NEUROGENOMICS

Discovery of genomic loci of the human cerebral cortex using genetically informed brain atlases

Carolina Makowski[1], Dennis van der Meer[2,3], Weixiu Dong[4], Hao Wang[1], Yan Wu[4]†, Jingjing Zou[5], Cin Liu[1], Sara B. Rosenthal[6], Donald J. Hagler Jr.[1], Chun Chieh Fan[1], William S. Kremen[7], Ole A. Andreassen[2], Terry L. Jernigan[8], Anders M. Dale[1,2], Kun Zhang[4], Peter M. Visscher[9], Jian Yang[9,10], Chi-Hua Chen[1]*

To determine the impact of genetic variants on the brain, we used genetically informed brain atlases in genome-wide association studies of regional cortical surface area and thickness in 39,898 adults and 9136 children. We uncovered 440 genome-wide significant loci in the discovery cohort and 800 from a post hoc combined meta-analysis. Loci in adulthood were largely captured in childhood, showing signatures of negative selection, and were linked to early neurodevelopment and pathways associated with neuropsychiatric risk. Opposing gradations of decreased surface area and increased thickness were associated with common inversion polymorphisms. Inferior frontal regions, encompassing Broca's area, which is important for speech, were enriched for human-specific genomic elements. Thus, a mixed genetic landscape of conserved and human-specific features is concordant with brain hierarchy and morphogenetic gradients.

Large-scale magnetic resonance imaging and genetics datasets have afforded the opportunity to discover common genetic variants contributing to the morphology of the human cortex. Studies in model organisms have revealed intricate genetic mechanisms underlying cortical area and thickness (i.e., laminar) patterning, although it has been challenging to define aspects of cortical development that are shared across mammals as opposed to those that are human-specific (*1*). Nevertheless, many studies have shown support for the radial unit hypothesis, which posits differential neurodevelopment programs shaping and regulating these two cortical measures (*2*).

Consistent with this, the ENIGMA (Enhancing Neuroimaging Genetics through Meta-Analysis) Consortium's genome-wide association study (GWAS) of the human cortex found many variants associated with surface area and thickness linked to neurodevelopmental processes during fetal development (*3*). Such evidence for neurodevelopmental programming indicates the need to investigate these questions at earlier ages, as previous cortical GWASs have almost exclusively been conducted in older adults.

Cortical expansion and regional patterning are largely genetically determined (*2*); therefore, we used data-driven genetically informed atlases in this study (*4*, *5*), rather than atlases primarily determined by sulcal-gyral patterns. These genetically determined atlases capture patterns of hierarchical genetic similarity following known developmental gradients that shape the cortex along their anterior-posterior (A-P) and dorsal-ventral (D-V) axes, including 12 surface area and 12 thickness regions (*2*, *4*, *5*), and increase discoverability of genetic variants underlying the cortex (*6*).

Results

Genetic variants underlying cortical thickness and area

In our discovery UK Biobank (UKB) sample of 32,488 individuals (table S1), we found 440 genome-wide significant [mixed linear model association tests (*7*), $P < 5 \times 10^{-8}$] variants after clumping each phenotype separately in PLINK (*8*) [linkage disequilibrium (LD) $R^2 = 0.1$, 250 kb], where 305 and 88 regional genetic variants were associated with the 12 surface area phenotypes and the 12 cortical thickness phenotypes, respectively (Fig. 1 and tables S2 and S3). Twenty-seven genetic variants were significantly associated with total surface area and 20 variants with mean cortical thickness (table S2). After correction for multiple comparisons, 234 genetic variants remained significant ($P < 2.27 \times 10^{-9}$, $5 \times 10^{-8}/t_e$, with $t_e = 22$ being the effective number of independent traits). We performed subsequent functional analyses for the 393 regional variants. Single-nucleotide polymorphisms (SNPs) were mapped to genes on the basis of their genomic position with FUMA (*9*). Across all phenotypes, SNPs were significantly enriched for noncoding regions (44.0% enriched for intronic variants, 33.4% for intergenic, and 17.7% for noncoding intronic RNA; Fisher's exact test, $P < 0.05$) (Fig. 1 and table S4).

Replication and generalization

Replication was performed on an admixed sample of 7410 individuals from UKB, including 2232 of European descent, using mixed linear model association (MLMA) analysis in GCTA (genome-wide complex trait analysis) (*10*). We modeled population structure using GENESIS (*11*) to estimate principal components and kinship. Estimated genetic effects in the discovery dataset were correlated with those in the replication dataset, as indexed by significant beta correlations (ranging from a correlation coefficient, *r*, of 0.66 to 0.95 after correcting for errors in the estimated SNP effects) (fig. S1), sign concordance rate (binomial test, $P < 0.05$), and proportion of variants replicated after multiple comparison correction (*12*).

MLMA and GENESIS were also used for generalization to data from 9136 individuals from the Adolescent Brain Cognitive Development (ABCD) Study (table S1), given the high degree of admixture and relatedness in this sample. Generalization to ABCD was quite high, as can be observed through significant beta correlations (*r* range: 0.46 to 0.92) (fig. S2), sign concordance rate, and proportion of variants replicated after correction for multiple comparisons (*12*). This suggests that the genetic architecture of the cortex found in adulthood is largely generalizable to earlier life stages of neurodevelopment, particularly for surface area. We also examined correspondence between the two datasets by calculating genetic correlations with LD score regression (LDSC) for each region. Eighteen of 24 phenotypes were significantly genetically similar between ABCD and UKB (genetic correlation, r_g, range: 0.38 to 1.21) (fig. S3).

Given the evidence of comparable results, we ran a joint meta-analysis of the three samples using METAL (*12*). After clumping each phenotype separately to obtain independent loci, the meta-analysis revealed 800 genome-wide significant regional loci, with 467 passing correction for multiple comparisons (table S5). Of 800 loci, 526 were found to be independent by merging hits from these 26 phenotypes into one file and clumping with PLINK (*8*) (R^2 = 0.1, 250 kb). With the exception of one SNP, all had a nonsignificant heterogeneity *P* value ($P > 1 \times 10^{-6}$) associated with Cochran's Q statistic, suggesting comparability among samples. SNPs from the meta-analysis with a significant heterogeneity $P < 1 \times 10^{-6}$ are listed in table S6.

[1]Center for Multimodal Imaging and Genetics, University of California, San Diego, CA, USA. [2]Norwegian Centre for Mental Disorders Research (NORMENT), Division of Mental Health and Addiction, Oslo University Hospital and Institute of Clinical Medicine, University of Oslo, Oslo, Norway. [3]School of Mental Health and Neuroscience, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, Netherlands. [4]Department of Bioengineering, University of California, San Diego, CA, USA. [5]Division of Biostatistics, Herbert Wertheim School of Public Health and Human Longevity Science, University of California, San Diego, CA, USA. [6]Center for Computational Biology and Bioinformatics, University of California, San Diego, CA, USA. [7]Department of Psychiatry and Center for Behavior Genetics of Aging, University of California, San Diego, CA, USA. [8]Center for Human Development, University of California, San Diego, CA, USA. [9]Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia. [10]School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China.
*Corresponding author. Email: chc101@ucsd.edu (C.-H.C.)
†Present address: Cellarity, Somerville, MA, USA.

Comparison to previous cortical GWAS

We used conditional and joint analysis (COJO) (*13*) to identify novel loci compared with the most recent GWAS of cortical architecture which identified 369 loci (*3*). Of these loci, 206 were found to be independent by clumping all 70 phenotypes together (R^2 = 0.1, 250 kb). COJO revealed that 63.6% of our 393 regional variants remained genome-wide significant and thus are considered novel associated variants (*12*) (table S7).

Assigning SNPs to genes and neuropsychiatric implications

All SNPs in LD (R^2 > 0.6) with the 393 regional variants were mapped to genes using positional, gene expression [expression quantitative trait locus (eQTL)], and chromatin interaction information in FUMA SNP2GENE (*9*). This mapped our genetic variants to 915 genes (tables S8 and S9). MAGMA gene-based analyses yielded 575 significant genes (mean χ^2 statistics, $P < 2.6 \times 10^{-6}$) (table S10). According to the National Institutes of Health Genetics Home Reference, many significant genes are related to neurodevelopmental disorders (autism, epilepsy, microcephaly) or dementia (table S11).

Further support for this conclusion was determined by investigating the shared genetic effects between our brain phenotypes and disorders by estimating genetic correlations through LDSC (fig. S4 and table S12). We found a significant association between global surface area and attention-deficit/hyperactivity disorder (ADHD) after multiple comparison correction, as well as nominal significant associations [e.g., temporal area with schizophrenia and autism spectrum disorder (ASD)]. To examine putative causal association, we performed Mendelian randomization (*14*) on global area and ADHD that showed the most significant r_g, and we did not find evidence of causality. We also examined ASD, a neurodevelopmental disorder with early onset, and its relationship with anteromedial temporal area indexed by a significant r_g. We found a significant unidirectional causation (b_{xy} = −0.36, $P = 9.5 \times 10^{-5}$), indicating that decreased anteromedial temporal area may cause ASD. These SNPs could be missed in classical GWAS of ASD, but nevertheless are important genetic factors in the pathogenesis

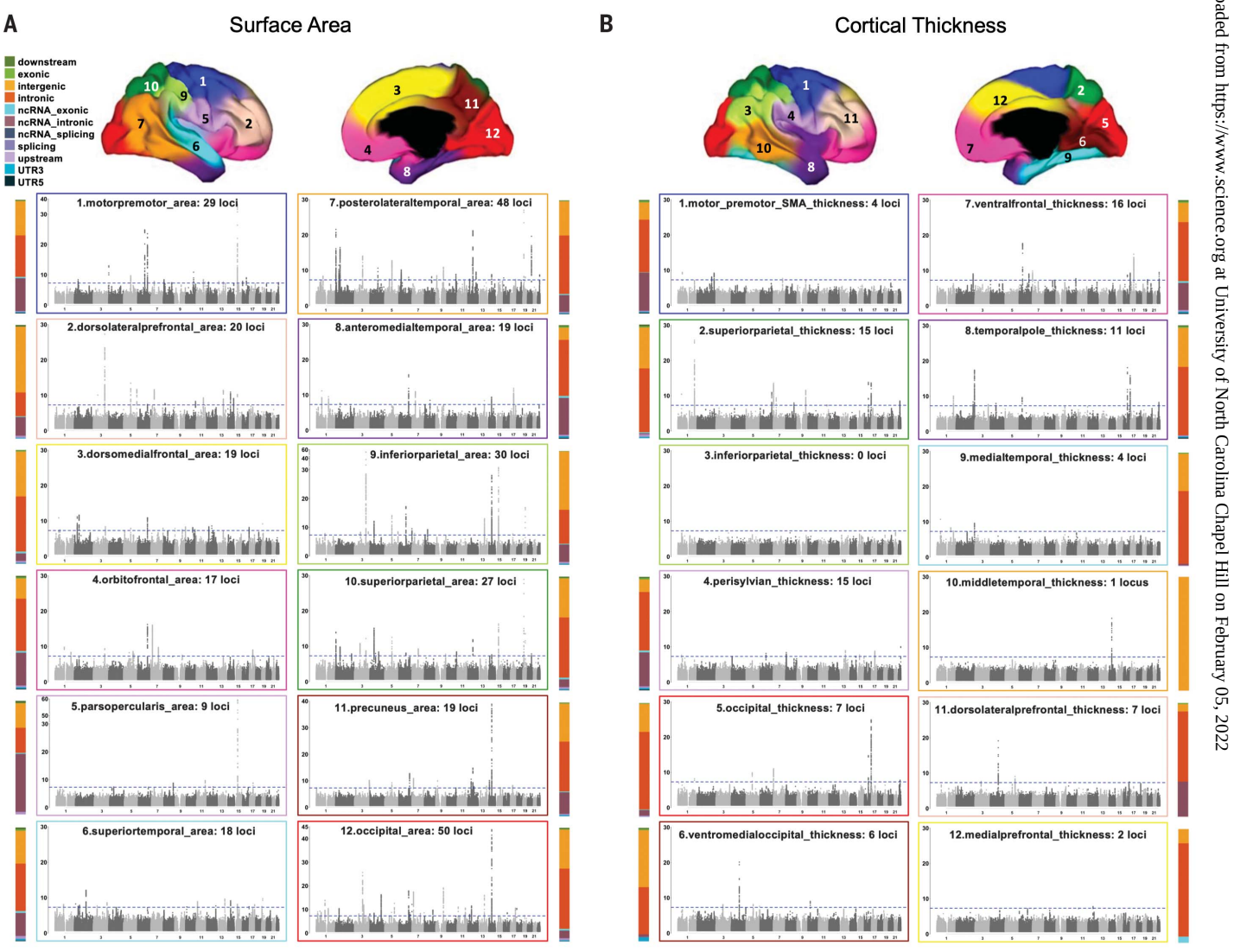

Fig. 1. Manhattan plots of genetic variants underlying surface area and cortical thickness. Results are shown separately for surface area (**A**) and cortical thickness (**B**). Numbers on brain atlases represent each brain region. Plots are color coded by brain atlas region. Number of significant genetic loci are listed in Manhattan subplot titles, with the horizontal dashed line denoting genome-wide significance. Vertical bar charts show breakdown of genomic position of SNPs, with corresponding legend at the top of (A). ncRNA, noncoding RNA; UTR3, 3′ untranslated region; UTR5, 5′ untranslated region.

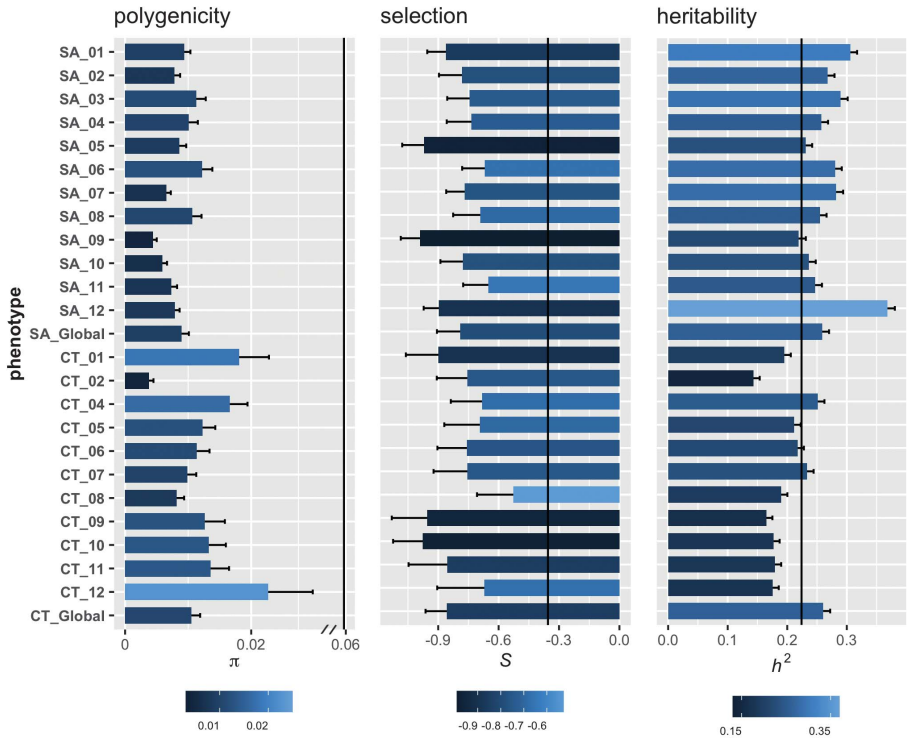

Fig. 2. Genetic architecture of the cortex. Cortical phenotypes generally have low polygenicity, medium to high heritability, and are under strong negative selection. Vertical black lines on each plot are average reference lines for relevant estimates of commonly studied traits taken from (*16*). Numbering of regions follows labels in Fig. 1. SA, surface area; CT, cortical thickness; π, polygenicity; *S*, selection; h^2, heritability.

of this disorder through their contributions to anteromedial temporal morphology.

Genetic architecture of the cortex

Compared with other common complex traits, cortical phenotypes tend to have low polygenicity (proportion of genome-wide SNPs with non-null effects; range: 0.0038 to 0.040; area: $\pi = 0.0085 \pm 0.0011$; thickness: $\pi = 0.015 \pm 0.0039$) and average-to-high SNP-based heritability (range: 0.14 to 0.37; area: $\overline{h^2} = 0.27 \pm 0.012$; thickness: $\overline{h^2} = 0.20 \pm 0.011$) (Fig. 2). Pedigree-based heritability for the UKB discovery sample (range: 0.31 to 0.95), calculated with multiple genetic relatedness matrices (*15*), and twin-based heritability approximated by Falconer's formula from the ABCD sample (range: 0.39 to 0.96) can be found in table S13. Negative selection signatures can be inferred from the relationship between minor allele frequency and effect size, quantified by the *S* parameter implemented in SBayesS (*16*) (fig. S5). We found that loci associated with our cortical phenotypes may be under strong negative selection pressures (*16*) compared with phenotypes with similar levels of heritability and polygenicity (range: −0.99 to 0.045; area: $\overline{S} = -0.79 \pm 0.11$; thickness: $\overline{S} = -0.72 \pm 0.18$). It should be noted that π is slightly dependent on sample size and thus should be interpreted with caution. However, others have shown similar estimates of polygenicity for brain phenotypes (*17*).

Partitioned heritability

Different functional regions of the genome can contribute disproportionately to complex human traits. Thus, we applied stratified LDSC regression to partition heritability estimates of our 24 cortical phenotypes for 97 annotations from the baseline model (*18*, *19*) (table S14), from which we focused on enriched annotations where regression coefficients are significantly positive ($z > 1.96$, two-tailed $P < 0.05$). We classified the annotations into three categories determined from conserved, developmental, and regulatory genomic partitions. We found seven conserved annotations (found in primates and other mammals) to be significantly enriched after multiple comparison correction ($P < 0.0025$, where $P < 0.05/t_e$) (Fig. 3A and table S15) across 16 cortical phenotypes, with notable enrichment for seven phylogenetically conserved cortical regions (e.g., medial temporal lobe, motor and orbitofrontal regions) for the annotation "ancient sequence age human promoter." This conserved promoter annotation reflects a genomic region that is evidenced to have existed before the evolutionary split of marsupial and placental mammals (*18*).

Seven regions, mostly indexing surface area, were significantly enriched for developmental annotations of fetal deoxyribonuclease I (DNAse I) hypersensitive sites (DHSs), a marker of accessible chromatin (*20*), along with enrichment of 15 cortical phenotypes for 13 regulatory annotations (table S15). We performed an additional partitioned heritability analysis using differential methylation regions (DMRs) that were previously found to be associated with present-day humans compared to Neanderthal and Denisovan genomes (*21*). Perisylvian thickness was nominally enriched for present-day human DMRs (LDSC Jackknife test, $P = 0.03$). By partitioning the genome into meaningful functional categories, we capture patterns of hierarchical brain organization with evolutionarily conserved (paralimbic, sensory motor) regions enriched for conserved and developmental annotations and association areas more strongly associated with regulatory annotations.

Gene Ontology enrichment

To elucidate the biological pathways associated with our discovered genetic variants, MAGMA-mapped genes were input into the Molecular Signatures Database to obtain Gene Ontology (GO) terms. Twenty-six GO terms, predominantly related to neurodevelopment, were significantly associated with our brain phenotypes after Bonferroni correction (Fig. 3B and table S16). Notable biological pathways included *WNT*/beta-catenin, *TCF*, *FGF*, and hedgehog signaling, which are important for axis specification and areal identity (*1*). For higher-order association regions, the dorsolateral prefrontal cortex was linked to cortical tangential migration.

Three-dimensional genetic characterization of the cortex

To better understand the relationship between our cortical phenotypes, we computed phenotypic and genetic correlation matrices using LDSC (Fig. 4A). Significant correspondence was observed between matrices (Mantel test: $r = 0.85$, $P = 0.001$), suggesting substantial genetic influences on cortical patterning. Hierarchical clustering was applied to genetic correlations of area and thickness separately, revealing a clear separation in genetic architecture between A-P divisions in area and between D-V divisions in thickness (Fig. 4, A and B, and fig. S6). Regions anatomically closer to each other tended to be more correlated with each other. However, homologous regions in contralateral hemispheres had high genetic correlations despite their physical distance (*4*) (table S18).

Given the observed correlations, we sought to estimate the shared genetic effects across phenotypes with genomic structural equation

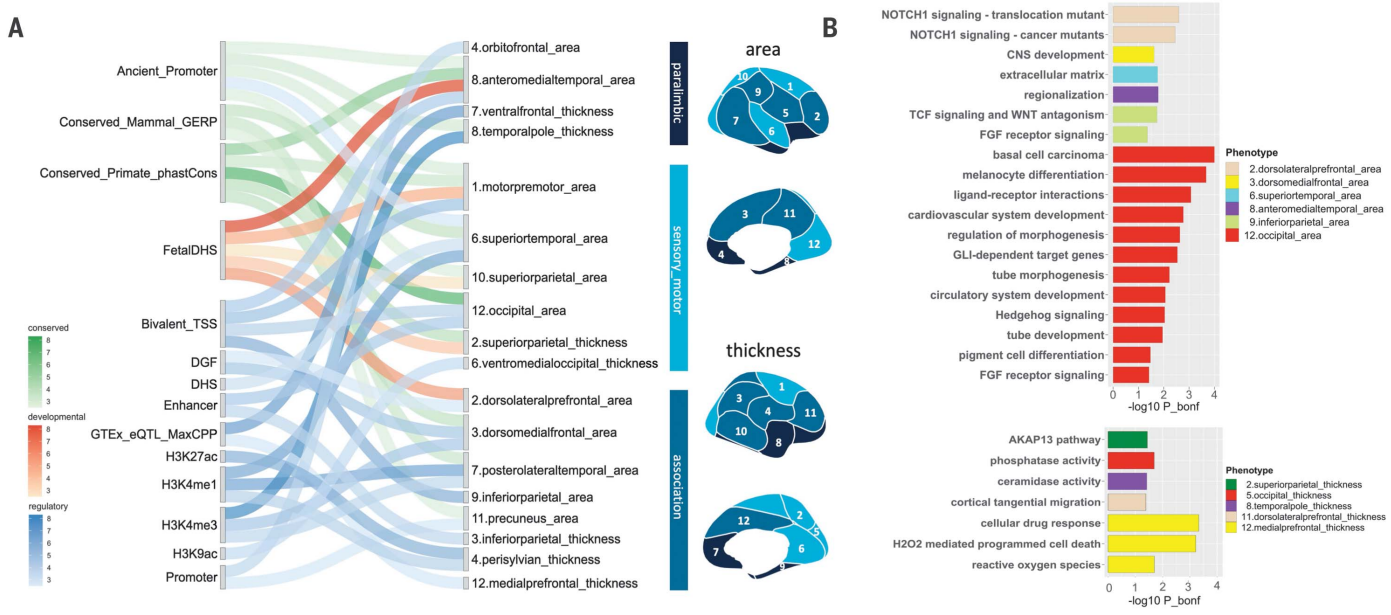

Fig. 3. Partitioned heritability and Gene Ontology (GO) enrichment. (**A**) Heritability of cortical phenotypes is significantly enriched for conserved, developmental, and regulatory annotations. The river plot depicts mapping between significant annotations (*18*) and cortical phenotypes. Color coding of the river plot is based on $-\log_{10}$ enrichment *P* values. (**B**) Significantly enriched GO terms from MAGMA gene set analysis for surface area (top) and cortical thickness (bottom). GERP, genomic evolutionary rate profiling; TSS, transcription start site; DGF, digital genomic footprint; GTEx, genotype-tissue expression; CNS, central nervous system; TCF, T cell factor; FGF, fibroblast growth factor; bonf, Bonferroni.

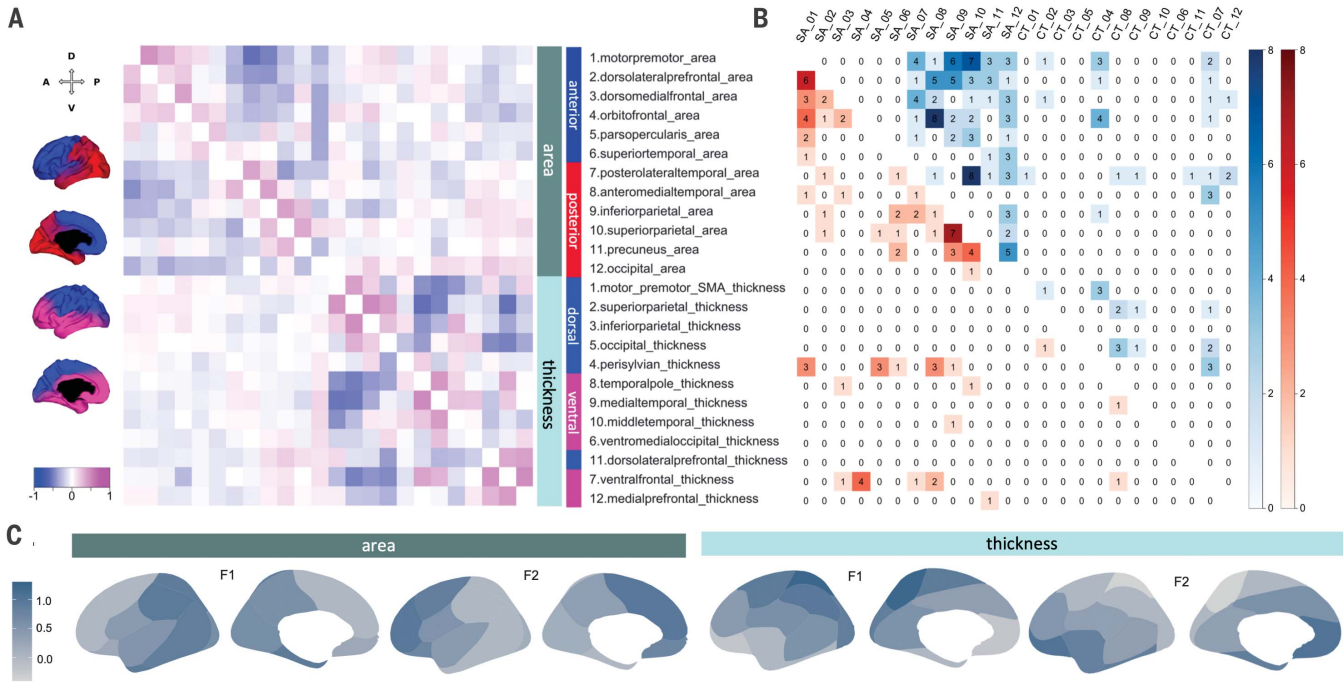

Fig. 4. Three-dimensional genetic characterization of the cortex. (**A**) Phenotypic and genotypic correlations between 24 regions, ordered by hierarchical clustering that shows A-P divisions for area and D-V divisions for thickness. Phenotypic correlations are in bottom left triangle, and genetic correlations are in the upper right. SMA, supplementary motor area. (**B**) Pleiotropic SNP counts for each pair of regions, using the same ordering as in (A). Agonistic or same direction of effects are in the lower red triangle, antagonistic or opposing effects are in upper blue triangle. (**C**) Brain maps of standardized effects of each latent factor (F1 and F2) derived from genomic SEM on each brain region. See fig. S7 for detailed statistics.

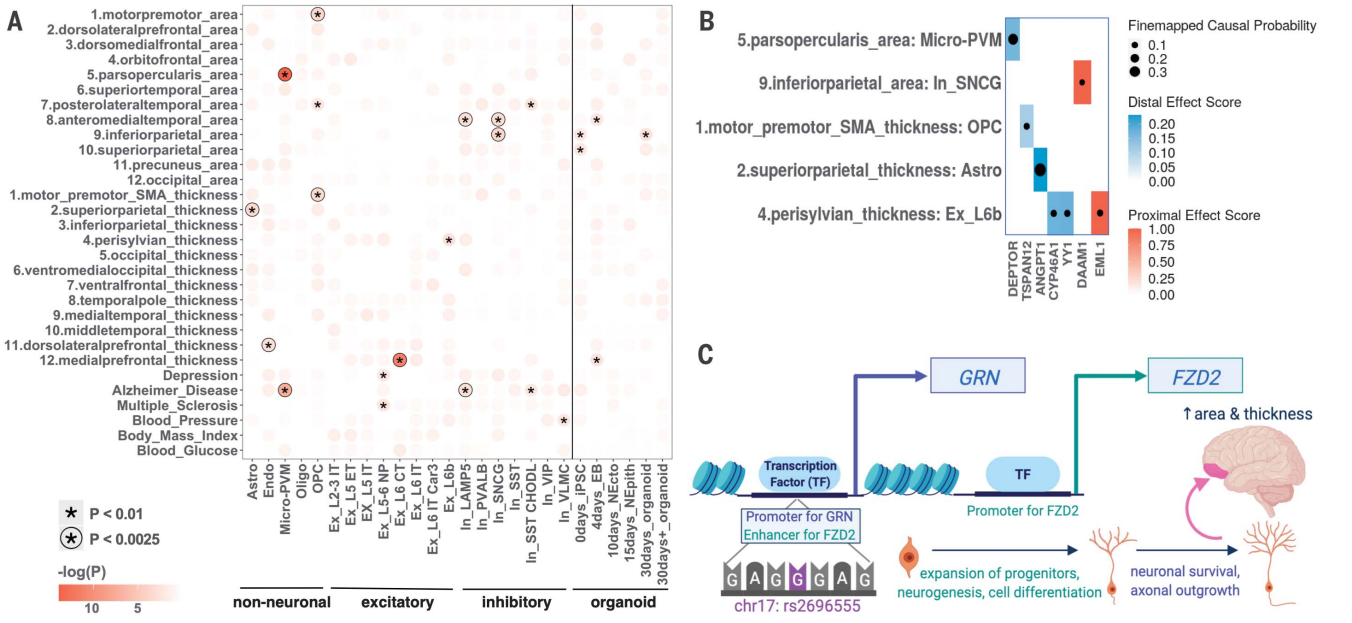

Fig. 5. Enrichment of cell type–specific accessible chromatin sites and fine-mapping to regulatory regions of genes. (**A**) Heatmap of enrichment for cortical phenotypes and cell type–specific accessible chromatin peaks. Phenotypes also include three metabolic (blood glucose, body mass index, and blood pressure) and three cortical-related (multiple sclerosis, Alzheimer's disease, and depression) controls. Vertical black line differentiates M1 cell types (left) from organoid developmental stages (right). Significant values are based on the bias-corrected enrichment statistic from g-chromVAR (*12*). (**B**) Mapped genes and the regulatory region (blue, enhancer; red, promoter) of the causal SNPs carried forward by positively enriched M1 cell type–cortical phenotype pairs ($z > 2.36$, $P < 0.01$). Size of dot reflects probability of SNP being causal. Colors represent peak to gene coaccessibilities, where a score of 1 reflects a peak being in the gene's promoter region. (**C**) A selected pleiotropic SNP (rs2696555) influencing both orbitofrontal area and ventral frontal thickness, mapped to target genes on the basis of coaccessibility with M1. Cell types are outlined in table S23.

modeling (SEM) (Fig. 4C and figs. S7 and S8) (*12*, *22*). We found that two-factor models fit our data well (comparative fit index of >0.98). The two latent factors recapitulated the A-P and D-V gradations of cortical patterning for area and thickness, respectively. The strongest association signals between the latent factors and variants reside in the 17q21.31 inversion region for area ($P < 1.48 \times 10^{-56}$), and more widespread effects across the genome with notable peaks on chromosomes 3 and 17 for thickness ($P < 3.39 \times 10^{-15}$) (fig. S8). We further performed association testing of inversion polymorphisms on 17q21.31 with our cortical phenotypes (table S19). We found the inverted allele to be highly associated with overall surface area reductions, with stronger effects in posterior regions along the A-P gradient and a modest positive correlation with increasing thickness in ventral regions. The opposing effects on area and thickness may in part account for the observation of a modest negative association between area and thickness ("cortical stretching") after accounting for total brain size (*23*).

After extracting salient latent factors underlying multiple brain regions, we searched for pleiotropic loci between pairs of regions. We used COJO to map SNPs with potential pleiotropic effects (i.e., that influence two regions), defined by the loci of region i that were no longer genome-wide significant when conditioned on the loci of region k (*13*). Using this approach, we found that 107 of our 393 loci had pleiotropic effects on two phenotypes (Fig. 4B and table S20). Surface area of parietal and posterolateral temporal regions shared eight SNPs with antagonistic effects (i.e., increasing area of one region while decreasing area of the other); these regions show good correspondence between ABCD and UKB and are both enriched for fetal DNAse hypersensitive sites (Fig. 3A). Two of these antagonistic SNPs, rs10878269 and rs142166430, are intronic variants of methionine sulfoxide reductase B3 (*MSRB3*), a gene that is important for protein repair and metabolism (*24*).

We also noted antagonistic pleiotropic effects of two SNPs, rs12676193 and rs6986885, in the 8p23.1 inversion polymorphism linked to motor-premotor area and perisylvian thickness (table S20). These SNPs were mapped to methionine sulfoxide reductase A (*MSRA*), a gene that is important for repair of oxidatively damaged proteins (*25*). Further, the 8p23.1 region is considered to be a potential hub for neurodevelopmental and psychiatric disorders (*26*). Another notable SNP with pleiotropic effects was rs888812, with antagonistic effects on precuneus and prefrontal area. This and other variants were mapped to *NR2F1*/*COUP-TF1*, a transcription factor influencing A-P patterning of the cortex in development (*1*).

Enrichment of cell type–specific accessible chromatin sites and fine-mapping to regulatory regions of genes

To map putative causal genes for our genetic variants—motivated by observed enrichment of our phenotypes for regulatory genomic regions—we computed cell type–specific enrichment for our fine-mapped GWAS SNPs on the basis of high-resolution accessible chromatin sites drawn from human primary motor cortex (M1) (*27*) and cerebral organoid data (*28*) using g-chromVAR (fig. S9). To quantify enrichment, we computed the accessibility deviations as the expected number of feature counts per peak per cell type, weighted by the fine-mapped variant posterior probabilities. This revealed 11 significantly positively enriched cell type–phenotype pairs after Bonferroni correction ($z > 2.8$, $P < 0.0025$) (Fig. 5A), including enrichment of the motor-premotor region for accessible chromatin sites in oligodendrocyte

precursor cells (OPCs). This result is particularly compelling given that OPCs give rise to mature oligodendrocytes which in turn myelinate axons in the central nervous system, and the motor cortex is known to be a region rich in intracortical myelin content (*29*). In control analyses, no significant enrichment was found for metabolic traits, suggesting that this approach is specific to cortical phenotypes. This approach is further supported by the consistent finding of the significant Alzheimer's-microglia pair (*30*).

For each significant M1 cell type–phenotype pair from Fig. 5A, we identified putative causal genes from a locus's genomic position relative to its gene targets and chromatin coaccessibility relationships (i.e., both the genomic locus and its gene target were simultaneously accessible). From the initial 25 target genes, five distal and two proximal genes remained (Fig. 5B) after filtering out genes with weak evidence of gene expression in the corresponding cell type (fig. S10 and table S21).

We applied the same mapping approach to pleiotropic SNPs and found three SNPs that overlapped with the M1 accessible chromatin peaks (Fig. 5 and table S22). Notably, rs2696555, a SNP in the 17q21.31 inversion region, was associated with increases in orbitofrontal area and ventral frontal thickness and mapped to the promoter region of *GRN*, a granulin precursor that helps preserve neuronal survival, axonal outgrowth, and neuronal integrity through its impact on inflammatory processes in the brain (*31*). This SNP was also mapped at a distal putative enhancer site of *FZD2*, which encodes a Frizzled receptor within the *WNT*/beta-catenin pathway and is expressed in cortical progenitor cells of the dorsal and ventral telencephalon of the developing brain (*32*). A schematic of how this single variant could influence area and thickness is depicted in Fig. 5C.

Discussion

This study advances understanding of the genetic architecture underlying the organization of the cerebral cortex and uniquely human traits. Our genetically informed atlases enhanced discovery of significant loci compared with previous cortical GWAS with traditional nongenetic atlases (*3*, *6*). The improved discovery is likely aided by the fact that our atlases conform to genetic cortical patterning (*4*, *5*), thereby increasing discoverability and heritability, while also having lower polygenicity.

Making use of two large cohorts of adults and children, we found that many genetic variants in our findings pinpoint genetic mechanisms influencing cortical patterning of the human brain in early development. Our data, particularly findings with *COUP-TF1*, support the protomap hypothesis whereby genes hold spatial and temporal instructions to initiate a cortical map by graded signaling from patterning centers in early development (*1*, *2*). Our results are consistent with reports of loss of *COUP-TF1* function leading to expansion of frontal motor areas at the expense of posterior sensory areas in the rodent brain (*1*), which is intriguing given the challenges in defining rodent-specific versus human-specific developmental mechanisms. These variants are promising candidates for future functional experiments.

We also uncovered latent factors describing our area phenotypes, suggesting genetic effects related to inversion polymorphisms. Recurrent inversions of genomic regions, such as 17q21.31 identified here along with 8p.23, have occurred through primate evolution and show that the inverted orientation is the ancestral state. Specifically, both 17q.21.31 and 8p.23 inversions appear to have occurred independently within the *Homo* and *Pan* lineages (*33*, *34*). 17q21.31 inversion contains microtubule associated protein tau (*MAPT*), a risk gene for neurodegeneration (*35*). The inverted (minor) allele has been associated with lower susceptibility for Parkinson's dementia but higher predisposition to developmental disorders (*33*).

We linked several of our findings to the *WNT*/beta-catenin pathway, which regulates cortical size by controlling whether progenitors continue to proliferate or exit the cell cycle to differentiate (*36*). Cell proliferation is thought to exponentially enlarge the progenitor pool and the number of cortical columns, which results in expansion of cortical surface area and gyrification. On the other hand, cortical thickness is largely determined by cell differentiation and a linear production of neurons within each cortical column (*2*, *36*). In addition to 17q21.31, our results revealed loci linked to various cortical regions in this pathway (e.g., *WNT3*, *GSK3B*), and their combined interactive effects may be differentially involved in shaping area and thickness.

The brain is particularly vulnerable to insults (genetic and environmental) during sensitive periods of neurodevelopment, and changes during this time can have lasting impacts on the brain over the life span. This perspective helps situate our findings of predominantly negative selection acting on our cortical phenotypes (Fig. 2), which may be linked to conserved genomic loci and those enriched for neuropsychiatric diseases (*18*, *19*). Here we uncovered a putative causal relationship of reduced anteromedial temporal area potentially giving rise to ASD. The medial temporal lobe has been linked to abnormal connectivity in some types of ASD and houses structures (e.g., amygdala, hippocampus) important in regulating emotion and social behaviors (*37*). We also found this region to be enriched for accessible chromatin sites in inhibitory neurons; thus, these findings may provide clues to the long-standing theory of excitatory-inhibitory imbalance in ASD (*38*).

Intriguingly, most of our phenotypes, especially paralimbic and sensory motor regions, exhibited enriched heritability for conserved genomic partitions (Fig. 3A) including promoter regions, rather than enhancers, consistent with the idea that the former are more evolutionarily conserved (*18*). However, we also identified brain regions that have evolved to support human-specific behaviors, such as language and communication. Differential methylation and human-specific SNPs in association with perisylvian thickness lead us to speculate that altered morphology of the perisylvian region, and potentially also motor-premotor regions, were important in the evolution of speech articulation (*39*).

Our results with genetically informed atlases demonstrate that human brain arealization and regionalization largely arise from phylogenetically conserved regions and multiple neurodevelopmental programs, but that a select few regulatory features, some of which may be specific to modern-day humans, have had widespread downstream effects on brain morphology and may have given rise to human-specific traits and diseases.

REFERENCES AND NOTES

1. Z. Molnár *et al.*, *J. Anat.* **235**, 432–451 (2019).
2. P. Rakic, *Nat. Rev. Neurosci.* **10**, 724–735 (2009).
3. K. L. Grasby *et al.*, *Science* **367**, eaay6690 (2020).
4. C.-H. Chen *et al.*, *Science* **335**, 1634–1636 (2012).
5. C.-H. Chen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17089–17094 (2013).
6. D. van der Meer *et al.*, *Cereb. Cortex* **30**, 5597–5603 (2020).
7. L. Jiang *et al.*, *Nat. Genet.* **51**, 1749–1755 (2019).
8. S. Purcell *et al.*, *Am. J. Hum. Genet.* **81**, 559–575 (2007).
9. K. Watanabe, E. Taskesen, A. van Bochoven, D. Posthuma, *Nat. Commun.* **8**, 1826 (2017).
10. J. Yang, S. H. Lee, M. E. Goddard, P. M. Visscher, *Am. J. Hum. Genet.* **88**, 76–82 (2011).
11. S. M. Gogarten *et al.*, *Bioinformatics* **35**, 5346–5348 (2019).
12. See supplementary materials.
13. J. Yang *et al.*, *Nat. Genet.* **44**, 369–375, S1–S3 (2012).
14. Z. Zhu *et al.*, *Nat. Commun.* **9**, 224 (2018).
15. N. Zaitlen *et al.*, *PLOS Genet.* **9**, e1003520 (2013).
16. J. Zeng *et al.*, *Nat. Genet.* **50**, 746–753 (2018).
17. N. Matoba, M. I. Love, J. L. Stein, *Hum. Brain Mapp.* **43**, 329–340 (2022).
18. M. L. A. Hujoel, S. Gazal, F. Hormozdiari, B. van de Geijn, A. L. Price, *Am. J. Hum. Genet.* **104**, 611–624 (2019).
19. S. Gazal *et al.*, *Nat. Genet.* **50**, 1600–1607 (2018).
20. G. Trynka, S. Raychaudhuri, *Curr. Opin. Genet. Dev.* **23**, 635–641 (2013).
21. D. Gokhman *et al.*, *Science* **344**, 523–527 (2014).
22. A. D. Grotzinger *et al.*, *Nat. Hum. Behav.* **3**, 513–525 (2019).
23. L. J. Hogstrom, L. T. Westlye, K. B. Walhovd, A. M. Fjell, *Cereb. Cortex* **23**, 2521–2530 (2013).
24. T.-J. Kwon *et al.*, *Hum. Mol. Genet.* **23**, 1591–1601 (2014).
25. A. N. Minniti *et al.*, *Antioxid. Redox Signal.* **22**, 48–62 (2015).
26. M.-T. Lo *et al.*, *Nat. Genet.* **49**, 152–156 (2017).
27. T. E. Bakken *et al.*, *Nature* **598**, 111–119 (2021).
28. S. Kanton *et al.*, *Nature* **574**, 418–422 (2019).
29. M. F. Glasser, D. C. Van Essen, *J. Neurosci.* **31**, 11597–11616 (2011).
30. B. B. Lake *et al.*, *Nat. Biotechnol.* **36**, 70–80 (2018).
31. X. Gao *et al.*, *Protein Cell* **1**, 552–562 (2010).
32. S. J. Harrison-Uy, S. J. Pleasure, *Cold Spring Harb. Perspect. Biol.* **4**, a008094 (2012).
33. M. C. Zody *et al.*, *Nat. Genet.* **40**, 1076–1083 (2008).
34. M. P. A. Salm *et al.*, *Genome Res.* **22**, 1144–1153 (2012).

35. K. H. Strang, T. E. Golde, B. I. Giasson, *Lab. Invest.* **99**, 912–928 (2019).
36. A. Chenn, C. A. Walsh, *Science* **297**, 365–369 (2002).
37. P. Rane *et al.*, *Harv. Rev. Psychiatry* **23**, 223–244 (2015).
38. G. Cellot, E. Cherubini, *Front Pediatr.* **2**, 70 (2014).
39. D. Gokhman *et al.*, *Nat. Commun.* **11**, 1189 (2020).
40. Adolescent Brain Cognitive Development Study (ABCD) 2.0.1 release #721, NIMH Data Archive (2021); https://doi.org/10.15154/1504041.10.15154/1504041.
41. Derived data from Discovery of genomic loci of the human cerebral cortex using genetically informed brain atlases release #1303, NIMH Data Archive (2022); https://doi.org/10.15154/1523026.

ACKNOWLEDGMENTS

The authors thank the research participants and staff involved in data collection of the UK Biobank and Adolescent Brain Cognitive Development (ABCD) Study. The ABCD Study is a multisite, longitudinal study designed to recruit more than 10,000 children ages 9 and 10 and follow them over 10 years into early adulthood. The ABCD Study is supported by the National Institutes of Health (NIH) 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, and 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/study-sites/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily all participate in 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. **Funding:** This research was supported by the NIH under grants R01MH118281, R56AG061163, R01MH122688, R01AG050595, and R01AG022381. C.M. is supported by the Canadian Institutes of Health Research (CIHR), Fonds de Recherche du Quebec-Sante (FRQS), and the Kavli Institute for Brain and Mind (KIBM). P.M.V. and J.Y. acknowledge funding from the National Health and Medical Research Council (1113400) and the Australian Research Council (FT180100186 and FL180100072). J.Y. is supported by the Westlake Education Foundation. D.v.d.M. is supported by the Research Council of Norway (project 276082). O.A.A. is funded by the Research Council of Norway (283798, 273291, 248778, and 223273) and KG Jebsen Stiftelsen. **Author contributions:** Study design: C.M., D.v.d.M., W.D., H.W., Y.W., K.Z., P.M.V., J.Y., and C.-H.C. Data analysis: C.M., D.v.d.M., W.D., Y.W., J.Z., H.W., S.B.R., and C.-H.C. Manuscript writing: C.M., D.v.d.M., W.D., and C.-H.C. UKB data processing: D.v.d.M. and O.A.A. ABCD data collection and processing: C.M., D.J.H., C.C.F., T.L.J., and A.M.D. Single-cell data processing: W.D., Y.W., and K.Z. Data visualization: C.M., D.v.d.M., W.D., Y.W., C.L., and C.-H.C. Manuscript preparation and revisions: All authors. **Competing interests:** O.A.A. has received speaker's honoraria from Lundbeck and Sunovion and is a consultant to HealthLytix. A.M.D. is a founder of and holds equity in CorTechs Labs, Inc., and serves on its scientific advisory board. A.M.D. is also a member of the scientific advisory board of Human Longevity, Inc., and receives funding through research agreements with General Electric Healthcare and Medtronic, Inc. The terms of these arrangements have been reviewed and approved by the University of California, San Diego, in accordance with its conflict of interest policies. The other authors declare no competing interests. **Data and materials availability:** The individual-level raw data used in this study can be obtained from two accessible data resources, UK Biobank (www.ukbiobank.ac.uk/) and ABCD Study (https://abcdstudy.org). ABCD Study data were processed from the raw structural imaging data held in the National Institute of Mental Health (NIMH) Data Archive (NDA). The ABCD data repository grows and changes over time. The ABCD data used in this report came from ABCD Collection Release 2.0.1 (*40*). Data access details can be found on the NDA website (https://nda.nih.gov/abcd/request-access). GWAS summary statistics are accessible in an NDA study (*41*). We made use of publicly available software and tools. The analysis code is available in the Bio-protocol.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abe8457
Materials and Methods
Supplementary Text
Figs. S1 to S11
Tables S1 to S24
References (*42–119*)
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

17 September 2020; resubmitted 12 March 2021
Accepted 28 December 2021
10.1126/science.abe8457

Science

Discovery of genomic loci of the human cerebral cortex using genetically informed brain atlases

Carolina MakowskiDennis van der MeerWeixiu DongHao WangYan WuJingjing ZouCin LiuSara B. RosenthalDonald J. Hagler Jr.Chun Chieh FanWilliam S. KremenOle A. AndreassenTerry L. JerniganAnders M. DaleKun ZhangPeter M. VisscherJian YangChi-Hua Chen

Science, 375 (6580), • DOI: 10.1126/science.abe8457

Genes control cortical surface area

Humans exhibit heritable variation in brain structure and function. To identify how gene variants affect the cerebral cortex, Makowski *et al*. performed genome-wide association studies in almost 40,000 adults and 9000 children. They identified more than 400 loci associated with brain surface area and cortical thickness that could be observed through magnetic resonance imaging analyses. Examining biological pathways linking gene variants to phenotypes identified region-specific enrichments of neurodevelopmental functions, some of which were associated with psychiatric disorders. Partitioning genes with heritable variants relative to evolutionary conservation helped to identify a hierarchy of brain development. This analysis identified a human-specific gene-phenotype association related to speech and informs upon what genes can be studied in various model organisms. —LMZ

Science (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.