup>2</sup> [https://www.fmrib.ox.ac.uk/ukbiobank/protocol/V4\\_23092014.pdf](https://www.fmrib.ox.ac.uk/ukbiobank/protocol/V4_23092014.pdf)

<sup>3</sup> [https://github.com/nihf-fmrib/abcd\\_protocols](https://github.com/nihf-fmrib/abcd_protocols) and see also Casey et al. 2018.

T1 weighted image only. Given systematic variation related to the inclusion of T2-FLAIR, this was included as a confound variable in downstream analyses. Cortical surfaces were reconstructed for each individual using FreeSurfer and registered using FreeSurfer's surface-based registration to fsaverage. The HCP's multi-modal parcellation v1.0<sup>20</sup> was resampled from fs\_LR to fsaverage using existing transformations<sup>88</sup> and from there back to the individual's surface meshes based on the FreeSurfer folding-based surface registration. Reconstruction quality was assessed using the Euler index<sup>89</sup> and included as a covariate in subsequent analyses (SN 4, SF 16).

Structural diffusion weighted imaging was obtained in processed form from UK BioBank and ABCD in similar fashion. As described in the UK Biobank Brain Imaging Documentation (v1.8)<sup>90</sup>, UKBiobank diffusion images were corrected for eddy currents, head motion and outlier-slices using Eddy tool (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/EDDY>). EPI correction was performed using a field map estimated from 3 b=0 images with standard (Anterior-Posterior) phase encoding and 3 b=0 images acquired with reversed phase encoding. Similarly, ABCD diffusion images were corrected for eddy currents (12 free parameters), head motion (rigid-body-registration) and EPI distortion using pairs of b=0 with opposite phase encoding polarities<sup>91</sup>. Neurite Orientation Dispersion and Density Indices (NODDI) parameters were estimated using the Accelerated Microstructure Imaging via Convex Optimization (AMICO)<sup>92</sup> processing approach from the minimally processed diffusion images. The subject specific T1 aligned (based on surface alignment procedures) parcellation template was co-registered to the diffusion weighted image using fsl FLIRT and regional values for FA, MD and the three NODDI parameters were extracted using AFNI's 3dROIstats function for all of the 360 cortical regions included in the Human Connectome Parcellation and averaged across hemisphere to reduce the number of regions to 180 bilateral regions. We also evaluated a direct surface-based registration approach in line with HCP protocols for surface based registration (see Supplementary note 4b).

In total 13 different imaging derived phenotypes were extracted using this pipeline:

1. Total Surface Area of the cortex (SA) (measured at midthickness)
2. Total volume of the cortex (volume)
3. Average thickness of the cortex (CT)
 

Measures of curvature: We calculated five measures of curvature. Assuming two principal curvatures ( $k_1$  and  $k_2$ ), we can define the five measures of curvature as follows.
4. Total mean curvature (MC) =  $(k_1 + k_2)/2$ . MC is typically thought to measure extrinsic curvature. In other words, this is not curvature that is intrinsic to the surface, but rather extrinsic to the surface.
5. Total gaussian curvature (GC) =  $k_1 * k_2$
6. Total intrinsic curvature index (ICI) =  $\max(K, 0)$ . In other words, if GC is positive, ICI is positive. If GC is negative, ICI is 0.
7. Total Folding Index (FI) =  $\text{ABS}(K_1) * (\text{ABS}(K_1) - \text{ABS}(K_2))$ .

8. Total Local Gyrification Index (LGI)<sup>93</sup>: Gyrification index quantifies the amount of the curvature that is buried within the sulcal folds, and is a measure of gyrification. This is computed by calculating the ratio of the area between an outer smoother surface and an inner surface tightly wrapping the pial surface. As it is a ratio, it is a unitless measure.

These measures have consistently been found to have high test-retest reliability (ICC: > 0.8) across sites, acquisition protocols and recent FreeSurfer versions<sup>94–98</sup>. The above properties measure primarily tissue macrostructure. To better understand cortical microstructure, we calculated five measures from the diffusion weighted images<sup>99</sup>. Since conventional diffusion parameters such as fractional anisotropy (FA) and mean diffusivity (MD) alone are not specific to the underlying microstructure of axons and dendrites (referred to, collectively, as neurites) we also extracted measures of Neurite Orientation Density and Dispersion Imaging (NODDI) 100,101.

9. Fractional anisotropy (FA)<sup>99</sup>: FA is thought to be a measure of microstructural integrity. Higher FA values are thought to indicate fiber tracts (i.e., greater anisotropy).

FA would be higher in areas of greater neurite density due to less isotropic water diffusion.

10. Mean diffusivity (MD)<sup>99</sup>: MD measures the degree of displacement (or diffusivity) of water. It can be a measure of membrane density and degree of myelination. Lower membrane density and greater myelination is thought to decrease MD.

We calculated the following three metrics using NODDI. NODDI assumes three microstructural environments for the diffusion of water - intracellular, extracellular, and CSF<sup>47</sup>. The intracellular environment is anisotropic and water diffusion in this environment can be quantified using

11. Intracellular volume fraction (ICVF): Also referred to as neurite density index or NDI, this is a measure of density of neurites (axons and dendrites). Higher ICVF values indicate that a greater fraction of the tissue consists of neurites.
12. Orientation dispersion index (ODI): This measures the orientation and spatial variation of the neurite fibres. Zero indicates perfectly aligned straight fibres, and one for completely isotropic fibres. Thus, larger values of ODI represent highly dispersed neurites and smaller values represent highly aligned neurites.
13. Isotropic volume fraction (ISOVF): This is a measure of water diffusion, typically representing cerebrospinal fluid and ventricles in the cortex.

All white matter metrics have also shown high test-retest reliability (ICC: > 0.8) and scanner consistency, both in a longitudinal subset of the UK BioBank dataset (n=2,817, mean scan-to-scan interval 2.25 years sd: 0.12)<sup>98</sup> as well as in a specifically designed test-retest cohort to evaluate both inter-vendor as well as scan-rescan reliability<sup>102</sup>.

We note that all phenotypes were standardised. Mean CT was calculated as the average across the 180 bilaterally averaged cortical regions. Due to this standardisation, the standardised scores from the average and total values will be identical.

We calculated the hemispheric asymmetry between the regional values using the widely-used Asymmetry Index<sup>73,103–105</sup>

$$\text{Asymmetry Index} = (\text{left} - \text{right}) / ((\text{left} + \text{right}) / 2)$$

## Genome-wide association analyses

**Genetic quality control in the UK Biobank**—Genetic quality control and imputation of the UKB was done by the UK Biobank team and described in detail elsewhere<sup>18</sup>.

After this, we included only individuals of self-reported white European ethnicity, and from this group of individuals, excluded individuals who were above  $\pm 5$  SD from the means of the first two genetic principal components. We further removed individuals whose genetic sex did not match their reported sex, or had excessive genetic heterozygosity, as provided by the UKB team. For the GWAS, we used all genotyped and imputed SNPs in the UK Biobank that had a minor allele frequency  $> 0.01\%$  and were in Hardy-Weinberg Equilibrium ( $p < 1 \times 10^{-6}$ ) and, for imputed SNPs, had an imputation  $r^2 > 0.4$ . After quality control, we retained a maximum of 31,977 participants and 15,916,802 SNPs. We conducted our analyses using people of predominantly European genetic ancestries as this represented the largest, relatively genetically homogeneous group. We did not conduct GWAS for individuals in other ethnic groups as there were fewer than 400 individuals with imaging and genetic data after quality control in each of the other ethnic groups, which is insufficient sample size for linear mixed effect models for GWAS. However, as greater data from other ethnic groups become available, we will revisit these analyses.

**Genetic quality control in ABCD**—Prior to imputation, we filtered SNPs with missingness  $> 90\%$ , and deviations from Hardy Weinberg Equilibrium ( $p < 1 \times 10^{-6}$ ). We removed individuals with missingness  $> 5\%$  and whose genetic sex did not match their reported sex. As HWE and heterozygosity are incorrectly calculated in populations with diverse genetic ancestries, these steps were conducted in relatively homogenous genetic ancestral groups identified using principal-component based clustering after combining the data with the 1000 Genomes phase 3 data<sup>106</sup>. Principal components were calculated using GENESIS<sup>107</sup> after accounting for relatedness between samples as calculated using KING<sup>108</sup>. We calculated genetic principal components using only genotyped SNPs that had passed quality control, and after pruning the SNPs to account for LD ( $r^2 > 0.1$ ), and after removing the MHC locus, a region of long-range LD. To identify genetically homogeneous groups, we used the first five principal components to identify clusters in the 1000 Genomes data using UMAP, identifying 7 broad populations - Non-finnish Europeans, Finnish Europeans, Africans, Americans, East Asians, South Asian, and Bengali. Then, using the first five PCs from the ABCD dataset, we projected individuals onto the seven clusters, identifying broadly homogeneous populations (SF 18). HWE based filtering ( $p < 1 \times 10^{-6}$ ), and removing individuals with excess heterozygosity ( $\pm 3$  SD) was then conducted. After clustering into genetically homogeneous groups, we additionally calculated genetic PCs specifically in the

subgroup of the ABCD participants that were predominantly of European genetic ancestries, again using a pruned set of genotyped SNPs and after excluding the MHC. The data was then merged, and phased (Eagle v 2.4) and imputed (Minimac4) using the TOPMED Imputation Server. From the imputed data, we removed SNPs with poor imputation ( $r^2 < 0.4$ ) and minor allele frequency  $< 0.1\%$  ( $N = 14,495,763$  SNPs). We restricted our analyses to individuals of primarily European ancestries ( $N = 4,866$ ).

**Genome-wide association analyses**—In both the UKB and ABCD we followed the same procedures outlined below. We conducted whole brain and regional GWAS analyses for the 13 phenotypes mentioned in the “Image Processing” section. For each region, we averaged the values bilaterally, resulting in a total of 180 regional phenotypes per phenotype. For ICI and FI we excluded regions “52”, “PI”, “PHA2” because of no variance. In total, we conducted 2,347 GWAS using FastGWA<sup>109</sup>. FastGWA can simultaneously account for both relatedness and subtle population stratification in the analyses.

All phenotypes were scaled to a mean of 0 and a standard deviation of 1. We removed individuals who scored above or below 5 SDs from the mean for all phenotypes, as these are most likely technical outliers. Furthermore, these outliers skew the phenotypic scores and cannot be used in fastGWA which can produce false positives at stringent p-value or for SNPs with low minor allele frequencies<sup>109</sup>. Additionally, we visually inspected histograms of all phenotypes and further removed outliers above or below 5 median absolute deviations for phenotypes with substantial skew, primarily for MD and FI. Additionally, to ensure that the GWAS were not confounded by fine-scale population stratification, among the individuals of European ancestry identified in UKB or ABCD, we removed individuals who were above or below 5 SDs from the mean of the first two genetic principal components. For all GWAS, we included age, age<sup>2</sup>, sex, agexsex, age<sup>2</sup>xsex, imaging centre, first 40 genetic principal components, mean framewise displacement (as obtained from the accompanying resting-state fMRI scan), maximum framewise displacement (as obtained from the accompanying resting-state fMRI scan), and Euler Index<sup>89</sup> as covariates. In addition, for structural MRI metrics derived from T1, we included the inclusion of T2 scans as covariates as this influenced the calculation of these metrics. To ensure this inclusion did not bias our results we also computed separate GWAS for individuals with both T1 and T2-FLAIR and individuals who only had a usable T1 but no additional T2-FLAIR. Genetic correlations between the T1-only sample and the T1 plus T2-flair weighted sample was indistinguishable from 1 for all 7 GWAS, indicating that the overall genetic architecture is identical. Furthermore, effect sizes of the genome-wide significant SNP were highly correlated between the T1 only GWAS and the combined GWAS ( $r = 0.9998$ ,  $p < 2E-16$ ).

For the regional GWAS, we chose not to include the respective global phenotypes for three reasons. First, adjusting for heritable and highly correlated phenotypes biases the GWAS estimates<sup>64,65</sup>. All global phenotypes were substantially heritable and highly correlated with the regional phenotypes (detailed in SN 3). Second, including highly correlated and heritable covariates may result in collider bias for downstream analyses such as Mendelian Randomisation<sup>110</sup>. Given that we wish to make the summary statistics available for researchers to conduct other analyses, including global phenotypes as a covariate can restrict the scope of downstream analyses. Finally, we were specifically interested in identifying

SNPs with effects across the cortex, which may not have been possible if we had adjusted for global phenotypes. We note, methods such as genomic-SEM<sup>28</sup>, mtCOJO<sup>111</sup>, and multivariable MR<sup>112</sup> all allow adjustment for global GWAS in downstream analyses. Here, we used genomic-SEM to regress out the genetic effects of the global phenotype for the majority of experiment-wide significant SNPs (definition of which is detailed below), to identify the fraction of SNPs that remained significant. We note that modelling of global vs. local genetic effects at a genome-wide level as conducted elsewhere<sup>68</sup> is beyond the scope of this study.

We meta-analysed results from the UK Biobank and ABCD using inverse-variance weighted meta-analyses in Plink v1.9<sup>113</sup>, excluding SNPs that were absent from the UKB, given the difference in sample sizes (and consequently, statistical power) between the UKB and ABCD. We checked for inflation in summary statistics using the attenuation ratio. In fastGWA which uses a linear mixed effects model, the LDscore intercept is not a good indicator of inflation in test statistics due to population stratification. Instead, as recommended<sup>114</sup>, we used the attenuation ratio:

$$\text{Attenuation ratio} = (\text{LDSC intercept} - 1) / (\text{mean } \chi^2 - 1).$$

We calculated the expected number of variants that would replicate in ABCD at a  $p < 0.05$  (60% power) using previously deve

We investigated if our variants or variants in high LD ( $r^2 > 0.8$  in CEU or GBP populations) were novel or significantly associated with neuroimaging phenotypes (including quality control) metrics sequentially using four different databases: The Oxford Brain Imaging Genetics PheWeb ([PheWeb \(ox.ac.uk\)](http://PheWeb.ox.ac.uk)), GWAS catalog ([GWAS Catalog \(ebi.ac.uk\)](http://GWAS.Catalog.ebi.ac.uk)), GWAS ATLAS ([Genome wide association study ATLAS \(ctglab.nl\)](http://Genome.wide.association.study.ATLAS.ctglab.nl)), and Brain Imaging Genetics Knowledge Portal [Brain Imaging Genetics Summary Statistics](http://Brain.Imaging.Genetics.Summary.Statistics).

**Multiple testing correction**—Using matrix decomposition<sup>115</sup>, we estimated that there were 1,092 independent phenotypes. This was estimated from all 2,347 global and regional phenotypes included in the study, and thus corrects for all the tests conducted in the study. Consequently, using the total number of independent phenotypes, we used Bonferroni correction to define an experiment-wide threshold of  $4.58 \times 10^{-11}$  ( $5 \times 10^{-8} / 1,092$ ) to correct for the multiple tests conducted. To identify significant loci, we clumped the GWAS using an  $r^2$  threshold of 0.1 over 1000kb. We used LD information available from a random sample of 5,000 unrelated individuals from the UK Biobank who were included in the GWAS.

## Genetic correlation and causal analyses

**Genetic correlation, SNP heritability estimation, clustering, and genomic structural equation modelling**—For the global phenotypes we used LDSC<sup>21,25</sup> to compute genetic correlations and SNP heritability for the meta-analysed GWAS statistics, using LD weights from the North West European populations. Intercepts were not constrained. Heritability and genetic correlation (among 180 regions per phenotype for all



13 phenotypes) of the regional GWAS were calculated using LDSC as incorporated within genomic-SEM<sup>28</sup>. Additionally, for the global phenotypes in the UK-biobank we conducted GCTA-GREML<sup>116</sup> based SNP heritability using genetic relationship matrix calculated using all imputed SNPs included in the GWAS, for 30,765 unrelated individuals (using a gcta-greml cutoff of 0.05) with neuroimaging GWAS. For the Asymmetry Indices, we calculated SNP heritability for a subset of approximately 9,650 unrelated individuals. We applied the same quality control and used the same covariates as for the GWAS.

For the global phenotypes, clustering of the phenotypic and genetic correlation matrices were conducted on the Euclidean distance. As the final hierarchical clustering is dependent on the clustering method used, we used three different clustering methods (Average, Ward D, and Complete Linkage) and visualised the different clusters obtained. Cophenetic correlations (in R, stats (version 3.6.2)) were obtained by comparing the phenotypic and genetic dendrograms produced by the different clustering methods.

Genomic structural equation modelling was conducted using genomic-SEM<sup>28</sup> using summary GWAS statistics of the global cortical phenotypes. We conducted exploratory factor analyses using the even chromosomes, identified factor models, and conducted confirmatory factor analyses using the odd chromosomes. The final model was selected after multiple iterations based on both fit indices and theoretical predictions. Fit indices and path diagrams are provided for models based on all chromosomes.

For the regional phenotypes, we conducted 1000 spin permutations<sup>72</sup> tests to investigate if SNP heritability of regions or genetic correlation among regions were higher in regions falling within functionally<sup>70</sup> or morphologically similar classes<sup>71</sup>. Spin permutation accounts for spatial correspondence between regions and generates null models using random rotations across the spherical cortical surface<sup>72</sup>.

We investigated if the genetic correlation among regions were correlated with topological geodesic distances among regions using Mantel test (within each phenotype separately). We investigated if clustering of regions based on genetic correlations were similar between phenotypes based on cophenetic correlation.

**Phenotypic correlation and principal component analysis**—Comparable to region specific genetic correlations, we also generated region to region phenotypic correlation matrices (“structural covariance”) for both UK Biobank and ABCD cohorts by taking the Pearson correlation across subjects on the scaled and filtered data. UK Biobank and ABCD were then combined into a single meta covariance matrix using the “psychmeta” package in R<sup>117</sup>.

We extracted the first principal component from the regional genetic correlation matrix and regional phenotypic correlation matrix for each of the thirteen phenotypes separately. This principal component analysis was done using a singular value decomposition of the centred and scaled similarity matrix using the “stats” package in R.

**Co-localisation**—To identify co-localised genomic regions among the 13 global phenotypes, we used Hyprcoloc<sup>118</sup>. Hyprcoloc is robust to participant overlap, and can

conduct multi-trait co-localisation using hundreds of GWAS. We restricted our analyses to experiment-wide significant loci, and mapped these onto predefined approximately independent LD blocks in individuals of European ancestry (approximately 1.6 Mb on average)<sup>119</sup>. We did not adjust for either participant or known correlation between phenotypes, as the method gives reasonable results comparable to adjusting for correlation between phenotypes. We used the branch and bound divisive clustering algorithm incorporated in Hyprcoloc to identify clusters of phenotypes that co-localise at any given locus. We used the default variant specific prior probabilities in Hyprcoloc<sup>118</sup>: prior 1 (probability that a SNP is associated with a single trait) as  $1 \times 10^{-4}$ , and prior c (prior probability that the SNP is associated with a second trait) as 0.02. We identified co-localised genomic regions if the genomic regional association probability was 0.6 or higher. We used this probability as simulation analyses by the authors of the method<sup>118</sup> demonstrate that at a regional association probability of 0.6, the empirical probability of identifying true clusters is greater than 90%. We used the same pipeline to investigate co-localisation for 180 regional GWAS and the global GWAS for each of the 13 phenotypes conducted separately.

**Mendelian Randomisation**—To investigate the causal effects of SA on other cortical macrostructural phenotypes, we conducted Mendelian Randomisation analyses<sup>29</sup> using global phenotypes. To avoid bias due to participant overlap, we randomly divided the UK Biobank into two groups of individuals (Group A: N = 15,884 of which males = 7,455; Group B, N = 15,899, of which males = 7,500) and conducted GWAS analyses in each of the groups separately for the eight cortical macrostructural phenotypes using the same pipeline as detailed above. We generated instruments which consisted of SNPs with  $p < 5 \times 10^{-8}$  in the exposure, with minor allele frequency  $> 1\%$ , and which were near-independent (clumping  $r^2 = 0.001$  using a 1000 kb window using data from 5,000 unrelated individuals from the UK Biobank). Where fewer than five SNPs met this criteria, we relaxed the p-value threshold to  $p < 1 \times 10^{-6}$ . Using SA GWAS generated in Group A as the exposure and the GWAS for the remaining 6 phenotypes in Group B as the outcome, we conducted inverse-variance weighted bi-directional Mendelian Randomisation analyses. To account for pleiotropy, we additionally conducted the following sensitivity analyses: (1) median weighted MR (majority-valid<sup>120</sup>); (2) MR-Egger (accounts for pleiotropy)<sup>121</sup>; (3) MR PRESSO (detects and excludes outliers in the instrument<sup>122</sup>).

Additionally, (4) for the significant MR results, to further account for correlated (vertical) pleiotropy, we conducted MR analyses using CAUSE<sup>123</sup> using two instruments: one with of SNPs with  $p < 5 \times 10^{-8}$ , and another at a more relaxed threshold of  $p < 0.001$ . We investigated heterogeneity in the instrument using Cochran's Q, and investigated if the Egger intercept was significant. We investigated if the orientation of the causal direction was correct using Steiger analyses<sup>124</sup>, and conducted additional sensitivity analyses after removing SNPs that did not have the correct causal orientation. Finally, we inspected the scatter plot, forest plot, and plots generated from leave-one-out analyses to identify if the results were driven by a subset of the SNPs. Analyses (1) and (2), and the sensitivity analyses were conducted using the R-package TwosampleMR<sup>125</sup>.

We repeated all MR analyses except for CAUSE using instruments generated in the UKB as the exposure and ABCD as the outcome. However, this was quasi-bidirectional, in that

in both directions, the exposure was instruments generated in the UKB and the outcome was SNPs in the ABCD. We did not conduct CAUSE in this instance due to sample size imbalance which reduces statistical power.

Given substantial pleiotropy between the phenotypes, we identified significant MR associations if: 1. The p-value was  $< 0.0035$  (bonferroni corrected threshold) in both the within UKB and the UKB-ABCD analyses for IVW, MR-PRESSO, and weighted median; 2 MR-Egger was in the consistent direction to the IVW (MR-Egger has lower statistical power so we did not require it to be statistically significant); 3. If Steiger analyses identified incorrect causal orientation, criteria 1 and 2 were met after Steiger filtering; and 4. Results were significant when MR was conducted using CAUSE which accounts for correlated pleiotropy. Analyses were conducted using the two-sample MR package (version 0.5.6)<sup>125</sup>. Power-calculations<sup>126</sup> were conducted assuming a standard deviation in the exposure results in a 0.33 unit standard deviation change in the outcome, which is a medium effect size.

### Gene-based association and enrichment analyses

**Gene based association**—We used MAGMA (version 1. 10)<sup>37</sup> to conduct gene-based association testing based on physical location. MAGMA assigns SNPs to the nearest gene. In line with previous analyses, we expanded the window to 35kb upstream and 10kb downstream of the gene to capture regulatory regions<sup>127</sup>. In addition, we used HMAGMA<sup>38</sup> (using MAGMA v 1.08) to identify genes based on Hi-C mapping. In contrast to MAGMA, HMAGMA is able to map SNPs to genes based on long-range interactions and can account for tissue specific regulatory effects. To map developmental trajectories, we used Hi-C data from postnatal and prenatal human cortex<sup>38,128</sup>. Subsequently, for enrichment analyses, we used Hi-C data from the prenatal cortex given that the majority of the phenotypes were either enriched for gene expression in the prenatal cortex or did not differ in gene expression between prenatal and postnatal cortex, and because many processes investigated occurred prenatally.

**Developmental trajectories**—To identify patterns of gene expression across cortical prenatal and postnatal windows, we used data from PsychEncode<sup>17</sup>. The data was divided into 9 developmental windows: Window 1. 8 - 9 post-conception weeks (PCW); Window 2. 12 - 13 PCW; Window 3. 16 - 17 PCW; Window 4. 19 - 22 PCW; Window 5. 35 PCW - 4 months; Window 6. 6 months to 2.5 years; Window 7. 3 - 11 years; Window 8. 13 - 19 years; Window 9. 21 - 40 years. Gene expression values were log base 2 transformed after adding a pseudocount and normalised. For 12 of the 13 phenotypes the transformed expression values of all genes with  $q < 0.05$  were averaged for each developmental window and smoothed LOESS curves were plotted. The excluded phenotype was FA as HMAGMA and MAGMA identified 1 and 0 genes with  $q < 0.05$  respectively.

**Enrichment analyses**—To investigate enrichment for cell types, signatures of genomic constraint, and gene sets associated with neurodevelopmental and cephalic disorders, we conducted the following analyses. Within each gene set, significant results were identified after correcting for all 13 phenotypes using Benjamini-Hochberg FDR correction ( $q < 0.05$ ).

To identify cell types in the prenatal and postnatal cortex we conducted enrichment analyses using (1) single cell gene expression data from PsychENCODE<sup>17</sup> using prenatal (5pcw to 125 days) and postnatal gene expression. To provide additional temporal resolution we also conducted analyses using (2) single cell gene expression data that spanned early cortical development (6 - 10 pcw)<sup>40</sup>; (3) single cell gene expression data spanning mid gestation period of cortical development (17 - 18 pcw)<sup>41</sup>; (4) Single-cell epigenomic data (scATAC-seq) from the midgestation period of cortical development<sup>42</sup>; and (5) Cell-type specific (fluorescent-activated nuclei sorting isolated) epigenomic (ATAC-seq and ChIP-seq) data from postnatal cortex<sup>46</sup>. Analyses for datasets 1 - 3 were conducted using MAGMA gene-set enrichment using genes identified by MAGMA and HMAGMA. Following previously described methods<sup>127</sup>, we filtered out genes with non-unique names and genes not expressed in any cell types. Gene expression values were log base 2 transformed after adding a pseudocount and normalised. Mean cell-type specific gene values were calculated, and this was divided by the mean expression of the gene in all cells to get relative cell-type expression. We then selected the top 10% of genes with the highest relative expression in each cell type to conduct enrichment analyses using MAGMA gene-set enrichment analyses<sup>37</sup>. Significant cell types were identified if  $q < 0.05$  in analyses using both HMAGMA and MAGMA identified genes. Analyses for datasets 4 and 5 were conducted using conditional partitioned heritability analyses in LDSC (i.e., enrichment for a cell type after conditioning on all other cell types and baseline annotations)<sup>129,130</sup>.

We used the same gene-enrichment pipeline as above to investigate gene enrichment for genes that are constrained ( $pLOUEF < 0.37$ )<sup>49</sup>, genes associated with neurodevelopmental disorders<sup>50</sup> (662 genes with  $FDR < 0.05$ ), and genes associated with severe microcephaly obtained from the Genomics England Panel (244 genes, signed off on March 2, 2022: Severe microcephaly (Version 2.304) (genomicsengland.co.uk)). Signatures of selection were identified using SBayesS<sup>48</sup>.

## Polygenic score association analyses

**Genetic quality control and polygenic score generation**—Polygenic scores (PGS) for SA, CT, and volume were calculated using the meta-analysed GWAS in a dataset of individuals with severe developmental disorders (DDD study,  $N = 6,916$ ) and autistic individuals and their families (SPARK dataset,  $N = 25,621$ ) using PRSs<sup>131</sup>. PRSs is a Bayesian algorithm that infers posterior effect sizes of SNPs using continuous shrinkage, and does not require defining p-value thresholds. Details of genetic quality control in the DDD cohort in individuals of predominantly European ancestries are provided elsewhere<sup>52</sup>. The data were re-imputed using the TOPMed reference panel and variants with low imputation quality ( $\text{minimac4 } r^2 < 0.8$ ) were excluded. We kept common (minor allele frequency  $> 1\%$ ) SNPs that are also in HapMap3 to calculate the PGS using PRSs. Genetic-ancestry QC of the SPARK dataset was conducted similar to the ABCD dataset and as detailed elsewhere<sup>132,133</sup>. We calculated PGS on individuals of predominantly European ancestries as identified by genetic principal components. All polygenic scores were standardised with a mean zero and a standard deviation of 1 for all analyses.

**Defining phenotypes in DDD and SPARK**—In the DDD study, we used HPO terms assigned by clinicians to define macrocephaly (N=396 with HPO term “HP:0040194”, “HP:0000256”, “HP:0004482”, “HP:0004481”, “HP:0004488”, or “HP:0005490”) and microcephaly (N=1,198 with HPO term “HP:0040195”, “HP:0000252”, “HP:0005484”, “HP:0004485”, “HP:0000253”, “HP:0011451”, or “HP:0040196”). We also analysed occipital frontal circumference data (N=6,146), which were calculated as standard deviations from the mean given the proband’s gestational age at birth, age at time of measurement and sex. In the SPARK dataset, information about macro- and microcephaly were obtained from parental/caregiver reports of medical diagnoses.

**Statistical analyses**—Linear or logistic mixed-effect regressions (random intercepts for family, in SPARK) were conducted using either: (1) PGS for volume; (2) PGS for SA and CT in a multiple regression framework. Primary analyses were conducted using logistic regression, separately for macrocephaly and microcephaly (coded as 1) compared to controls (i.e. individuals in the cohort without microcephaly or macrocephaly) (coded as 0). In the DDD, we also conducted linear regression using standardised occipital frontal circumference. Additionally, we conducted linear regression with macrocephaly coded as 1, microcephaly as -1 and no diagnosis as 0. In the DDD study, we included sex, genetic diagnosis, and the first 10 genetic principal components as covariates. Specifically, we considered probands to be “diagnosed” if they had at least one variant reported to DECIPHER that had been confirmed as pathogenic or likely pathogenic (C/LP) by a clinician, or that had been predicted as P/LP by a computational algorithm based on the American College of Medical Genetics criteria, as described in Wright et al.<sup>134</sup>. In SPARK, age, sex, autism diagnosis, and the first 10 genetic principal components were included as covariates. Significant results were identified after Benjamini-Hochberg FDR correction ( $q < 0.05$ ) across all models.

### **Fine-mapping, Summary Mendelian Randomisation, and prioritising candidate genes**

For all exome-wide significant loci in the global GWAS (N = 90), we conducted functionally informed fine-mapping using Polyfun<sup>54</sup>, using SuSiE<sup>135</sup> as the fine-mapping method and with up to 5 causal variants per locus, with each locus defined 500 kb upstream and downstream of the sentinel variant. In-sample LD was obtained from 5,000 unrelated individuals included in the GWAS from the UK Biobank. We used precomputed prior causal probabilities from the UK Biobank as provided in Polyfun.

To link the variants in the 95% credible sets to genes, we used Hi-C data from (1) the prenatal brain germinal zone<sup>128</sup>, (2) the prenatal brain cortical plate<sup>128</sup>, (3) neurons from postnatal cortex<sup>136</sup>, and (4) glia from postnatal cortex<sup>136</sup>. Additionally, we (5) used Ensembl Variant Effect Predictor<sup>137</sup> to identify genes containing damaging missense (deleterious in SIFT and/or damaging/probably damaging/possibly damaging in PolyPhen) and protein truncating variants from the list of the 95% credible sets.

To identify candidate genes using relevant eQTL and methylation data, we further conducted summary Mendelian randomisation (SMR)<sup>138</sup>. SMR was conducted for all 13 phenotypes, using cis-eQTL data from postmortem (6) prenatal<sup>139</sup> and (7) postnatal brains<sup>140</sup>, and

additionally (8) methylation data from postnatal brains<sup>141</sup>. Within each phenotype, we identified significant genes by using Bonferroni correction for the total number of genes tested. We excluded significant genes with evidence to indicate that the MR association results are due to pleiotropy using the HEIDI test (HEIDI  $p < 0.01$ )<sup>138</sup>.

Finally, (9) we identified the closest gene to each sentinel variant (i.e., the SNP with the lowest  $p$ -value in each locus). Where the variant was intergenic, we included both the closest upstream and downstream gene. From these nine methods, we identify a list of prioritised candidate genes if they are supported by at least two methods. We conducted GO enrichment analyses to identify biological pathways enriched for the prioritised candidates genes.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

VW is supported by St. Catharine's College Cambridge. E-MS is supported by a Ph.D. studentship awarded by the Friends of Peterhouse. EAWS is supported by the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. RAIB is supported by the Autism Research Trust. SBC received funding from the Wellcome Trust 214322/Z/18/Z. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. SBC also received funding from the Autism Centre of Excellence, SFARI, the Templeton World Charitable Fund, the MRC, and the NIHR Cambridge Biomedical Research Centre. The research was supported by the National Institute for Health Research (NIHR) Applied Research Collaboration East of England. JS was supported by NIMH T32MH019112-29 and K08MH120564. E.T.B. was supported by an NIHR Senior Investigator award and the Wellcome Trust collaborative award for the Neuroscience in Psychiatry Network. A.F.A.-B. was supported by NIMH K08MH120564. R.R.G. was supported by EMERGIA Junta de Andalucía program (EMERGIA20\_00139). S.L.V. was supported by Max Planck Gesellschaft, (Otto Hahn Award) and the Helmholtz Association's Initiative and Networking Fund under the Helmholtz International Lab grant agreement InterLabs-0015, and the Canada First Research Excellence Fund (CFREF Competition 2, 2015–2016) awarded to the Healthy Brains, Healthy Lives initiative at McGill University, through the Helmholtz International BigBrain Analytics and Learning Laboratory (HIBALL). GKM was supported by MRC MR/W020025/1.

Data were curated and analysed using a computational facility funded by an MRC research infrastructure award (MR/M009041/1) to the School of Clinical Medicine, University of Cambridge and supported by the mental health theme of the NIHR Cambridge Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NIH, NHS, the NIHR or the Department of Health and Social Care. Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive DevelopmentSM (ABCD) Study (<https://abcdstudy.org>), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children age 9-10 and follow them over 10 years into early adulthood.

The ABCD Study® is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. A full list of supporters is available at <https://abcdstudy.org/federal-partners.html>. A listing of participating sites and a complete listing of the study investigators can be found at [https://abcdstudy.org/consortium\\_members/](https://abcdstudy.org/consortium_members/). ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. The ABCD data repository grows and changes over time. The ABCD data used in this report came from [NIMH Data Archive Digital Object Identifier (10.15154/1503209)]. DOIs can be found at <http://dx.doi.org/10.15154/1503209>.

For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.



We thank Luke Kweku Abraham and Jennifer Asimit for helpful discussions. The authors do not declare any competing interests.

## Data availability

All summary statistics for the GWAS meta-analyses will be made available upon publication without restriction on <https://portal.idc-cam.org.uk/overview/483>.

## Code availability

<https://github.com/ucam-department-of-psychiatry/UKB>

<https://github.com/ucam-department-of-psychiatry/ABCD>

## References

1. Bethlehem RAI, et al. Brain charts for the human lifespan. *Nature*. 2022; 604: 525–533. [PubMed: 35388223]
2. Thompson PM, et al. ENIGMA and global neuroscience: A decade of large-scale studies of the brain in health and disease across more than 40 countries. *Transl Psychiatry*. 2020; 10: 100. [PubMed: 32198361]
3. Gilmore JH, Knickmeyer RC, Gao W. Imaging structural and functional brain development in early childhood. *Nat Rev Neurosci*. 2018; 19: 123–137. [PubMed: 29449712]
4. Paus T, Keshavan M, Giedd JN. Why do many psychiatric disorders emerge during adolescence? *Nat Rev Neurosci*. 2008; 9: 947–957. [PubMed: 19002191]
5. Elliott LT, et al. Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature*. 2018; 562: 210–216. [PubMed: 30305740]
6. Grasby KL, et al. The genetic architecture of the human cerebral cortex. *Science*. 2020; 367
7. Stein JL, et al. Identification of common variants associated with human hippocampal and intracranial volumes. *Nat Genet*. 2012; 44: 552–561. [PubMed: 22504417]
8. Zhao B, et al. Common genetic variation influencing human white matter microstructure. *Science*. 2021; 372
9. Makowski C, et al. Discovery of genomic loci of the human cerebral cortex using genetically informed brain atlases. *Science*. 2022; 375: 522–528. [PubMed: 35113692]
10. Jansen PR, et al. Genome-wide meta-analysis of brain volume identifies genomic loci and genes shared with intelligence. *Nat Commun*. 2020; 11: 5606. [PubMed: 33154357]
11. Smith SM, et al. An expanded set of genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nat Neurosci*. 2021; 24: 737–745. [PubMed: 33875891]
12. Naqvi S, et al. Shared heritability of human face and brain shape. *Nat Genet*. 2021; 53: 830–839. [PubMed: 33821002]
13. Whalley HC, et al. Polygenic risk and white matter integrity in individuals at high risk of mood disorder. *Biol Psychiatry*. 2013; 74: 280–286. [PubMed: 23453289]
14. Reus LM, et al. Association of polygenic risk for major psychiatric illness with subcortical volumes and white matter integrity in UK Biobank. doi: 10.1101/080283
15. Stauffer E-M, et al. Grey and white matter microstructure is associated with polygenic risk for schizophrenia. *Mol Psychiatry*. 2021; doi: 10.1038/s41380-021-01260-5
16. Jayaraman D, Bae B-I, Walsh CA. The Genetics of Primary Microcephaly. *Annu Rev Genomics Hum Genet*. 2018; 19: 177–200. [PubMed: 29799801]
17. Li M, et al. Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science*. 2018; 362
18. Bycroft C, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018; 562: 203–209. [PubMed: 30305743]

19. Casey BJ, et al. The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev Cogn Neurosci*. 2018; doi: 10.1016/J.DCN.2018.03.001
20. Glasser MF, et al. A multi-modal parcellation of human cerebral cortex. *Nature*. 2016; 536: 171–178. [PubMed: 27437579]
21. Bulik-Sullivan BK, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. 2015; 47: 1236–1241. [PubMed: 26414676]
22. Hedman AM, van Haren NEM, Schnack HG, Kahn RS, Hulshoff Pol HE. Human brain changes across the life span: a review of 56 longitudinal magnetic resonance imaging studies. *Hum Brain Mapp*. 2012; 33: 1987–2002. [PubMed: 21915942]
23. Brouwer RM, et al. Genetic variants associated with longitudinal changes in brain structure across the lifespan. *Nat Neurosci*. 2022; 25: 421–432. [PubMed: 35383335]
24. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010; 26: 2190–2191. [PubMed: 20616382]
25. Bulik-Sullivan BK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet*. 2015; 47: 291–295. [PubMed: 25642630]
26. Hofer E, et al. Genetic correlations and genome-wide associations of cortical structure in general population samples of 22,824 adults. *Nat Commun*. 2020; 11: 4796. [PubMed: 32963231]
27. Sodini SM, Kemper KE, Wray NR, Trzaskowski M. Comparison of Genotypic and Phenotypic Correlations: Cheverud's Conjecture in Humans. *Genetics*. 2018; 209: 941–948. [PubMed: 29739817]
28. Grotzinger AD, et al. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature Human Behaviour*. 2019; 1
29. Sanderson E, et al. Mendelian randomization. *Nature Reviews Methods Primers*. 2022; 2: 1–21.
30. Rakic P. Specification of cerebral cortical areas. *Science*. 1988; 241: 170–176. [PubMed: 3291116]
31. Ronan L, et al. Differential tangential expansion as a mechanism for cortical gyrification. *Cereb Cortex*. 2014; 24: 2219–2228. [PubMed: 23542881]
32. Garcia KE, Kroenke CD, Bayly PV. Mechanics of cortical folding: stress, growth and stability. *Philos Trans R Soc Lond B Biol Sci*. 2018; 373
33. Richman DP, Stewart RM, Hutchinson JW, Caviness VS Jr. Mechanical model of brain convolucional development. *Science*. 1975; 189: 18–21. [PubMed: 1135626]
34. Tallinen T, Chung JY, Biggins JS, Mahadevan L. Gyrification from constrained cortical expansion. *Proc Natl Acad Sci U S A*. 2014; 111: 12667–12672. [PubMed: 25136099]
35. Reillo I, de Juan Romero C, García-Cabezas MÁ, Borrell V. A role for intermediate radial glia in the tangential expansion of the mammalian cerebral cortex. *Cereb Cortex*. 2011; 21: 1674–1694. [PubMed: 21127018]
36. Kriegstein A, Noctor S, Martínez-Cerdeño V. Patterns of neural stem and progenitor cell division may underlie evolutionary cortical expansion. *Nat Rev Neurosci*. 2006; 7: 883–890. [PubMed: 17033683]
37. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol*. 2015; 11: 1–19.
38. Sey NYA, et al. A computational tool (H-MAGMA) for improved prediction of brain-disorder risk genes by incorporating brain chromatin interaction profiles. *Nat Neurosci*. 2020; 23: 583–593. [PubMed: 32152537]
39. Akbarian S, et al. The PsychENCODE project. *Nat Neurosci*. 2015; 18: 1707–1712. [PubMed: 26605881]
40. Eze UC, Bhaduri A, Haeussler M, Nowakowski TJ, Kriegstein AR. Single-cell atlas of early human brain development highlights heterogeneity of human neuroepithelial cells and early radial glia. *Nat Neurosci*. 2021; 24: 584–594. [PubMed: 33723434]
41. Polioudakis D, et al. A Single-Cell Transcriptomic Atlas of Human Neocortical Development during Mid-gestation. *Neuron*. 2019; 103: 785–801. e8 [PubMed: 31303374]
42. Ziffra RS, et al. Single-cell epigenomics reveals mechanisms of human cortical development. *Nature*. 2021; 598: 205–213. [PubMed: 34616060]

43. Florio M, Huttner WB. Neural progenitors, neurogenesis and the evolution of the neocortex. *Development*. 2014; 141: 2182–2194. [PubMed: 24866113]
44. Geschwind DH, Rakic P. Cortical evolution: judge the brain by its cover. *Neuron*. 2013; 80: 633–647. [PubMed: 24183016]
45. Gertz CC, Lui JH, LaMonica BE, Wang X, Kriegstein AR. Diverse behaviors of outer radial glia in developing ferret and human cortex. *J Neurosci*. 2014; 34: 2559–2570. [PubMed: 24523546]
46. Nott A, et al. Brain cell type-specific enhancer-promoter interactome maps and disease-risk association. *Science*. 2019; 366: 1134–1139. [PubMed: 31727856]
47. Fukutomi H, et al. Neurite imaging reveals microstructural variations in human cerebral cortical gray matter. *Neuroimage*. 2018; 182: 488–499. [PubMed: 29448073]
48. Zeng J, et al. Widespread signatures of natural selection across human complex traits and functional genomic categories. *Nat Commun*. 2021; 12: 1164. [PubMed: 33608517]
49. Karczewski KJ, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020; 581: 434–443. [PubMed: 32461654]
50. Fu JM, et al. Rare coding variation illuminates the allelic architecture, risk genes, cellular expression patterns, and phenotypic context of autism. *bioRxiv*. 2021; doi: 10.1101/2021.12.20.21267194
51. Study, D. D. D. & Deciphering Developmental Disorders Study. Prevalence and architecture of de novo mutations in developmental disorders. *Nature*. 2017; 542: 433–438. DOI: 10.1038/nature21062 [PubMed: 28135719]
52. Niemi MEK, et al. Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. *Nature*. 2018; 562: 268–271. [PubMed: 30258228]
53. SPARK Consortium. Electronic address: pfeliciano@simonsfoundation.org & SPARK Consortium. *Neuron*. Vol. 97. SPARK: A US Cohort of 50,000 Families to Accelerate Autism Research; 2018. 488–493.
54. Weissbrod O, et al. Functionally informed fine-mapping and polygenic localization of complex trait heritability. *Nat Genet*. 2020; 52: 1355–1363. [PubMed: 33199916]
55. Kabeche L, Nguyen HD, Buisson R, Zou L. A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. *Science*. 2018; 359: 108–114. [PubMed: 29170278]
56. Kaczmarczyk A, Sullivan KF. CENP-W plays a role in maintaining bipolar spindle structure. *PLoS One*. 2014; 9 e106464 [PubMed: 25329824]
57. Koolen DA, et al. The Koolen-de Vries syndrome: a phenotypic comparison of patients with a 17q21.31 microdeletion versus a KANSL1 sequence variant. *European Journal of Human Genetics*. 2016; 24: 652–659. DOI: 10.1038/ejhg.2015.178 [PubMed: 26306646]
58. Zhou X, et al. Cellular and molecular properties of neural progenitors in the developing mammalian hypothalamus. *Nat Commun*. 2020; 11: 4063. [PubMed: 32792525]
59. Kuwayama N, et al. A role for Hmga2 in the early-stage transition of neural stem-progenitor cell properties during mouse neocortical development. *bioRxiv*. 2021; 2020.05.14.086330 doi: 10.1101/2020.05.14.086330
60. De Crescenzo A, et al. A splicing mutation of the HMGA2 gene is associated with Silver-Russell syndrome phenotype. *J Hum Genet*. 2015; 60: 287–293. [PubMed: 25809938]
61. Chenn A, Walsh CA. Regulation of cerebral cortical size by control of cell cycle exit in neural precursors. *Science*. 2002; 297: 365–369. [PubMed: 12130776]
62. Xiang Y-Y, et al. Versican G3 domain regulates neurite growth and synaptic transmission of hippocampal neurons by activation of epidermal growth factor receptor. *J Biol Chem*. 2006; 281: 19358–19368. [PubMed: 16648628]
63. Dobyns WB, et al. MACF1 Mutations Encoding Highly Conserved Zinc-Binding Residues of the GAR Domain Cause Defects in Neuronal Migration and Axon Guidance. *Am J Hum Genet*. 2018; 103: 1009–1021. [PubMed: 30471716]
64. Aschard H, Vilhjálmsson BJ, Joshi AD, Price AL, Kraft P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *Am J Hum Genet*. 2015; 96: 329–339. [PubMed: 25640676]

65. Day FR, Loh P-R, Scott RA, Ong KK, Perry JRB. A Robust Example of Collider Bias in a Genetic Association Study. *American journal of human genetics*. 2016; 98: 392–393. [PubMed: 26849114]
66. Chen S, et al. A genome-wide mutational constraint map quantified from variation in 76,156 human genomes. *bioRxiv*. 2022; 2022.03.20.485034 doi: 10.1101/2022.03.20.485034
67. Demange PA, et al. Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. *Nat Genet*. 2021; 53: 35–44. [PubMed: 33414549]
68. Grotzinger AD, et al. Multivariate genomic architecture of cortical thickness and surface area at multiple levels of analysis. *bioRxiv*. 2022; doi: 10.1101/2022.02.19.22271223
69. Bhaduri A, et al. An atlas of cortical arealization identifies dynamic molecular signatures. *Nature*. 2021; 598: 200–204. [PubMed: 34616070]
70. Yeo BTT, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol*. 2011; 106: 1125–1165. [PubMed: 21653723]
71. Mesulam MM. From sensation to cognition. *Brain*. 1998; 121 (Pt 6) 1013–1052. [PubMed: 9648540]
72. Alexander-Bloch AF, et al. On testing for spatial correspondence between maps of human brain structure and function. *Neuroimage*. 2018; 178: 540–551. [PubMed: 29860082]
73. Sha Z, et al. The genetic architecture of structural left-right asymmetry of the human brain. *Nature Human Behaviour*. 2021; 5: 1226–1239. DOI: 10.1038/s41562-021-01069-w
74. Rubenstein JL, Rakic P. Genetic control of cortical development. *Cereb Cortex*. 1999; 9: 521–523. [PubMed: 10498269]
75. Cox SR, et al. Ageing and brain white matter structure in 3,513 UK Biobank participants. *Nat Commun*. 2016; 7 13629 [PubMed: 27976682]
76. Sexton CE, et al. Accelerated changes in white matter microstructure during aging: a longitudinal diffusion tensor imaging study. *J Neurosci*. 2014; 34: 15425–15436. [PubMed: 25392509]
77. Pletikos M, et al. Temporal specification and bilaterality of human neocortical topographic gene expression. *Neuron*. 2014; 81: 321–332. [PubMed: 24373884]
78. Zhu Y, et al. Spatiotemporal transcriptomic divergence across human and macaque brain development. *Science*. 2018; 362
79. Yoon B, Shim Y-S, Lee K-S, Shon Y-M, Yang D-W. Region-specific changes of cerebral white matter during normal aging: a diffusion-tensor analysis. *Arch Gerontol Geriatr*. 2008; 47: 129–138. [PubMed: 17764763]
80. Shi Y, et al. Diffusion Tensor Imaging-Based Characterization of Brain Neurodevelopment in Primates. *Cereb Cortex*. 2012; 23: 36–48. [PubMed: 22275483]
81. Fasching L, et al. Early developmental asymmetries in cell lineage trees in living individuals. *Science*. 2021; 371: 1245–1248. [PubMed: 33737484]
82. Mancini M, et al. An interactive meta-analysis of MRI biomarkers of myelin. *Elife*. 2020; 9
83. Coalson TS, Van Essen DC, Glasser MF. The impact of traditional neuroimaging methods on the spatial localization of cortical areas. *Proc Natl Acad Sci U S A*. 2018; 115: E6356–E6365. [PubMed: 29925602]
84. Kharabian Masouleh S, et al. Influence of Processing Pipeline on Cortical Thickness Measurement. *Cereb Cortex*. 2020; 30: 5014–5027. [PubMed: 32377664]
85. Alfaro-Almagro F, et al. Confound modelling in UK Biobank brain imaging. *bioRxiv*. 2020; 2020.03.11.987693 doi: 10.1101/2020.03.11.987693
86. Barch DM, et al. Demographic, physical and mental health assessments in the adolescent brain and cognitive development study: Rationale and description. *Dev Cogn Neurosci*. 2018; 32: 55–66. [PubMed: 29113758]
87. Fischl B, et al. Automatically parcellating the human cerebral cortex. *Cereb Cortex*. 2004; 14: 11–22. [PubMed: 14654453]
88. Van Essen DC, Glasser MF, Dierker DL, Harwell J, Coalson T. Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases. *Cereb Cortex*. 2012; 22: 2241–2262. [PubMed: 22047963]
89. Rosen AFG, et al. Quantitative assessment of structural image quality. *Neuroimage*. 2018; 169: 407–418. [PubMed: 29278774]

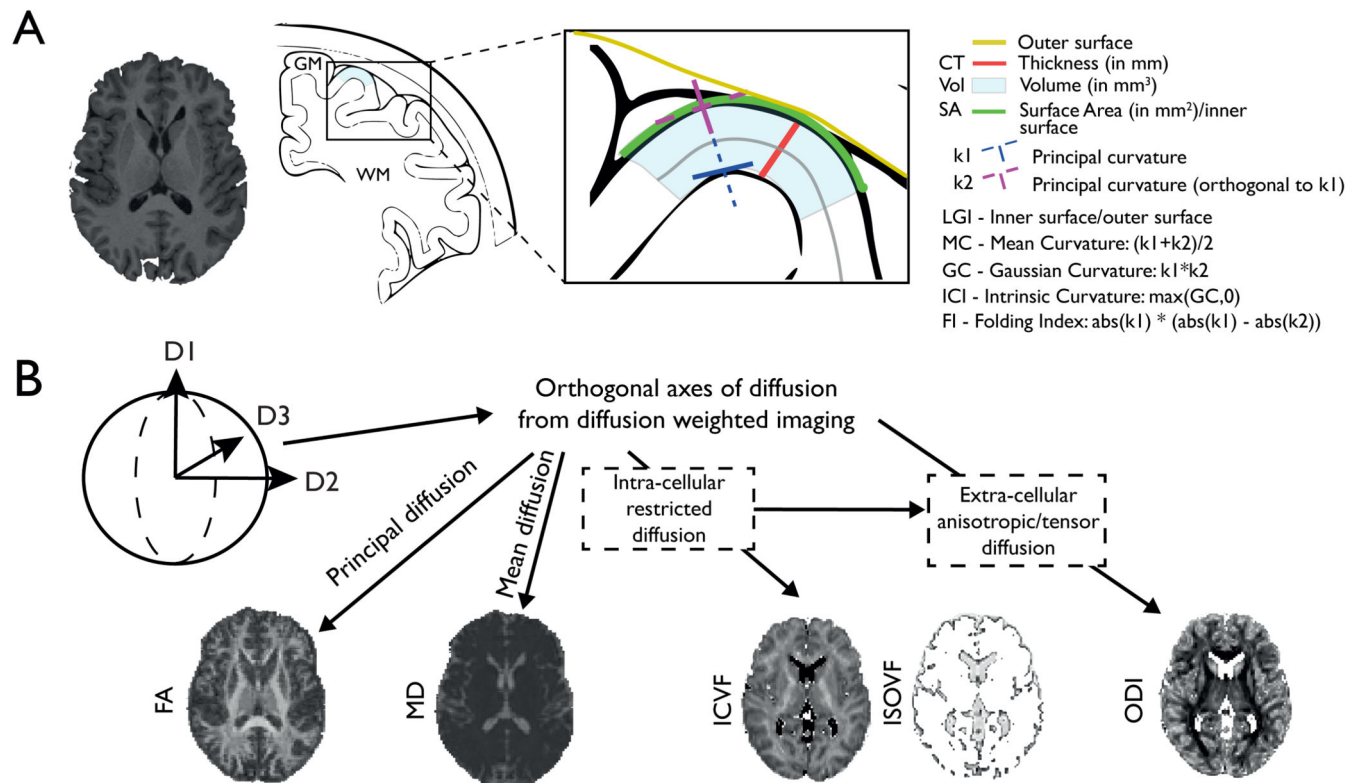
90. Alfaro-Almagro F, et al. Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage*. 2018; 166: 400–424. [PubMed: 29079522]
91. Hagler DJ Jr, et al. Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *Neuroimage*. 2019; 202: 116091 [PubMed: 31415884]
92. Daducci A, et al. Accelerated Microstructure Imaging via Convex Optimization (AMICO) from diffusion MRI data. *Neuroimage*. 2015; 105: 32–44. [PubMed: 25462697]
93. Schaer M, et al. How to measure cortical folding from MR images: a step-by-step tutorial to compute local gyrification index. *J Vis Exp*. 2012. e3417 [PubMed: 22230945]
94. Knussmann GN, et al. Test-retest reliability of FreeSurfer-derived volume, area and cortical thickness from MPRAGE and MP2RAGE brain MRI images. *Neuroimage Rep*. 2022; 2
95. Haddad E, et al. Multisite test-retest reliability and compatibility of brain metrics derived from FreeSurfer versions 7.1, 6.0, and 5.3. *Hum Brain Mapp*. 2022; doi: 10.1002/hbm.26147
96. Hedges EP, et al. Reliability of structural MRI measurements: The effects of scan session, head tilt, inter-scan interval, acquisition sequence, FreeSurfer version and processing stream. *Neuroimage*. 2022; 246: 118751 [PubMed: 34848299]
97. Madan CR, Kensinger EA. Test-retest reliability of brain morphology estimates. *Brain Inform*. 2017; 4: 107–121. [PubMed: 28054317]
98. Duff E, et al. Reliability of multi-site UK Biobank MRI brain phenotypes for the assessment of neuropsychiatric complications of SARS-CoV-2 infection: The COVID-CNS travelling heads study. *PLoS One*. 2022; 17: e0273704 [PubMed: 36173949]
99. O'Donnell LJ, Westin C-F. An introduction to diffusion tensor image analysis. *Neurosurg Clin N Am*. 2011; 22: 185–96. [PubMed: 21435570]
100. Zhang H, Schneider T, Wheeler-Kingshott CA, Alexander DC. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage*. 2012; 61: 1000–1016. [PubMed: 22484410]
101. Tariq M, Schneider T, Alexander DC, Gandini Wheeler-Kingshott CA, Zhang H. Bingham-NODDI: Mapping anisotropic orientation dispersion of neurites using diffusion MRI. *Neuroimage*. 2016; 133: 207–223. [PubMed: 26826512]
102. Andica C, et al. Scan-rescan and inter-vendor reproducibility of neurite orientation dispersion and density imaging metrics. *Neuroradiology*. 2020; 62: 483–494. [PubMed: 31883043]
103. Kong X-Z, et al. Mapping cortical brain asymmetry in 17,141 healthy individuals worldwide via the ENIGMA Consortium. *Proc Natl Acad Sci U S A*. 2018; 115: E5154–E5163. [PubMed: 29764998]
104. Kurth F, Gaser C, Luders E. A 12-step user guide for analyzing voxel-wise gray matter asymmetries in statistical parametric mapping (SPM). *Nat Protoc*. 2015; 10: 293–304. [PubMed: 25591011]
105. Leroy F, et al. New human-specific brain landmark: the depth asymmetry of superior temporal sulcus. *Proc Natl Acad Sci U S A*. 2015; 112: 1208–1213. [PubMed: 25583500]
106. 1000 Genomes Project Consortium. et al. A global reference for human genetic variation. *Nature*. 2015; 526: 68–74. [PubMed: 26432245]
107. Gogarten SM, et al. Genetic association testing using the GENESIS R/Bioconductor package. *Bioinformatics*. 2019; 35: 5346–5348. [PubMed: 31329242]
108. Manichaikul A, et al. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010; 26: 2867–2873. [PubMed: 20926424]
109. Jiang L, et al. A resource-efficient tool for mixed model association analysis of large-scale data. *Nat Genet*. 2019; 51: 1749–1755. [PubMed: 31768069]
110. Hartwig FP, Tilling K, Davey Smith G, Lawlor DA, Borges MC. Bias in two-sample Mendelian randomization when using heritable covariable-adjusted summary associations. *Int J Epidemiol*. 2021; 50: 1639–1650. [PubMed: 33619569]
111. Zhu Z, et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat Commun*. 2018; 9: 224. [PubMed: 29335400]



112. Burgess S, Thompson SG. Multivariable Mendelian Randomization: The Use of Pleiotropic Genetic Variants to Estimate Causal Effects. *American Journal of Epidemiology*. 2015; 181: 251–260. DOI: 10.1093/aje/kwu283 [PubMed: 25632051]
113. Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81: 559–575. [PubMed: 17701901]
114. Loh P-R, Kichaev G, Gazal S, Schoech AP, Price AL. Mixed-model association for biobank-scale datasets. *Nat Genet*. 2018; 50: 906–908. [PubMed: 29892013]
115. Zheng J, et al. PhenoSpD: an integrated toolkit for phenotypic correlation estimation and multiple testing correction using GWAS summary statistics. *Gigascience*. 2018; 7
116. Yang J, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet*. 2010; 42: 565–569. [PubMed: 20562875]
117. Dahlke JA, Wiernik BM. psychmeta: An R Package for Psychometric Meta-Analysis. *Appl Psychol Meas*. 2019; 43: 415–416. [PubMed: 31235986]
118. Foley CN, et al. A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. *Nat Commun*. 2021; 12: 764. [PubMed: 33536417]
119. Berisa T, Pickrell JK. Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics*. 2016; 32: 283–285. [PubMed: 26395773]
120. Bowden J, Smith GD, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genetic Epidemiology*. 2016; 40: 304–314. DOI: 10.1002/gepi.21965 [PubMed: 27061298]
121. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015; 44: 512–525. [PubMed: 26050253]
122. Verbanck M, Chen C-Y, Neale B, Do R. Publisher Correction: Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet*. 2018; 50: 1196.
123. Morrison J, Knoblach N, Marcus JH, Stephens M, He X. Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. *Nat Genet*. 2020; 52: 740–747. [PubMed: 32451458]
124. Hemani G, Tilling K, Smith GD. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLOS Genetics*. 2017; 13 e1007081 doi: 10.1371/journal.pgen.1007081 [PubMed: 29149188]
125. Hemani G, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018; 7
126. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. *Int J Epidemiol*. 2014; 43: 922–929. [PubMed: 24608958]
127. Bryois J, et al. Genetic Identification of Cell Types Underlying Brain Complex Traits Yields Novel Insights Into the Etiology of Parkinson's Disease. doi: 10.1101/528463
128. Won H, et al. Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature*. 2016; 538: 523–527. [PubMed: 27760116]
129. Finucane HK, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet*. 2015; 47: 1228–1235. [PubMed: 26414678]
130. Finucane HK, et al. Heritability enrichment of specifically expressed genes identifies disease relevant tissues and cell types. *Nat Genet*. 2018; 50: 621–629. [PubMed: 29632380]
131. Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun*. 2019; 10: 1776. [PubMed: 30992449]
132. Warrier V, et al. Gene-environment correlations and causal effects of childhood maltreatment on physical and mental health: a genetically informed approach. *The Lancet Psychiatry*. 2021; 8: 373–386. [PubMed: 33740410]
133. Warrier V, et al. Genetic correlates of phenotypic heterogeneity in autism. *Nat Genet*. 2022; doi: 10.1038/s41588-022-01072-5

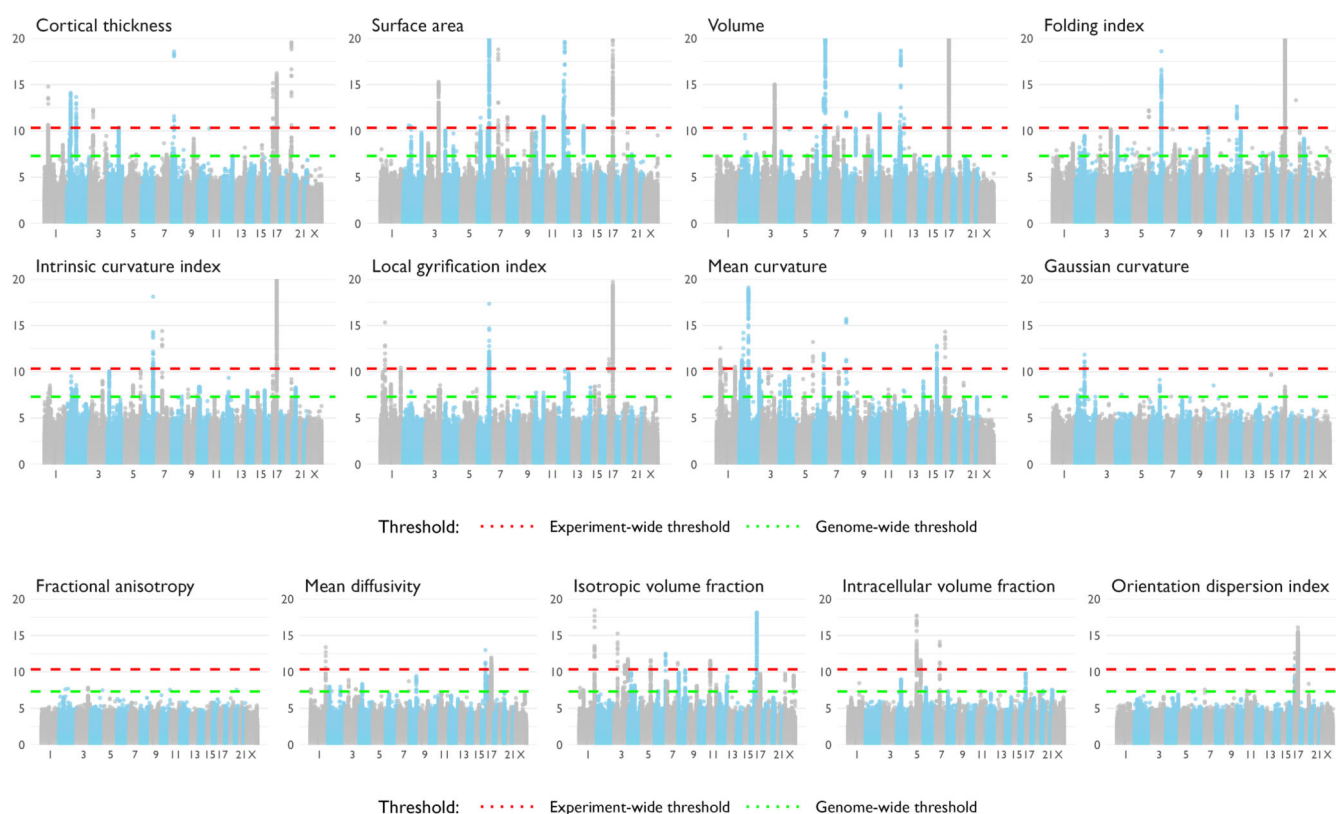


134. Wright CF, et al. Optimising diagnostic yield in highly penetrant genomic disease. *bioRxiv*. 2022; doi: 10.1101/2022.07.25.22278008
135. Wang G, Sarkar A, Carbonetto P, Stephens M. A simple new approach to variable selection in regression, with application to genetic fine mapping. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2020; 82: 1273–1300. DOI: 10.1111/rssb.12388 [PubMed: 37220626]
136. Hu B, et al. Neuronal and glial 3D chromatin architecture informs the cellular etiology of brain disorders. *Nat Commun*. 2021; 12: 3968. [PubMed: 34172755]
137. McLaren W, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016; 17: 122. [PubMed: 27268795]
138. Zhu Z, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet*. 2016; 48: 481–487. [PubMed: 27019110]
139. O'Brien HE, et al. Expression quantitative trait loci in the developing human brain and their enrichment in neuropsychiatric disorders. *Genome Biol*. 2018; 19: 194. [PubMed: 30419947]
140. Yang J, Qi T, Wu Y, Zhang F, Zeng J. Genetic control of RNA splicing and its distinctive role in complex trait variation. *Research Square*. 2021; doi: 10.21203/rs.3.rs-155233/v1
141. Qi T, et al. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. *Nat Commun*. 2018; 9: 2282. [PubMed: 29891976]



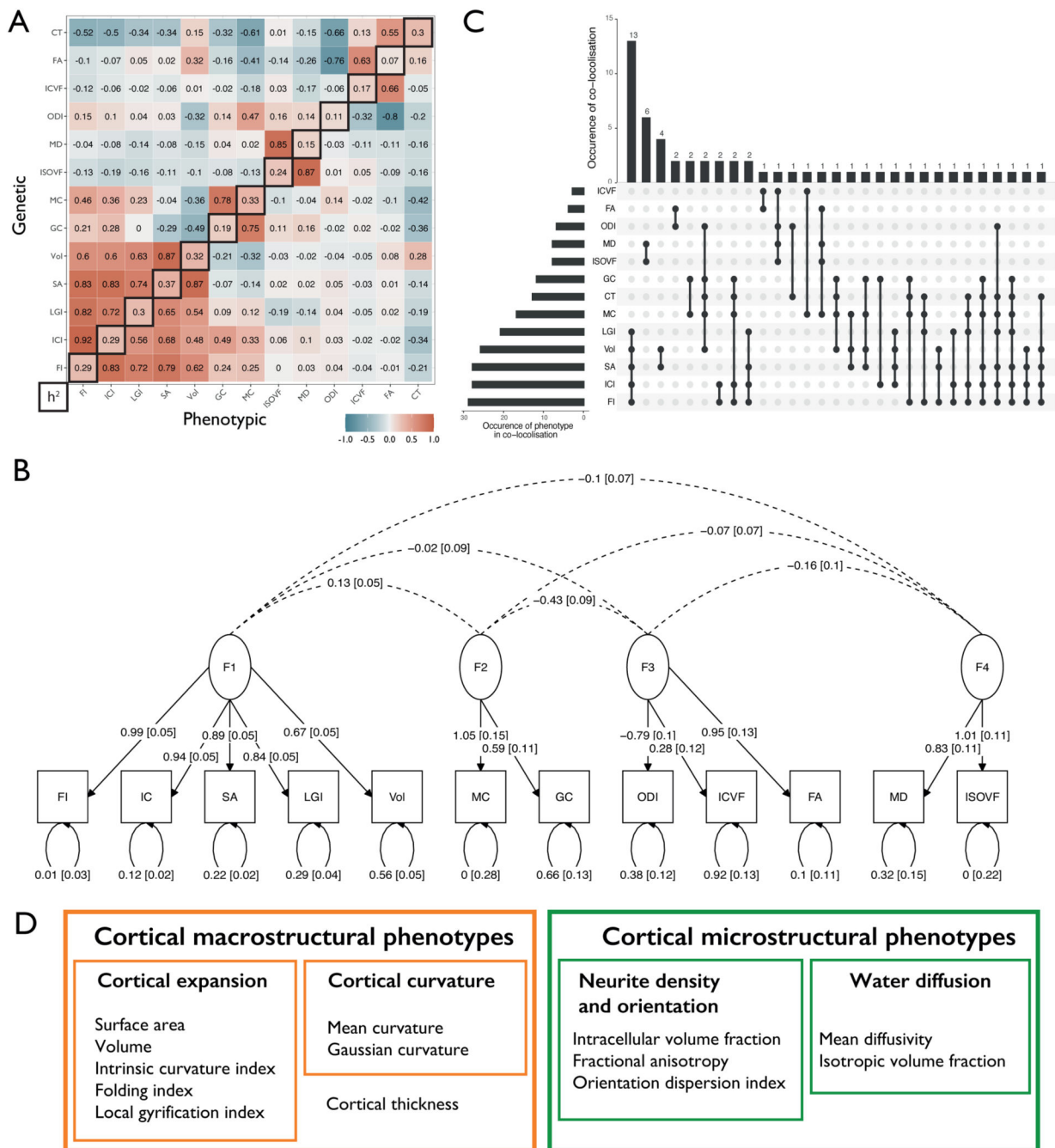
**Figure 1. Schematic overview of 13 brain MRI phenotypes and the genetic analyses.**

(A) We considered eight cortical macrostructural phenotypes: cortical thickness (CT), cortical surface area (SA), please note that this is for illustrative purposes and that the surface area is measured at midthickness, grey matter volume (volume), folding index (FI), intrinsic curvature index (ICI), local gyrification index (LGI), mean curvature (MC), and Gaussian curvature (GC). (B) We also considered five cortical microstructural phenotypes: fractional anisotropy (FA), mean diffusivity (MD), intracellular volume fraction (ICVF, also called Neurite Density Index [NDI]), isotropic volume fraction (ISOVF), and orientation dispersion index (ODI). (C) Each phenotype was measured globally (total or mean for the whole cortex) and regionally at each of 180 bilaterally averaged cortical regions defined by the Human Connectome Project parcellation scheme. We conducted genome-wide association studies of all phenotypes after removing outliers, and investigated latent structure of all phenotypes, developmental trajectories, and cell type specificity and genetic organisation.



**Figure 2. Manhattan plots of GWAS meta-analysis of 13 global MRI phenotypes.**

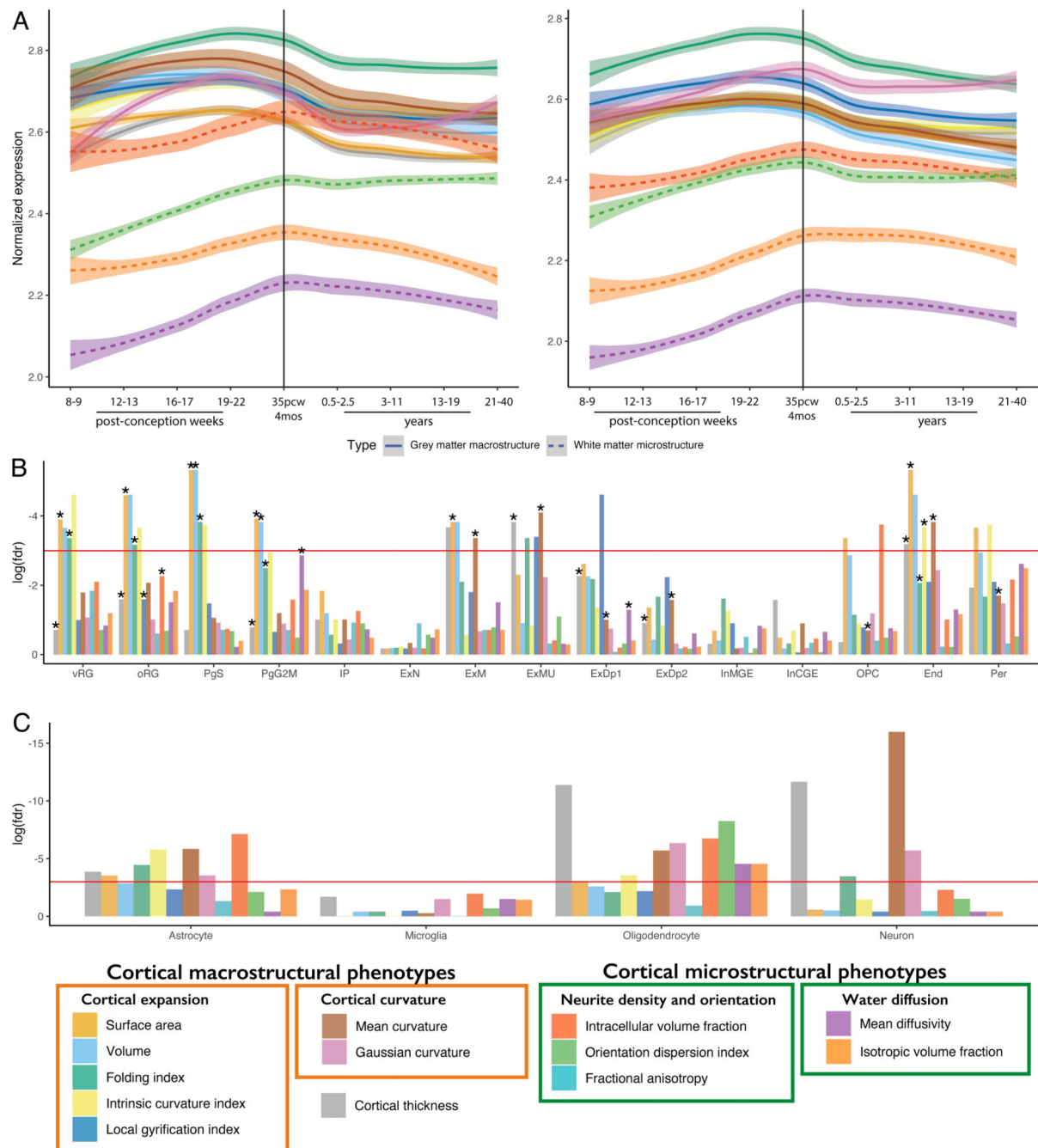
Green dotted line indicates the threshold for genome-wide significance ( $p = 5 \times 10^{-8}$ ), and the red dotted line indicates the threshold for experiment-wide significance ( $p = 4.58 \times 10^{-11}$ ). Each dot on the x axis indicates a SNP, and the y axis indicates the negative log<sub>10</sub> p-value.



**Fig 3. Pleiotropy among the 13 global phenotypes demonstrated by genetic/phenotypic correlations, structural equation modelling and co-localisation analysis.**

A: Phenotypic and genetic correlation matrices. The upper matrix triangle shows bivariate genetic correlations for each pair of phenotypes estimated using LDSC, the lower triangle shows the pairwise phenotypic correlations (Spearman's coefficient). The diagonal indicates the SNP heritability of each phenotype based on LDSC. Phenotypes are ordered based on hierarchical clustering of the genetic bivariate correlation (hierarchical clustering on the phenotypic correlation matrix resulted in a near identical ordering). B: Genomic SEM path diagram demonstrating the underlying latent structure of 12 of the 13 global

phenotypes and the interfactor genetic correlations. Covariance relationships = dotted arrows connecting two variables, variance estimates = double-headed arrows connecting variable to itself, regression relationships = one-headed arrows pointing from independent variable to dependent variable. Circles indicate latent variables, squares indicate measured phenotypes. The model was identified using unit variance identification such that the variance of the latent factors was set to one, and the dotted arrows across the factors can be interpreted as genetic correlation estimates. C: UpSet plot of the results of co-localisation analysis demonstrating the numbers of genomic loci that co-localise between the 13 phenotypes. The dots correspond to the co-localised clusters, with the number of clusters in the vertical bars. The number of times a phenotype co-localises is provided in the horizontal bars. D: Summary of clusters identified through GSEM and colocalisation analyses and their relationship with other terms used in this study. Abbreviations: cortical thickness (CT), cortical surface area (SA), grey matter volume (Vol), folding index (FI), intrinsic curvature index (ICI), local gyrification index (LGI), mean curvature (MC), and Gaussian curvature (GC), fractional anisotropy (FA), mean diffusivity (MD), intracellular volume fraction (ICVF), isotropic volume fraction (ISOVF), and orientation dispersion index (ODI).

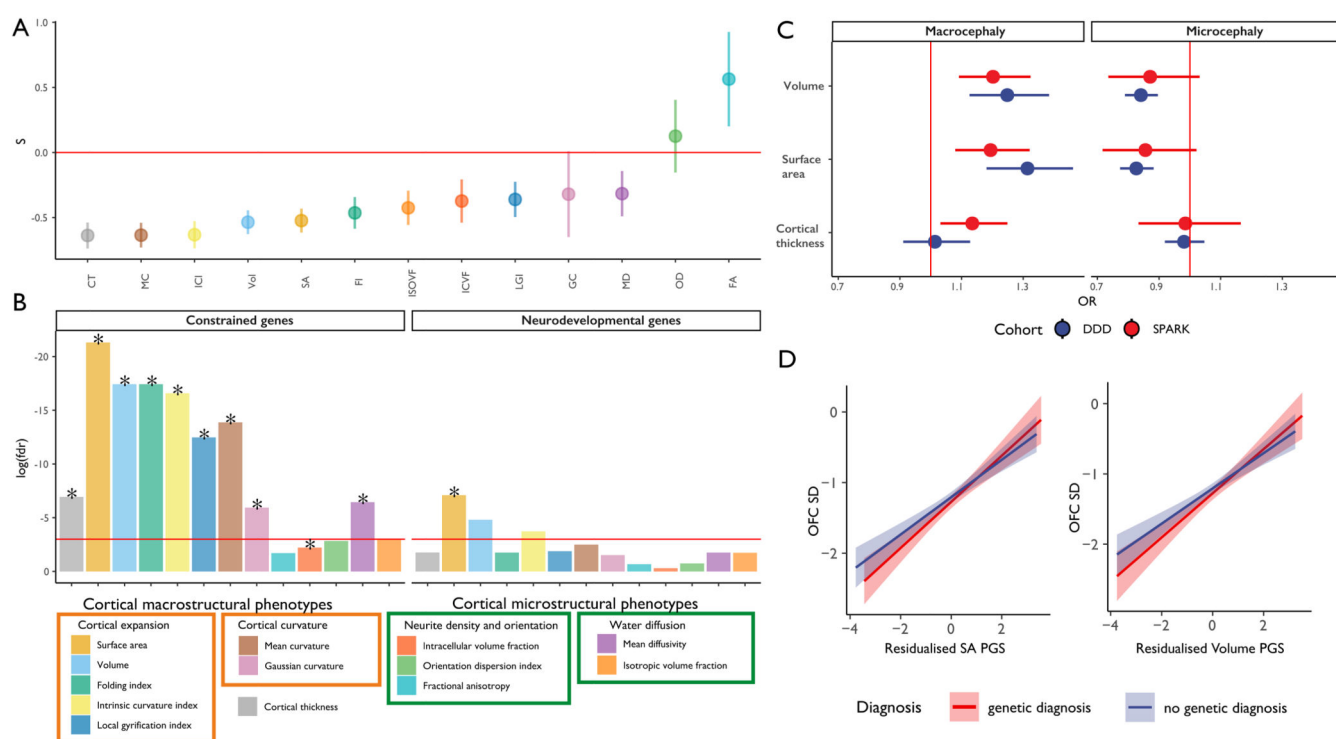


**Fig 4. Enrichment of GWAS signals in different cell-types during development**

A: Developmental trajectories of average gene expression in cortical postmortem-bulk RNA data (PsychEncode) for all significant genes ( $FDR < 0.05$ ) identified using HMAGMA (left panel) or MAGMA (right panel) for 12 of the 13 global phenotypes. Data for FA are not shown as too few genes were identified as significant. B: Results of enrichment analyses for cell-specific gene expression from mid gestation. FDR corrected  $\log_{10}$  p-values for gene enrichment using genes identified from MAGMA are plotted. The red line indicates the significance threshold at FDR corrected p-value  $< 0.05$ . Additionally,

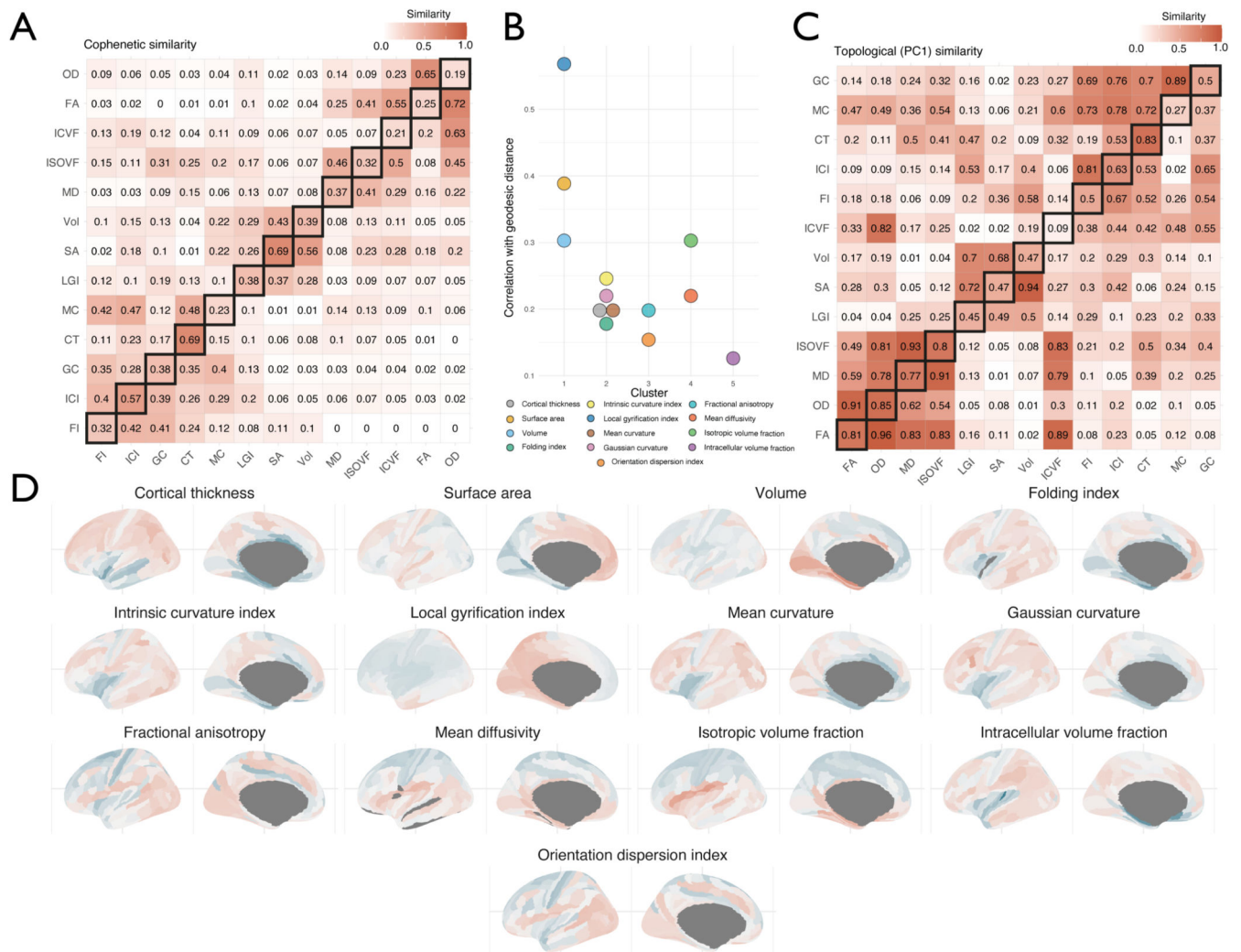


significant enrichments identified using HMAGMA genes are indicated with an asterisk C: Results of enrichment analyses from cell-specific epigenetic marks from postnatal cortex identified using LDSC-based enrichment. Abbreviations used: vRG = ventral radial glia, oRG = outer radial glia, PgS and PgG2M = cycling progenitors, S phase and G2-M phase respectively, IP = intermediate progenitors, ExN = Migrating excitatory neurons, ExM = Maturing Excitatory neurons, ExMU = Maturing excitatory neuron, upper enriched, ExDp1 = Excitatory deep layer neurons 1, ExDp2 = Excitatory deep layer neurons 2, InMGE = MGE Interneuron, InCGE = CGE interneuron, OPC = oligodendrocyte precursor cells, End = endothelial cells, Per = pericytes.



**Figure 5. Signatures of constraint and links to neurodevelopment for the global phenotypes.**

**A:** Estimates of selection for the 13 cortical phenotypes. Selection coefficients (S) are provided on the y axis (points). Bars indicate standard deviations for the selection coefficients. Negative values indicate that lower-MAF alleles tend to have larger effect sizes **B:** Results of the enrichment analyses for constrained genes and genes associated with neurodevelopmental disorders using genes identified from MAGMA. FDR corrected  $\log_{10}$  p-values for gene enrichment using genes identified from MAGMA are shown (y-axis). The red line indicates the significance threshold at FDR adjusted p-value = 0.05. Additionally, significant enrichments identified using HMAGMA genes are indicated with an asterisk **C:** Odds ratio (OR) and 95% confidence intervals for macrocephaly and microcephaly compared to individuals with neither for 1 standard deviation increase in polygenic scores for volume, surface area, and cortical thickness in the DDD and SPARK cohorts. **D:** Line plots demonstrating the linear association between genetic principal component corrected polygenic scores (surface area (SA) and volume) and standardised (compared to general population) occipital-frontal circumference (OFC SD) for individuals with or without a genetic diagnosis in the DDD cohort.



**Fig 6. Topographic similarity and principal component structure of cortical phenotypes.**

A: Cophenetic similarity matrix depicting the similarity between the region\*region similarity matrices. The upper triangle shows the cophenetic genetic similarity, the lower triangle shows the cophenetic phenotypic similarity and diagonals show the phenotype-genotype cophenetic similarity across features. B: Correlation between network topology and geodesic distance organised by hierarchical clustering of the cophenetic similarity. C: Spatial correlation between the first principal component of each regional similarity matrix. The upper triangle shows the genetic similarity, the lower triangle shows the phenotypic similarity and diagonals show the phenotype-genotype correlation across. D: Topology of the first genetic principal components, with colour depicting the relative PCA eigenvalues. The colour thus indicates to what extent regions show more homogenous similarity (i.e., regions with more similar colour have more similar covariance), but the actual sign and magnitude are relative within each phenotype.