RESEARCH ARTICLE SUMMARY

NEUROSCIENCE

Common genetic variation influencing human white matter microstructure

Bingxin Zhao†, Tengfei Li†, Yue Yang, Xifeng Wang, Tianyou Luo, Yue Shan, Ziliang Zhu, Di Xiong, Mads E. Hauberg, Jaroslav Bendl, John F. Fullard, Panagiotis Roussos, Yun Li, Jason L. Stein, Hongtu Zhu*

INTRODUCTION: White matter in the human brain serves a critical role in organizing distributed neural networks. Diffusion magnetic resonance imaging (dMRI) has enabled the study of white matter in vivo, showing that interindividual variations in white matter microstructure are associated with a wide variety of clinical outcomes. Although white matter differences in general population cohorts are known to be heritable, few common genetic variants influencing white matter microstructure have been identified.

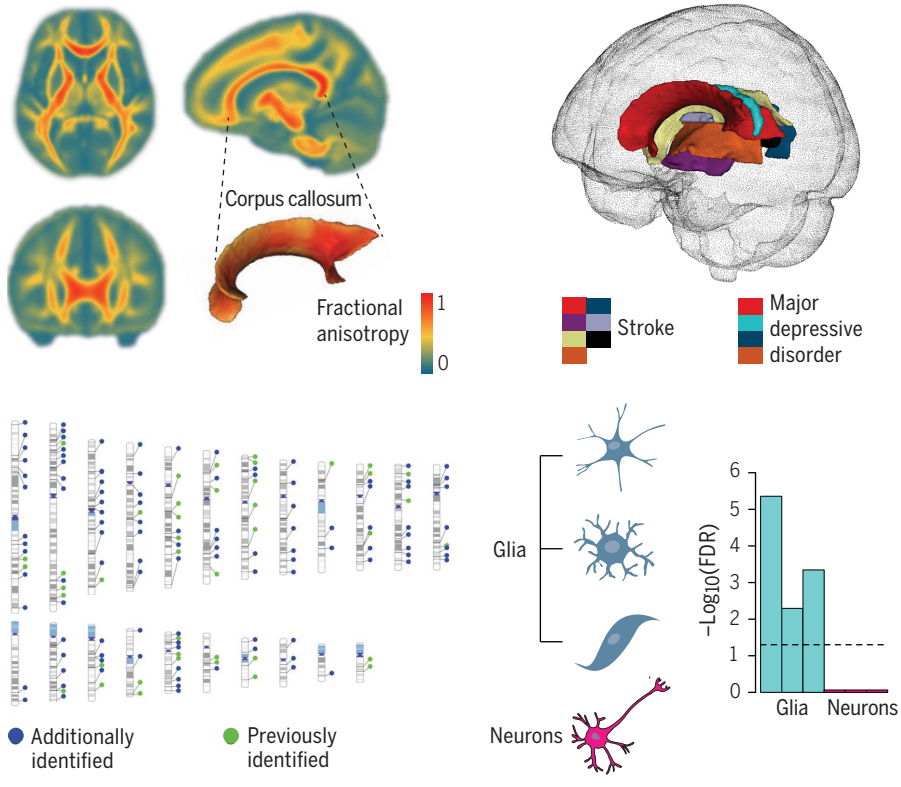
RATIONALE: To identify genetic variants influencing white matter microstructure, we conducted a genome-wide association study (GWAS) of dMRI data from 43,802 individuals across five data resources. We analyzed five major diffusion tensor imaging (DTI) model–derived parameters along 21 cerebral white matter tracts.

RESULTS: In the discovery GWAS with 34,024 individuals of British ancestry, we replicated 42 of the 44 genomic regions discovered in the largest previous GWAS and identified 109 additional regions associated with white matter microstructure ($P < 2.3 \times 10^{-10}$, adjusted for the number of phenotypes studied). These results indicate strong polygenic influences on white matter microstructure. Of the 151 regions, 52 passed the Bonferroni significance level ($P < 5 \times 10^{-5}$) in our analysis of nine independent validation datasets, including four with subjects of non-European ancestry.

On average, common genetic variants explained 41% (standard error = 2%) of the variation in white matter microstructure. The 151 identified genomic regions can explain 32.3% of heritability for white matter microstructure, whereas the 44 previously identified genomic regions can only explain 11.7% of heritability. As a biological validation of our GWAS findings, we observed heritability enrichment within regulatory elements active in oligodendrocytes and other glia, whereas no enrichment was observed in neurons. These results are expected and suggest that genetic variation leads to changes in white matter microstructure by affecting gene regulation in glia.

We observed genetic correlations and colocalizations of white matter microstructure with a wide range of brain-related complex traits and diseases, such as cognitive functions, cardiovascular risk factors, as well as various neurological and psychiatric diseases. For example, of the 25 reported genetic risk regions of glioma, 11 were also associated with white matter microstructure, which illustrates the close genetic relationship between glioma and white matter integrity. Additionally, we found that 14 white matter microstructure–associated genes ($P < 1.2 \times 10^{-8}$) were targets for 79 commonly used nervous system drugs, such as antipsychotics, antidepressants, anticonvulsants, and drugs for Parkinson's disease and dementia.

CONCLUSION: This large-scale study of dMRI scans from 43,802 subjects improved our understanding of the highly polygenic genetic architecture of human brain white matter tracts. We identified 151 genomic regions associated with white matter microstructure. The GWAS findings were supported by enrichments within cell types that make up white matter microstructure. Moreover, we uncovered genetic relationships between white matter and various clinical endpoints, such as stroke, major depressive disorder, schizophrenia, and attention deficit hyperactivity disorder. The targets of many drugs commonly used for disabling cognitive disorders have genetic associations with white matter, which suggests that the neuropharmacology of many disorders can potentially be improved by studying how these medications work in the brain white matter. ■



Identifying genetic variants influencing human brain white matter microstructure. (Top left) Quantifying the microstructure in white matter tracts using DTI models. (Bottom left) Genomic locations of common genetic variants associated with white matter microstructure. (Top right) Selected genetic correlations between white matter microstructure and brain disorders (stroke and major depressive disorder). (Bottom right) Partitioned heritability enrichment analysis in brain cell types. FDR, false discovery rate.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: htzhu@email.unc.edu
†These authors contributed equally to this work.
Cite this article as B. Zhao *et al.*, *Science* **372**, eabf3736 (2021). DOI: 10.1126/science.abf3736

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abf3736

Downloaded from http://science.sciencemag.org/ on June 17, 2021

RESEARCH ARTICLE

NEUROSCIENCE

Common genetic variation influencing human white matter microstructure

Bingxin Zhao[1]†, Tengfei Li[2,3]†, Yue Yang[4], Xifeng Wang[4], Tianyou Luo[4], Yue Shan[4], Ziliang Zhu[4], Di Xiong[4], Mads E. Hauberg[5,6,7,8], Jaroslav Bendl[5,9,6], John F. Fullard[5,9,6], Panagiotis Roussos[5,9,6,10], Yun Li[4,11,12], Jason L. Stein[11,13], Hongtu Zhu[3,4]*

Brain regions communicate with each other through tracts of myelinated axons, commonly referred to as white matter. We identified common genetic variants influencing white matter microstructure using diffusion magnetic resonance imaging of 43,802 individuals. Genome-wide association analysis identified 109 associated loci, 30 of which were detected by tract-specific functional principal components analysis. A number of loci colocalized with brain diseases, such as glioma and stroke. Genetic correlations were observed between white matter microstructure and 57 complex traits and diseases. Common variants associated with white matter microstructure altered the function of regulatory elements in glial cells, particularly oligodendrocytes. This large-scale tract-specific study advances the understanding of the genetic architecture of white matter and its genetic links to a wide spectrum of clinical outcomes.

Brain functions depend on effective communications across brain regions (*1*). White matter makes up roughly half of the human brain and contains most of the brain's long-range communication pathways (*2*). White matter tracts build a complex network of structural connections (*3*, *4*). Cerebral white matter tracts form connections among cerebral cortical areas and between cortical and subcortical regions (*5*, *6*). Aspects of this connectivity can be estimated through measurements of cellular microstructure (*7*). Evidence from neuroscience has further suggested that cerebral white matter microstructure may underpin brain function and dysfunction and that connectivity differences are relevant to a wide variety of neurological and psychiatric disorders, such as attention deficit hyperactivity disorder (ADHD) (*8*), major depressive disorder (MDD) (*9*), schizophrenia (*10*), and Alzheimer's disease (*11*). White matter microstructural differences and abnormalities can be captured in vivo by diffusion magnetic resonance imaging (dMRI). Using dMRI data, microstructural connectivity can be quantified in diffusion tensor imaging (DTI) models (*12*) and measured by several DTI-derived metrics, including fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), radial diffusivity (RD), and mode of anisotropy (MO) (Fig. 1A). Among these, FA serves as the primary metric of interest in many studies (*13*). It is a robust measure of overall integrity and directionality and is highly sensitive to general connectivity changes. On the other hand, MD, AD, and RD directly quantify the absolute magnitude of directionalities. They are more sensitive to specific types of microstructural changes and have also been used in a variety of clinical studies (*14*). Additionally, MO can characterize the anisotropy type and describe whether the shape of the diffusion tensor is more linear or planar (*15*).

White matter differences in general population cohorts are under strong genetic control. Both family- and population-based studies have reported that DTI measurements of white matter microstructure have generally high heritability, with estimates varying across different age groups (*16*) and tracts (*17*). For example, heritability estimates of tract-averaged FA ranged from 53 to 90% in a twin study of the Human Connectome Project (HCP) (*18*). Recent genome-wide association studies (GWAS) of the UK Biobank (UKB) reported an average single-nucleotide polymorphism (SNP)–based heritability of 48.7% across different tracts (*19*). Several GWAS (*17*, *19*–*23*) have been performed to identify loci associated with interindividual variation in white matter microstructure, but these studies have shared at least two major limitations: (i) sample size and (ii) spatial specificity. First, the largest currently published GWAS of dMRI phenotypes had a sample size of 17,706 individuals (*19*). Similar to other brain-related traits, white matter has a complex and extremely polygenic genetic architecture (*19*). Large sample size is therefore essential to boost GWAS power to identify many common risk variants with small effect sizes and to uncover the shared genetic influences with other complex traits. Second, previous GWAS mainly focused on global dMRI measures of the whole brain (*20*, *21*) or tract-averaged (mean) values (*17*, *19*). Global and tract-averaged measures can capture the largest variations in white matter while reducing the burden to test multiple neuroimaging traits, which is particularly suitable for GWAS with limited sample sizes; however, these measures may lose information because microstructural differences and changes may not have a uniformly consistent pattern across the whole tract. Heterogeneous variation patterns typically exist within voxel-wise DTI maps of the three-dimensional (3D) tract curve, which may be more relevant to specific underlying biological processes. For example, a previous study found that the association between bipolar disorder and FA is specific to one given segment of the anterior limb of the internal capsule tract connecting the prefrontal cortex with the thalamus and brainstem (*24*). Because of these limitations, a large number of genetic factors influencing white matter may still be undiscovered. Consequently, with few exceptions [e.g., stroke (*20*) and cognitive traits (*19*)], the shared genetic influences between white matter and other complex traits are unknown. Uncovering these potential genetic links may identify important brain regions that are involved in clinical outcomes, especially for brain disorders.

To overcome these limitations, we collected individual-level dMRI data from five data resources: UKB (*25*); Adolescent Brain Cognitive Development (ABCD) (*26*); HCP (*27*); Pediatric Imaging, Neurocognition, and Genetics (PING) (*28*); and Philadelphia Neurodevelopmental Cohort (PNC) (*29*). We harmonized image processing by using the ENIGMA-DTI pipeline (*30*, *31*) and obtained voxel-wise DTI maps for 43,802 subjects (after quality controls), including 36,624 from UKB. We mainly focused on 21 predefined cerebral white matter tracts (Fig. 1B) and generated two groups of phenotypes (*19*). The first group contains 110 tract-averaged parameters for FA, AD, MD, MO, and RD in 21 tracts and across all white matter tracts (5 × 22). For the second group, we applied functional principal components

[1]Department of Statistics, Purdue University, West Lafayette, IN 47907, USA. [2]Department of Radiology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [3]Biomedical Research Imaging Center, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [4]Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [5]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [6]Department of Genetics and Genomic Science and Institute for Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [7]iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research, 8210 Aarhus, Denmark. [8]Centre for Integrative Sequencing (iSEQ), Aarhus University, 8000 Aarhus, Denmark. [9]Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [10]Mental Illness Research, Education, and Clinical Center (VISN 2 South), James J. Peters VA Medical Center, Bronx, NY 10468, USA. [11]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [12]Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [13]UNC Neuroscience Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
*Corresponding author. Email: htzhu@email.unc.edu
†These authors contributed equally to this work.

Downloaded from http://science.sciencemag.org/ on June 17, 2021

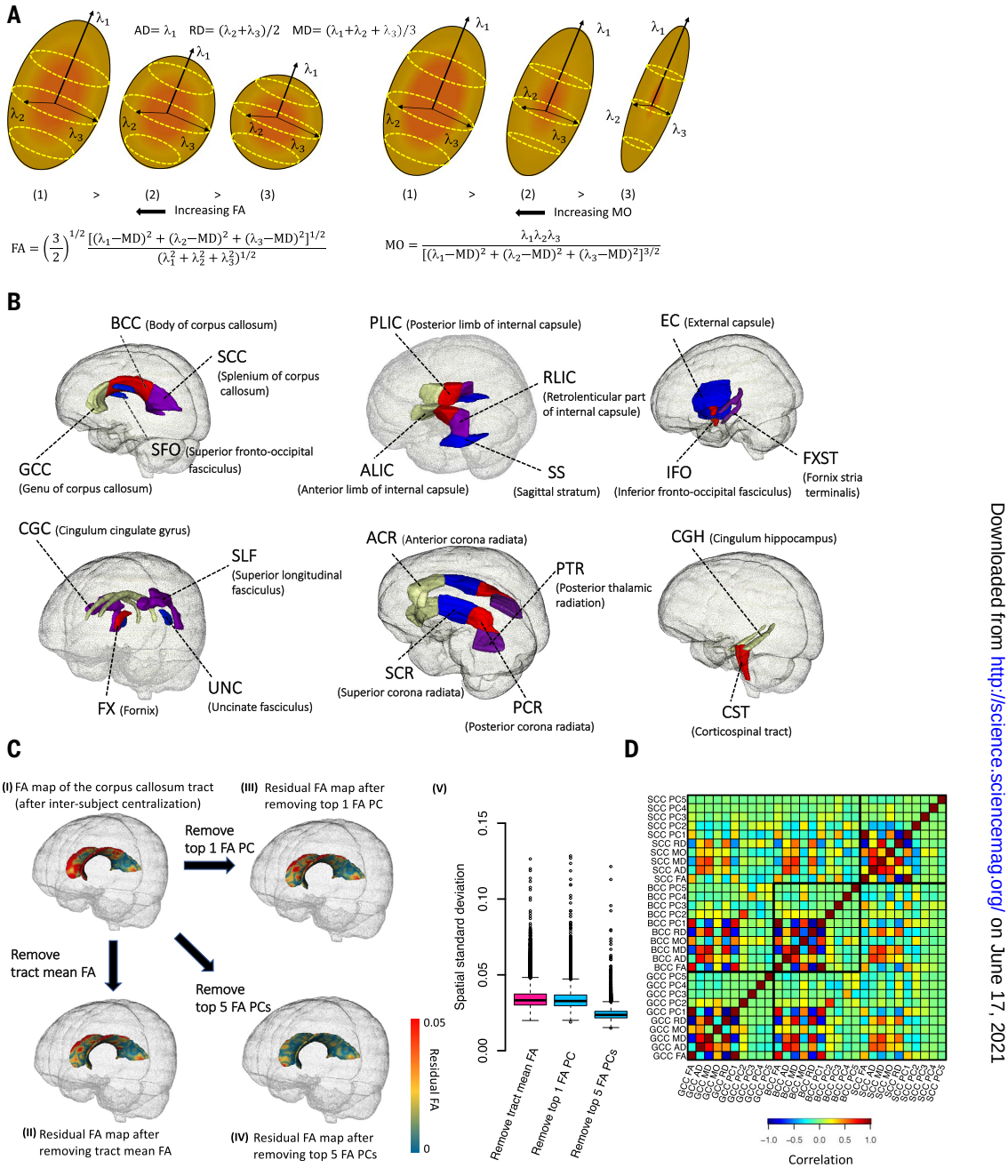


Fig. 1. Illustration of white matter traits and DTI parameters. (**A**) Schematic representation of the five DTI-derived metrics. Axial diffusivity (AD) is the diffusion along the long axis (λ_1), radial diffusivity (RD) is the diffusion of two small axes (average of λ_2 and λ_3), and mean diffusivity (MD) is the average diffusion regardless of direction (mean of λ_1, λ_2, and λ_3). Fractional anisotropy (FA) and mode of anisotropy (MO) are two more complicated measures of general directionality. (**B**) Annotation of the 21 white matter tracts in the human brain. (**C**) Comparison between tract mean FA and tract-specific FA PCs on the corpus callosum tract [(CC), including the genu of corpus callosum (GCC), body of corpus callosum (BCC), and splenium of corpus callosum (SCC)]. (I) illustrates an example skeletonized FA map within the corpus callosum tract after intersubject centralization, and (II) displays the residual FA map after removing the within-subject tract mean FA. In (III) and (IV), instead of removing the within-subject mean as in (II), we removed the top one and five FA PCs, respectively. (V) illustrates the standard deviation across the voxels in a residual FA map for each subject in the UKB (n = 36,624). The standard deviations are similar between the residual FA maps after removing tract mean FA [in (II)] and the first FA PC [in (III)], which suggests that this PC mainly accounts for the mean FA. Comparing (II) with (IV), the other four FA PCs can capture more spatial variations that are ignored by the tract mean FA and thus reduce the standard deviations of residuals in (V). (**D**) Correlation between the DTI parameters among GCC, BCC, and SCC tracts. Ten parameters are generated in each tract, including the mean FA, mean MD, mean AD, mean RD, and mean MO, as well as the top five FA PCs. The five FA PCs are orthogonal to each other, and the first FA PC can be highly correlated with mean FA.

analysis (FPCA) (*32*) to generate 105 tract-specific principal components (PCs) for FA by taking the top five PCs of the voxel-wise map within each tract. FPCA is a well-established approach in neuroimaging and functional data analysis (*33*). In DTI models, FPCA can characterize the strongest variation components of FA within each tract, which are expected to provide more microstructural details about axonal organization and myelination omitted by tract-averaged values while relieving multiple testing burdens. These PCs may represent FA changes that are more relevant to specific clinical outcomes. We used the UKB imaging data released up through 2018 (referred to as UKB phases 1 and 2 hereafter; n = 17,706 individuals) to generate the loadings of FPCA and applied them to compute the FA PCs in other independent cohorts. Figure 1C illustrates the difference between tract-averaged FA and tract-specific FA PCs for the corpus callosum, which is the largest white matter tract in the human brain. The first FA PC performed similarly to (but slightly better than) the tract-averaged FA, and the other top-ranked FA PCs (which are orthogonal to the first FA PC; Fig. 1D) further captured FA variations in subfields of the corpus callosum. We then performed a GWAS for these 215 phenotypes to discover the genetic architecture of white matter and explore the genetic links to a plethora of clinical endpoints in different trait domains. Our GWAS results have been made

Downloaded from http://science.sciencemag.org/ on June 17, 2021

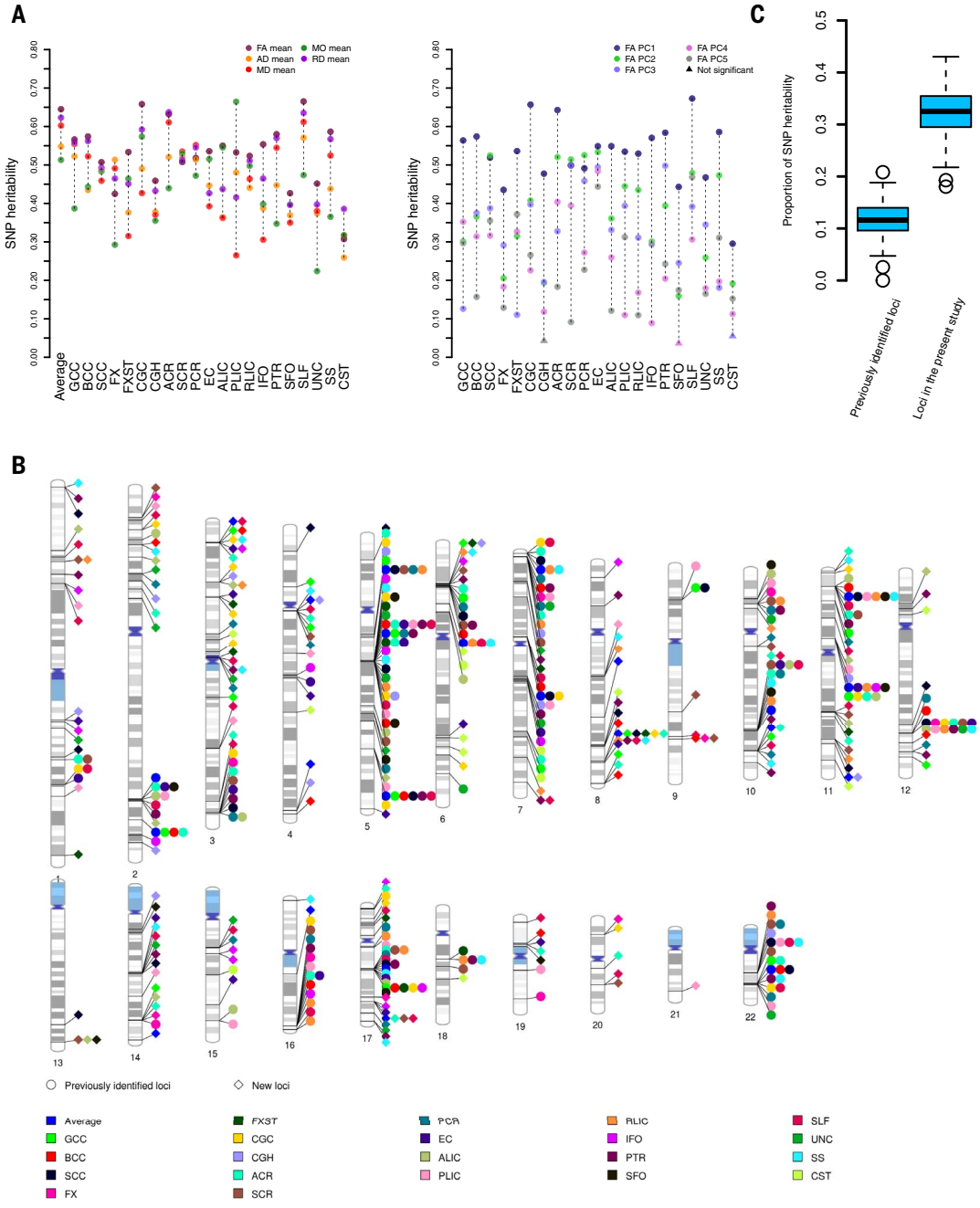


Fig. 2. SNP heritability and the associated genomic regions of DTI parameters. (**A**) SNP heritability of 215 DTI parameters, including 110 mean parameters (left panel, mean values of AD, RD, MD, FA, and MO in 21 tracts and the whole brain; 5 × 22) and 105 FA PCs (right panel, top five FA PCs of 21 tracts; 5 × 21). Average indicates the global average across 21 tracts; full names of the 21 white matter tracts can be found in Fig. 1B. (**B**) Ideogram of genomic regions influencing DTI parameters ($P < 2.3 \times 10^{-10}$), including 42 previously identified regions and 109 additional regions identified in the present study. The colors represent the 21 white matter tracts (and the global average). Each signal point indicates that at least one of the 10 DTI parameters (five mean parameters and five FA PCs) of this tract is associated with the genomic region. (**C**) Proportion of SNP heritability of the 215 DTI parameters that can be explained by the 44 genomic regions identified in the previous study (11.7%) and 151 regions identified in the present study (32.3%).

publicly available and can be browsed through our Brain Imaging Genetics Knowledge Portal (BIG-KP) at https://bigkp.org/.

GWAS discovery and validation for 215 DTI parameters

Our heritability analysis utilized data from UKB subjects of British ancestry (n = 34,024). All of the 110 DTI mean parameters had significant SNP heritability (*34*) (h^2) after Bonferroni adjustment ($P < 4.1 \times 10^{-25}$; Fig. 2A and table S1). The h^2 estimates varied from 22.4 to 66.5% (mean h^2 = 47.6%), which were comparable with previous results (*17*, *19*). For the 105 tract-specific FA PC parameters, we found that 102 had significant h^2 (h^2 range = 8.9 to 67.3%; mean h^2 = 34.9%; $P < 6.5 \times 10^{-6}$). For example, all of the top five PCs in the external capsule, superior longitudinal fasciculus, and splenium of the corpus callosum had moderate to high heritability (h^2 range = 30.6 to 67.3%; $P < 2.3 \times 10^{-47}$). The third PC of the corticospinal tract (5.5%), fifth PC of the cingulum hippocampus (4.3%), and fourth PC of the superior fronto-occipital fasciculus (3.6%) had nominally significant h^2 estimates ($P < 0.03$), which became insignificant after Bonferroni adjustment. To examine the robustness of heritability estimates of the 105 FA PCs, we reran the heritability analysis separately using the UKB phases 1 and 2 data (n = 17,706) and the newly released independent data from 2020 (named the UKB phase 3; n = 15,918, removing the relatives of subjects in the early released dataset). The correlation between the two sets of heritability estimates was 0.92 (mean h^2 = 34.6 versus 33.6%; two-sample t test, P = 0.64; table S2), which suggests that the FA PCs are heritable and have consistent heritability in independent cohorts. Overall, heritability analysis showed that the major microstructural variations of voxel-wise FA maps captured by unconventional tract-specific FA PCs are under genetic control. Those heritable FA variation patterns may have higher power than mean FA

Downloaded from http://science.sciencemag.org/ on June 17, 2021

to identify the shared genetic influences with other complex traits and clinical outcomes.

We used a multiple-stage design in our GWAS of the 215 DTI parameters (fig. S1). First, we checked the robustness of GWAS results within the UKB British cohort. Specifically, GWAS were performed separately in two independent datasets: the UKB phases 1 and 2 data (n = 17,706) and the independent UKB phase 3 data (n = 15,918). We calculated the replication slope between the two set of GWAS results, which was the correlation of the standardized effect size of variants estimated from two independent GWAS (*35*). This analysis was restricted to top ($P < 1 \times 10^{-6}$) independent lead variants after linkage disequilibrium (LD)–based clumping [window size, 250 kilobases; LD coefficient of determination (R^2) = 0.01]. The replication slope was 0.8 (standard error = 0.01; $P < 2 \times 10^{-16}$), indicating strong similarity between these top variant effect size estimates. In addition, for each of the 215 DTI parameters, we used LD score regression (LDSC) (*36*) to calculate genetic correlation (gc) between DTI measurements from the two GWAS. The mean gc estimate was 1.01 (standard error = 0.17; table S3) across these parameters, 209 of which were significant after adjusting for multiple comparisons by using the Benjamini-Hochberg procedure to control false discovery rate (FDR) at the 5% level. Genetic correlation estimates near 1 indicate a consistent genetic basis for these phenotypes measured in different cohorts. These findings indicate that GWAS results of DTI parameters are highly consistent within the UKB British cohort.

Second, we performed GWAS for these 215 DTI parameters using the combined UKB data of British ancestry (n = 34,024). All Manhattan and quantile-quantile plots can be browsed through our BIG-KP server (https://bigkp.org/). The LDSC intercepts (*37*) were close to one, which suggests that there was no genomic inflation of test statistics as a result of confounding factors (mean intercept = 1.008; standard error = 0.008; table S4). At a stringent significance level of 2.3×10^{-10} (5×10^{-8}/215, additionally adjusted for the 215 phenotypes studied), FUMA (*38*) identified 539 independent significant lead variants (LD $R^2 < 0.1$) involved in 1312 significant associations with 173 DTI parameters (106 mean and 67 PC parameters; table S5). These variant-level associations were further summarized into 996 significant locus-trait associations (table S6). Genetic variants had broad effects across all of the 21 white matter tracts. We replicated 42 of the 44 genomic regions discovered in the largest previous GWAS (*19*) and identified 109 additional regions associated with DTI parameters (Fig. 2B). On average, the 151 (42 + 109) significant genomic regions can explain 32.3% of SNP heritability for the 215 DTI parameters, whereas the 44 previously identified genomic regions can only explain 11.7% of SNP heritability (Fig. 2C and table S7). Of the 109 genomic regions we identified, 30 were only detected by FA PC parameters. For example, the third FA PC of the posterior thalamic radiation detected three additional regions (fig. S2). Our results illuminate the broad genetic control of white matter microstructural differences and the contribution of tract-specific FA PCs in identifying genetic variants associated with white matter tracts. The genetic effects are spread across a large number of genomic regions, consistent with the observed polygenic genetic architecture of many brain-related traits (*39*).

Next, to replicate our results from the UKB British GWAS, we performed additional analysis in nine independent validation datasets. GWAS was separately performed on five datasets consisting of individuals of European ancestry, including the UKB White but non-British (UKBW, n = 1938), ABCD European (ABCDE, n = 3821), HCP (n = 334), PING (n = 461), and PNC (n = 537), and on four non-European validation datasets: the UKB Asian (UKBA, n = 435), UKB Black (UKBBL, n = 227), ABCD Hispanic (ABCDH, n = 768), and ABCD African American (ABCDA, n = 1257). For each DTI parameter, the genetic correlation between discovery GWAS and the meta-analyzed European validation GWAS (total n = 7091) was estimated by LDSC (*36*). The mean genetic correlation estimate was 0.93 (standard error = 0.36) across the 215 DTI parameters, 136 of which were significant at the FDR 5% level (table S8). We then checked whether the locus-trait associations detected in UKB British GWAS can be replicated in these validation datasets. For the 996 significant associations, 203 (20.4%; in 50 genomic regions) passed the 5×10^{-5} (0.05/996) Bonferroni significance level in the meta-analyzed European validation GWAS, and 532 (53.4%; in 92 regions) were significant at the 1×10^{-3} level. We also meta-analyzed the four non-European GWAS (total n = 2687) and found that 64 associations (6.4%; in 19 regions) passed the Bonferroni significance level, and 367 (36.8%; in 75 regions) were significant at the 1×10^{-3} level. Moreover, we performed a third meta-analysis to combine all of the nine validation datasets, after which the number of replicated associations moved up to 258 (25.9%; in 52 regions) and 678 (68%; in 115 regions) at Bonferroni and the 1×10^{-3} significance levels, respectively (table S9). Overall, our results suggest that the associated genetic loci discovered in the UKB British GWAS can be strongly validated in independent studies, despite the fact that these studies may use different cohorts and/or MRI scanners. Therefore, we performed a final meta-analysis to combine all of these datasets (n = 43,802). At the 2.3×10^{-10} significance level, the meta-analyzed GWAS detected 193 significant genomic regions, including 144 (95.4%) of the 151 regions identified in the UKB British GWAS (table S10).

Additionally, we tested for replication by using polygenic risk scores (PRS) (*40*) derived from the UKB British GWAS. At the FDR 5% level (215 × 5 tests), the mean number of significant PRS in the five European validation GWAS datasets was 201 (range = 193 to 210; table S11). Almost all (213/215) DTI parameters had significant PRS in at least one dataset, and 173 had significant PRS in all of them, which demonstrates the high generalizability of our discovery GWAS results. Across the five validation datasets, the mean additional variance that can be explained by PRS (incremental R^2, adjusting for the effects of sex, age, and top 10 genetic PCs) was 1.8% (range = 0.6 to 4%) for the 173 consistently significant DTI parameters. The largest R^2 was on the fourth PC of the external capsule (range = 1.8 to 5.3%; P range = 1.2×10^{-19} to 5.6×10^{-12}). Finally, we constructed PRS on four non-European validation datasets. At the FDR 5% level, the number of significant PRS decreased to 172, 46, 144, and 147 in UKBA, UKBBL, ABCDH, and ABCDA, respectively (table S12). These findings show that the UKB British GWAS results have high generalizability in European cohorts, but the generalizability is reduced in cross-population applications, which highlights the importance of recruiting sufficient samples from global diverse populations in future genetic discovery of white matter.

Concordance with previous GWAS

We carried out association lookups for independent significant variants (*38*) ($P < 2.3 \times 10^{-10}$; and variants within LD, $R^2 \geq 0.6$) detected in our UKB British discovery GWAS. In the NHGRI-EBI GWAS catalog (*41*), white matter microstructure had shared genetic influences with a wide range of other complex traits in different trait domains, including brain structures, glioma [glioblastoma (GBM) and non-GBM tumors], stroke, vascular risk factors (such as diabetes, high blood pressure, obesity, and smoking), cognitive traits, neurological disorders (such as Alzheimer's disease and Parkinson's disease), psychiatric disorders [such as schizophrenia and autism spectrum disorder (ASD)], psychological traits (such as neuroticism and well-being spectrum), sleep, anthropometric measurements, bone mineral density, and educational attainment. These results are summarized in table S13.

Our results tagged 17 genomic regions that had been associated with brain structures, such as 3q28 and 17q21.31 with intracranial volume (*42*); 22q13.1, 10p12.31, 16q24.2, 3q28, and 12q23.3 with lateral ventricular volume (*43*); 5q14.3, 6q25.1, 13q34, 2p16.1, 17q25.1, 10q24.33, and 2q33.2 with white matter hyperintensity (*20*, *44*); 10q26.13, 14q23.1, and 5q12.3

Downloaded from http://science.sciencemag.org/ on June 17, 2021

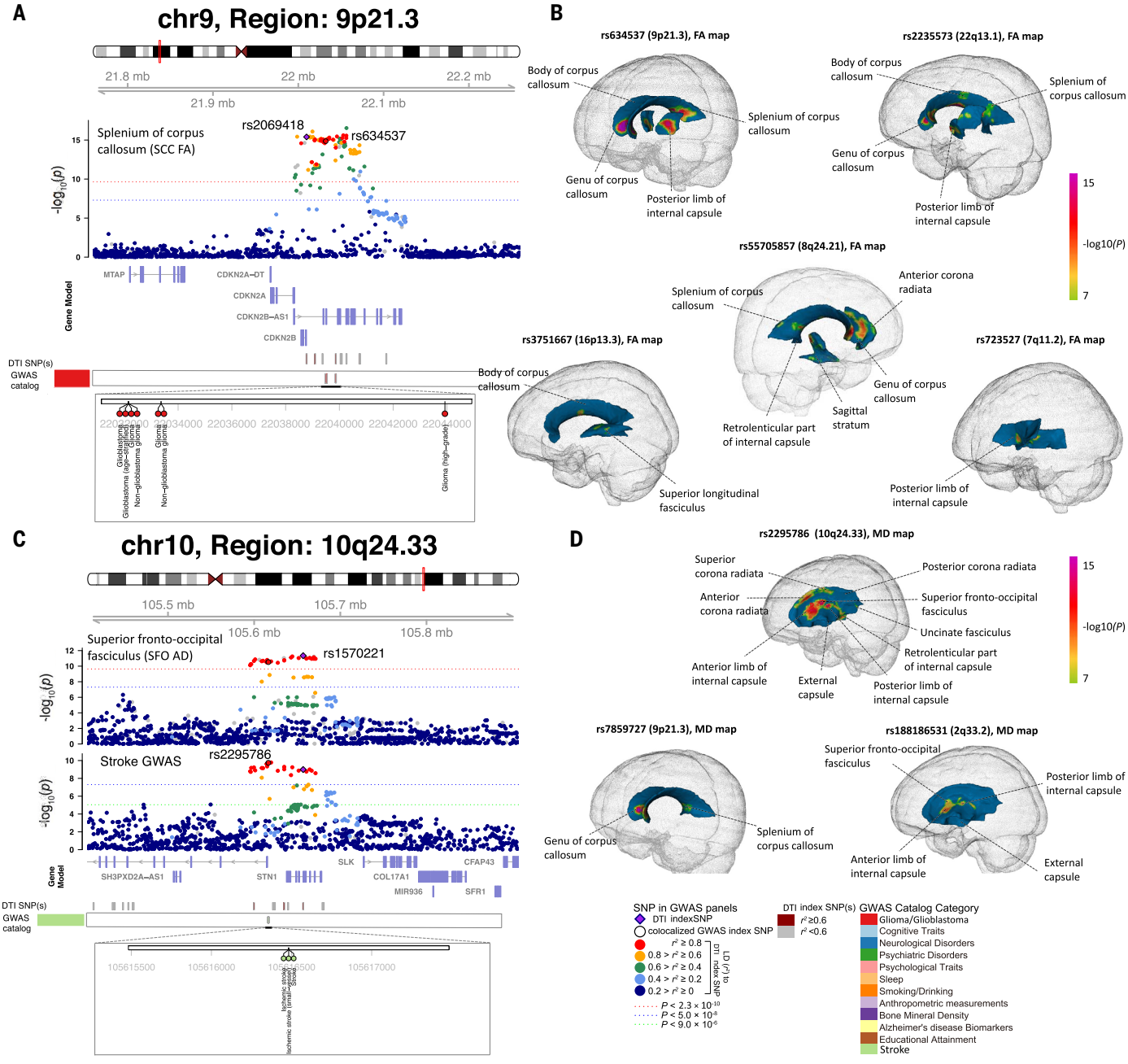


Fig. 3. Selected genetic loci that were associated with both DTI parameters and brain diseases. (**A**) In 9p21.3, we observed colocalization (LD $R^2 \geq 0.6$) between the mean FA of the splenium of corpus callosum (SCC FA; index variant rs2069418) and glioma (both GBM and non-GBM tumors; index variant rs634537). chr9, chromosome 9; mb, megabase. (**B**) We illustrate the voxel-wise genetic effects of five colocalized glioma GWAS index variants (rs634537, rs2235573, rs55705857, rs3751667, and rs723527) on FA. The genetic effects were obtained by performing voxel-wise target-variant analysis for the five colocalized glioma significant variants. We display the voxels passing the Bonferroni significant level ($P < 8.5 \times 10^{-8}$) in the voxel-wise target-variant analysis. White matter tracts that had significant voxels are labeled in each map. (**C**) In 10q24.33, we observe colocalization between the mean AD of the superior fronto-occipital fasciculus (SFO AD; index variant rs1570221) and stroke (index variant rs2295786). (**D**) We illustrate the voxel-wise genetic effects of three colocalized stroke GWAS index variants (rs2295786, rs7859727, and rs18818651) on MD.

with hippocampal subfield volumes (*45*); and 15q14 with brain imaging measurements (*17*). Moreover, we observed colocalizations with glioma in 9p21.3, 22q13.1, 16p13.3, 7p11.2, and 8q24.21 (*46*) (Fig. 3A and figs. S3 to S6). These colocalizations were detected by multiple mean parameters and FA PCs. To illustrate the genetic effect pattern in spatial brain maps, we took the lead index variants from the five glioma-associated regions and performed target-variant analysis on voxel-wise FA and MD data. After Bonferroni adjustment ($P < 8.5 \times 10^{-8}$), glioma risk variants were associated with multiple white matter tracts, including the corpus callosum, internal capsule, superior longitudinal fasciculus, anterior corona radiata, and sagittal stratum (Fig. 3B and fig. S7A). Reduced microstructure integrity and altered brain structural connectivity have been observed in glioma (*47*). Early invasive glioma growth along white matter tracts is one of the

Downloaded from http://science.sciencemag.org/ on June 17, 2021

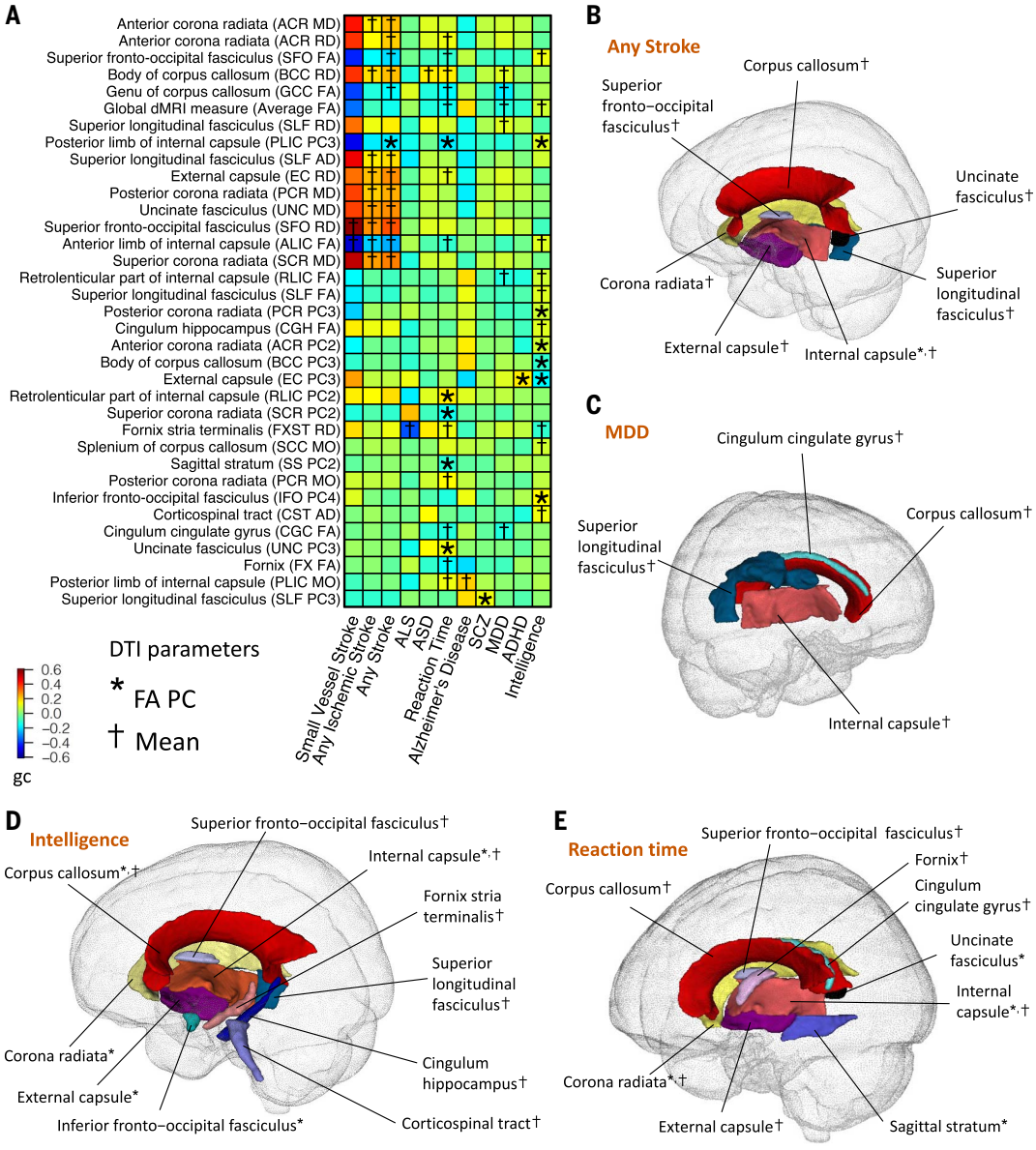


Fig. 4. Selected pairwise genetic correlations between DTI parameters of white matter tracts and brain disorders and cognitive functions. (**A**) The asterisks (for FA PCs) and daggers (for mean parameters) highlight significant genetic correlations after controlling the FDR at 5% level. The *y* axis lists the DTI parameters of white matter tracts and the *x* axis provides the names of brain-related traits and disorders. The colors represent genetic correlations (gc). PC, principal component of FA; SCZ, schizophrenia. (**B** to **E**) Location of the white matter tracts whose DTI parameters were genetically correlated with stroke (any subtype) (B), MDD (C), intelligence (D), and reaction time (E). The colors describe different white matter tracts.

main challenges in treating glioma (*48*). The corpus callosum serves as a hub between hemispheres, and glial cells may migrate along the corpus callosum during glioma cell infiltration (*49*). Additionally, three regions were colocalized with stroke, including 10q24.33, 9p21.3, and 2q33.2 (*50*) (Fig. 3, C and D; fig. S7B; and fig. S8). Stroke risk variants from these regions had genetic effects in the corpus callosum, superior fronto-occipital fasciculus, corona radiata, internal capsule, and external capsule. Ischemic white matter disorganization and the associated neuropsychiatric disorders (such as depression) have been frequently reported after stroke (*51*). White matter integrity is closely related to the pathologic stages of stroke and can predict the outcome of stroke recovery. For example, the MD of ischemic regions stays reduced during the acute phase of stroke, then increases over time to become pseudonormalized in the later-subacute phase, and continues rising in the chronic phase (*52*). Moreover, these three colocalized regions were also associated with cardiovascular disease, type 2 diabetes (T2D), obesity, and high blood pressure, which suggests that the genetic relationship between white matter microstructure and stroke may be partially mediated through vascular risk factors.

The 17q21.31 inversion region and the 6p22.1 and 6p22.2 in the extended major histocompatibility complex (MHC) region have complicated LD patterns. In these regions, white matter–associated variants colocalized with risk variants of many brain disorders and cognition, such as schizophrenia (*53*), Alzheimer's disease (*54*), corticobasal degeneration (*55*), Parkinson's disease (*56*), progressive supranuclear palsy (*55*), ASD (*57*), alcohol use disorder (*58*), depression (*59*), educational attainment (*60*), and cognitive ability (*61*) (figs. S9 to S11). Colocalizations were also found in 3q28 with Alzheimer's disease biomarkers (fig. S12), in 3p21.31 with intelligence (fig. S13), in 3p22.1 with amyotrophic lateral sclerosis (ALS), and in 1q23.2 with multiple sclerosis. Many colocalized loci of cognitive traits and educational attainment can only be detected by FA PCs. We also observed colocalizations with psychological traits [such as neuroticism (*62*), well-being spectrum (*63*), and general risk tolerance (*64*)], anthropometric traits [such as body mass index (BMI) (*35*)], smoking and drinking (*65*), and sleep [including sleep duration (*66*) and chronotype (*67*)]. Overall, genetic links were found between white matter microstructure and a wide spectrum of clinical outcomes and complex traits. Integrating the GWAS of white matter with these clinical outcomes may help us to understand the underlying mechanisms that lead

Downloaded from http://science.sciencemag.org/ on June 17, 2021

to the changes in brain structural connectivity and risk of brain disorders.

An atlas of genetic correlations with other complex traits

Because of the shared loci associated with both white matter microstructure and other complex traits, we systematically examined their pairwise genetic correlations by using publicly available summary-level data of 70 other complex traits through LDSC (table S14). We mainly focused on brain-related complex traits with large GWAS sample sizes, such as neurological and psychiatric disorders, cognitive traits, and cardiovascular risk factors. There were 727 significant pairs between 57 complex traits and 171 DTI parameters at the FDR 5% level (70 × 215 tests; P range = 3×10^{-13} to 2.4×10^{-3}; table S15), 41% (298/727) of which were detected by PC parameters. We found that DTI parameters were widely correlated with brain disorders (such as stroke, MDD, ADHD, schizophrenia, ALS, Alzheimer's disease, and ASD), cognition, educational attainment, chronotype, insomnia, neuroticism, risk tolerance, automobile speeding, high blood pressure, T2D, coronary artery disease, BMI, white matter hyperintensities, drinking, smoking, manual occupation, and lung function (figs. S14 to S16).

We replicated previously reported genetic correlations with cognitive or educational traits (*19*), drinking behavior (*19*), stroke (*17*, *20*), and MDD (*19*, *20*), and more tract-specific details were revealed. For example, stroke (any subtype) and ischemic stroke subtypes (*50*) (large artery stroke, cardioembolic stroke, and small vessel stroke) showed broad genetic correlations with the corpus callosum, corona radiata, internal capsule, external capsule, superior longitudinal fasciculus, superior fronto-occipital fasciculus, and uncinate fasciculus (|gc| range = 0.3 to 0.42; P range = 3.2×10^{-9} to 2.4×10^{-3}; Fig. 4, A and B), matching previous findings in the literature (*68*, *69*). We further observed that the small vessel stroke subtype had specific, but stronger genetic correlations with the anterior limb of the internal capsule and the superior fronto-occipital fasciculus (|gc| range = 0.56 to 0.69; P range = 3.7×10^{-4} to 1.2×10^{-3}). By contrast, there were no significant genetic correlations detected for large artery and cardioembolic strokes, which demonstrates the potentially much stronger genetic links between white matter tracts and the small vessel stroke subtype. This finding is consistent with previous observations that white matter changes are associated with small vessel strokes but not with non–small vessel events (*70*).

In addition, we observed genetic correlations between MDD and the corpus callosum, superior longitudinal fasciculus, internal capsule, and cingulum cingulate gyrus (|gc| range = 0.16 to 0.19; P range = 1.8×10^{-7} to 2.1×10^{-3}; Fig. 4C). Anatomical deficits in the corpus callosum are related to interhemispheric disconnections involved in the pathology of MDD (*71*). Lower FA in the superior longitudinal fasciculus, internal capsule, and cingulum has also been reported in patients with MDD (*14*, *72*). For intelligence and reaction times, we found genetic correlations with the FA of multiple white matter tracts, such as the internal capsule, corpus callosum, superior fronto-occipital fasciculus, corona radiata, external capsule, fornix, and superior longitudinal fasciculus (|gc| range = 0.15 to 0.18; P range = 3×10^{-9} to 2.2×10^{-3}; Fig. 4, D and E). Higher FA can reduce interhemispheric transfer time and increase information processing speed, resulting in more efficient cognitive functioning and faster reaction (*73*).

Moreover, many genetic correlations were uncovered for brain disorders, including schizophrenia, ADHD, ALS, Alzheimer's disease, and ASD. For example, significant genetic correlations were found between the posterior limb of the internal capsule and Alzheimer's disease, the external capsule and ADHD, the fornix stria terminalis and ALS, the superior longitudinal fasciculus and schizophrenia, and the body of corpus callosum and ASD (|gc| > 0.16; $P < 2 \times 10^{-3}$). Alterations of white matter integrity have been consistently observed in the corpus callosum among ASD patients (*74*). Fornix degeneration is involved in the pathology pathway of memory impairment in ALS, and the external capsule is related to ADHD symptoms (*75*). Additionally, white matter changes in the posterior limb of the internal capsule have been observed in people with Alzheimer's disease (*76*). For schizophrenia, case-control differences of FA were frequently reported in the superior longitudinal fasciculus (*77*).

We also found correlations for nonbrain traits, many of which were risk factors for stroke ($P < 2.4 \times 10^{-3}$). For example, high blood pressure was genetically correlated with many white matter tracts, including the superior longitudinal fasciculus, internal capsule, superior fronto-occipital fasciculus, uncinate fasciculus, and external capsule (P range = 5.7×10^{-11} to 2.3×10^{-3}). Hypertension can result in vascular stiffness and impaired cerebral perfusion, leading to microstructural white matter disruption (*78*) and stroke. We also observed strong genetic correlations with white matter hyperintensities (P range = 3.6×10^{-12} to 2.3×10^{-3}), which have been implicated in the connection between stroke and white matter (*79*). Additionally, we found that BMI, bone mineral density, smoking, and T2D were all genetically correlated with the corticospinal tract (P range = 3×10^{-13} to 2.3×10^{-3}). The corticospinal tract controls primary motor activity and is known to be closely related to stroke (*80*). Higher BMI, smoking, and T2D are risk factors for stroke and are associated with white matter changes (*81*, *82*). Low bone mineral density may also have correlation with white matter changes that contribute to stroke (*83*). Moreover, genetic correlations were observed between drinking and the corpus callosum (P range = 6×10^{-5} to 2.4×10^{-3}), whose atrophy is frequently observed in people with alcohol use disorders (*84*).

Next, we explored the genetic correlations between 215 DTI parameters and 101 regional brain volumes (*85*). After Bonferroni adjustment, there were 678 significant pairs between 132 DTI parameters and 62 regional brain volumes (215 × 101 tests; |gc| range = 0.18 to 0.85; $P < 2.2 \times 10^{-6}$; table S16). These results suggest the existence of widespread genetic overlaps between basic brain morphology and structural connectivity (figs. S17 and S18). Notably, most of the significant genetic correlations were observed between white matter tracts and their neighboring brain regions, which indicates the strong spatial colocalization of genetic covariation.

For example, lateral ventricle volumes were genetically correlated with the FA of white matter tracts surrounding the lateral ventricles, including the fornix, corona radiata, corpus callosum, posterior thalamic radiation, superior fronto-occipital fasciculus, internal capsule, and external capsule (Fig. 5, A and B). Both ventricular enlargement and altered white matter integrity were biomarkers of aging (*86*, *87*) and brain disorders, such as normal pressure hydrocephalus (*88*, *89*), Alzheimer's disease (*90*, *91*), and Parkinson's disease (*87*, *92*). Additionally, the FA of the corpus callosum, corona radiata, and posterior thalamic radiation had genetic correlations with the volumes of nearby cerebral cortex regions, such as the posterior cingulate, rostral anterior cingulate, lingual, isthmus cingulate, paracentral, and precuneus (Fig. 5C). These cortex regions have connections with many other brain regions through white matter tracts to support a wide range of cognitive functions. For example, the cingulate gyrus overlies the corpus callosum and serves as a neural hub among emotion, action, and memory (*93*), and the corona radiata consists of prominent projection fibers carrying information between the cerebral cortex and brainstem. Moreover, putamen and pallidum volumes were correlated with the FA of the internal capsule, external capsule, fornix, superior fronto-occipital fasciculus, and corona radiata. The putamen and pallidum lie on the lateral aspect of the internal capsule and separate the internal capsule from the external capsule (Fig. 5D), all of which are critical for human sensory motor functions (*94*). Similarly, volumes of other subcortical structures (such as the thalamus proper, accumbens area, and caudate) and the brainstem also had genetic correlations with adjacent white matter tracts, especially the corona radiata (fig. S19).

Downloaded from http://science.sciencemag.org/ on June 17, 2021

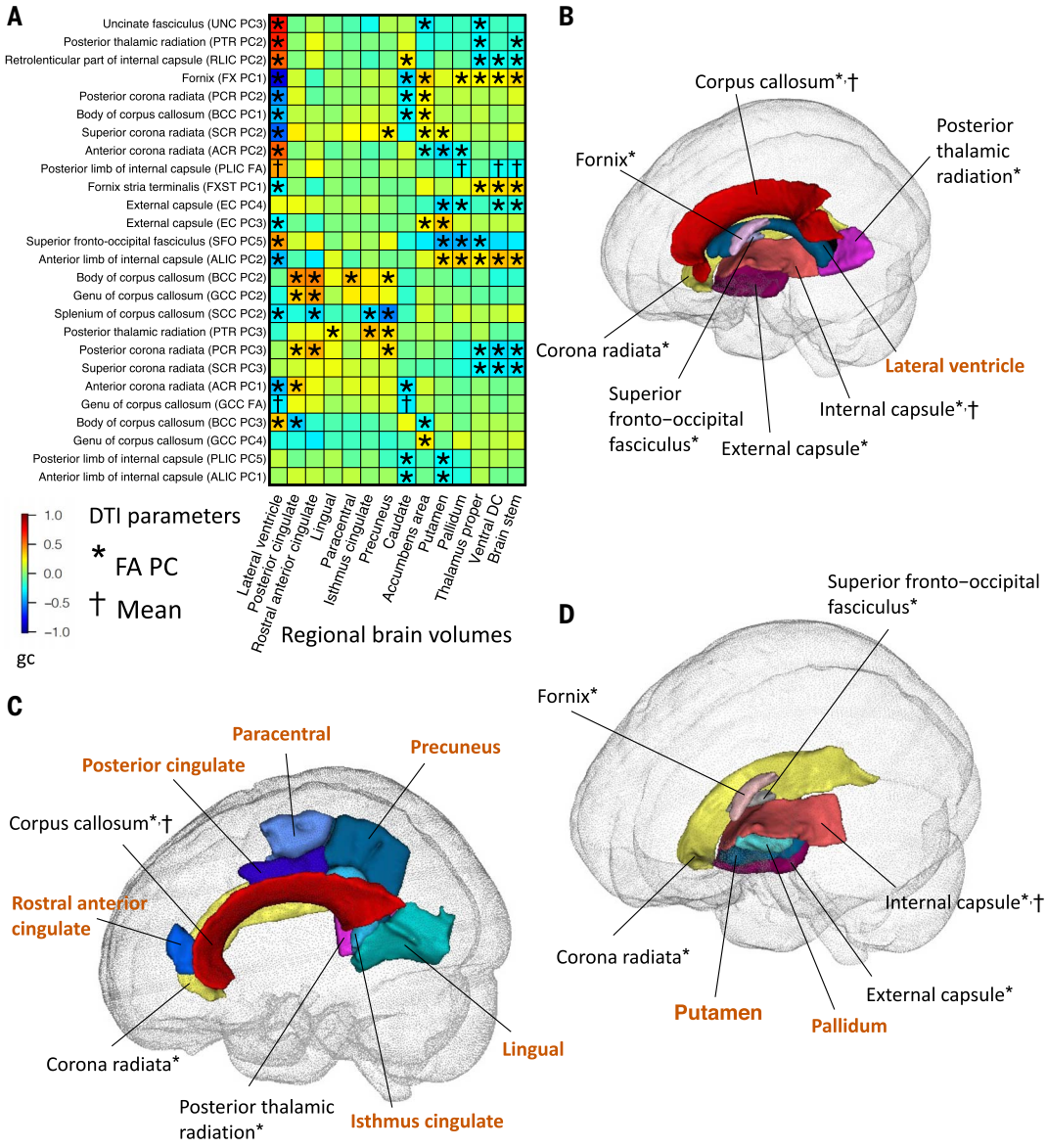


Fig. 5. Selected pairwise genetic correlations between FA of white matter tracts and regional brain volumes. (**A**) The asterisks (for FA PCs) and daggers (for mean parameters) highlight significant genetic correlations after Bonferroni adjustment for multiple testing. The *y* axis lists the DTI parameters of white matter tracts and the *x* axis provides the name of regional brain volumes. The colors represent genetic correlations (gc). (**B**) Location of the lateral ventricle region and its neighboring white matter tracts whose FA parameters were genetically correlated with the volume of the lateral ventricle. The colors describe different brain regions (brown labels) and white matter tracts. (**C**) Location of the cortex regions and their neighboring white matter tracts whose FA parameters were genetically correlated with the volume of these regions. (**D**) Location of the putamen and pallidum regions and their neighboring white matter tracts whose FA parameters were genetically correlated with the volume of these regions.

Gene-level analysis

We carried out MAGMA (*95*) gene-based association analysis for the 215 DTI parameters using our discovery GWAS summary statistics. There were 2772 significant gene-level associations ($P < 1.2 \times 10^{-8}$; adjusted for 215 phenotypes) between 487 genes and 178 DTI parameters (table S17), where 96 of the associated genes could only be discovered by PC parameters. We replicated 101 of 112 MAGMA genes reported in (*19*) and some other white matter–associated genes reported in previous studies (*20*–*22*, *44*), such as *SH3PXD2A* (10q24.33), *NBEAL1* (2q33.2), *TRIM47* (17q25.1), *VCAN* (5q14.3), *ALDH2* (12q24.12), *PLEKHG1* (6q25.1), *GNA12* (7p22.2), and *GNA13* (17q24.1). Many of our MAGMA significant genes have been linked to other complex traits and brain disease (table S18). For example, *EGFR* (7p11.2), *OBFC1* (10q24.33), *LMF1* (16p13.3), *SLC16A8* (22q13.1), and *CDKN2B* (9p21.3) are risk genes for glioma (*46*). Colocalization analysis (*96*) found that *EGFR* and *SLC16A8* had high colocalization probability (>0.9) with expression quantitative trait loci (eQTL) signals in GTEx (v8) (*97*) brain tissues (table S19). *EGFR* is critical for oligodendrocyte development and white matter recovery after neonatal hypoxia (*98*). It is known that *EGFR* plays an important role in the initiation of primary GBM and the progression of lower-grade glioma (*46*). In addition, many stroke risk genes, such as *CDH11* (16q21), *NTN4* (12q22), *COL4A1* (13q34), *FAT4* (4q28.1), *CACNB2* (10p12.31), *SUPT3H* (6p21.1), *ICA1L* (2q33.2), *CDKN2C* (1p32.3), *SH3PXD2A* (10q24.33), *ALDH2* (12q24.12), and *ADAMTS7* (15q25.1) (*50*, *99*), were associated with DTI parameters. These genes were also widely associated with hypertension, T2D, coronary artery disease, and BMI, which suggests the shared genetic components among stroke, white matter microstructure, and vascular risk factors.

Next, we mapped significant variants ($P < 2.3 \times 10^{-10}$) to genes according to physical position, eQTL association, and 3D chromatin (Hi-C) interaction through FUMA (*38*). FUMA yielded 1183 associated genes (1545 in total) that were not discovered in MAGMA analysis (table S20), replicating 282 of the 292 FUMA genes identified in (*19*) and more genes from previous studies of white matter, such as *CLDN23* (*100*) (8p23.1) and *EFEMP1* (*20*, *21*, *44*) (2p16.1). In particular, 911 FUMA genes were solely mapped by significant Hi-C interactions in brain tissues (table S21), demonstrating the power of integrating chromatin interaction profiles in GWAS of white matter. More overlapping genes were observed between white matter microstructure and other traits

Downloaded from http://science.sciencemag.org/ on June 17, 2021

(table S22). For example, glioma risk genes from five more regions were associated with white matter microstructure, including *RAVER2* (1p31.3), *MDM4* (1q32.1), *ZBTB16* (11q23.2), *AKT3* (1q44), and *MAML2* (11q21) (*46*). *AKT3* is hyperexpressed and activated in glioma (*101*) and is also associated with brain developmental disorders (*102*), learning and memory deficits (*103*), and disrupted integrity of white matter (*104*). Together, our variant-level and gene-level analyses uncovered the shared genetic influences with glioma in 11 different genomic regions. Considering the largest glioma GWAS (*46*) reported 25 risk regions, our results reveal the close genetic relationship between glioma and white matter integrity.

We performed drug target lookups in a recently established drug target network (*105*), which included 273 nervous system drugs [where the anatomical therapeutic chemical (ATC) code starts with N] and 241 targeted genes. We found that 14 white matter–associated genes were targets for 79 drugs, 23 of which were antipsychotics (ATC: N05A, target such as *CALM1*) to manage psychosis like schizophrenia and bipolar disorder, 38 were antidepressants (ATC: N06A, *SLC6A4*) to treat MDD and other conditions, 8 were used to treat addictive disorders (ATC: N07B, *ALDH2*), 6 were anti-Parkinson drugs (ATC: N04B, *HTR2B*), 4 were antidementia drugs (ATC: N06D, *CHRNB4*), and 3 were anticonvulsants (ATC: N03A, *GABBR1*) used in the treatment of epileptic seizures (table S23). Additionally, we treated white matter–associated genes as an annotation and performed partitioned heritability enrichment analysis (*106*) for the other 70 complex traits studied in our genetic correlation analysis. At the FDR 5% level, heritability of 49 complex traits was significantly (P range = 1.2×10^{-8} to 3.2×10^{-2}) enriched in regions influencing white matter microstructure, such as stroke, schizophrenia, ADHD, bipolar disorder, Alzheimer's disease, T2D, high blood pressure, and coronary artery disease (fig. S20 and table S24). Overall, these results suggest the potential clinical values of the genes identified for white matter microstructure.

Biological annotations

To identify tissues and cell types where genetic variation leads to changes in white matter microstructure, we performed partitioned heritability analyses (*106*) from the GWAS of global FA and MD within tissue type and cell type specific regulatory elements. First, we used regulatory elements across multiple adult and fetal tissues (*107*). As expected, both FA and MD had the strongest enrichment of heritability in active gene regulation regions of brain tissues (fig. S21 and table S25). To identify gross cell types, we again performed partitioned heritability analysis using chromatin accessibility data of two brain cell types, neurons (NeuN+) and glia (NeuN−) sampled from 14 brain regions, including both cortical and subcortical (*108*). For all regions, we found that enrichment of FA and MD heritability existed in glial but not neuronal regulatory elements (Fig. 6A). These results are expected as white matter is largely composed of glial cell types. For further resolution on cell types, we tested partitioned heritability enrichment within differentially accessible chromatin of glial cell subtypes, oligodendrocyte (NeuN−/Sox10+) and microglia and astrocyte (NeuN−/Sox10−); and neuronal cell subtypes, GABAergic (γ-aminobutyric acid–releasing) (NeuN+/Sox6+) and glutamatergic (NeuN+/Sox6−) neurons (*109*). Heritability of FA and MD was enriched in oligodendrocyte and microglia/astrocyte annotations. By contrast, no enrichment was observed in neurons (Fig. 6B). We also performed regional-specific enrichment analysis for FA in each white matter tract. Glial cell enrichment was observed in different white matter tracts, and the strongest enrichments were found in the posterior corona radiata, posterior limb of the internal capsule, and genu of the corpus callosum (Fig. 6C). These analyses imply that common variants associated with white matter microstructure alter the function of regulatory elements in glial cells, particularly oligodendrocytes, the cell type expected to influence white matter microstructure, supporting the biological validity of the genetic associations.

Next, we performed MAGMA gene property (*95*) analysis for 13 GTEx (v8) brain tissues to examine whether tissue-specific gene expression levels were related to significance of the correlations between genes and DTI parameters. At the FDR 5% level, 63 significant (P range = 7×10^{-16} to 8.4×10^{-4}) associations were detected between DTI parameters and gene expression data sampled in GTEx v8 brain tissues (table S26), which suggests that genes with higher transcription levels in brain tissues also had stronger genetic associations with DTI parameters. We also applied DEPICT (*110*) gene-set enrichment testing for 10,968 preconstituted gene sets to prioritize enriched biological pathways. At FDR 5% levels, DEPICT found 381 significant gene sets (10,968 × 215 tests; $P < 1 \times 10^{-5}$; table S27). Several gene sets of rat sarcoma (Ras) proteins, small guanosine triphosphate hydrolases (GTPases), and Rho family GTPases were prioritized by DEPICT, such as "small GTPase mediated signal transduction" [gene ontology (GO) term: 0007264], "Rho GTPase binding" (GO: 0017048), "Ras protein signal transduction" (GO: 0007265), "reactome Rho GTPase cycle" (M27078), and "reactome signaling by Rho GTPases" (M501). The activity of Ras proteins is involved in developmental processes and abnormalities of neural cells in the central nervous system (*111*); small and Rho family GTPases play crucial roles in basic cellular processes during the entire neurodevelopment process and are closely connected to several neurological disorders (*112*, *113*). We also observed enrichment in pathways related to glial cells and the nervous system, such as "glial cell differentiation" (GO: 0010001), "abnormal radial glial cell morphology" [mammalian phenotype (MP) term: 0003648], "brain development" (GO: 0007420), "abnormal forebrain morphology" (MP: 0000783), "open neural tube" (MP: 0000929), "regulation of neurogenesis" (GO: 0050767), "abnormal neuronal migration" (MP: 0006009), "central nervous system neuron development" (GO: 0021954), and "abnormal neural fold formation" (MP: 0004837).

Discussion

In this study, we analyzed the genetic architecture of brain white matter using dMRI scans of 43,802 subjects collected from five publicly accessible data resources. We identified 109 additional genomic regions for white matter microstructural differences. Many previously reported genetic hits were confirmed in our discovery GWAS, and we further validated our discovery GWAS in several replication cohorts. We evaluated the genetic relationships between white matter and a wide variety of complex traits in association lookups, genetic correlation estimations, and gene-level analyses. A large proportion of our findings were revealed by unconventional tract-specific PC parameters. Bioinformatics analyses found tissue- and cell-specific functional enrichments and many enriched biological pathways. Together, these results suggest the value of large-scale neuroimaging data integration and the application of tract-specific FPCA in studying the genetics of the human brain.

Many efforts have been made to study the gray matter abnormalities associated with brain diseases and cognition. Our results indicate that a better understanding of the etiopathogenesis and treatment of many brain disorders could be achieved from a more balanced perspective that includes both gray matter regions and white matter tracts. Moreover, we found that some commonly used centrally-acting medications have effects on genes with associations to white matter microstructure, which suggests that the neuropharmacology of brain disorders may potentially benefit from knowledge of the relationship between these medications and white matter. Many drugs for cognitive disorders have not been fully understood, and better knowledge of how these drugs work in the white matter of the brain could substantially improve patient care.

One limitation of the present study is that most of the publicly available dMRI data are from subjects of European ancestry, and our discovery GWAS focused on UKB British individuals. Such GWAS strategy can efficiently avoid false discoveries as a result of population

Downloaded from http://science.sciencemag.org/ on June 17, 2021

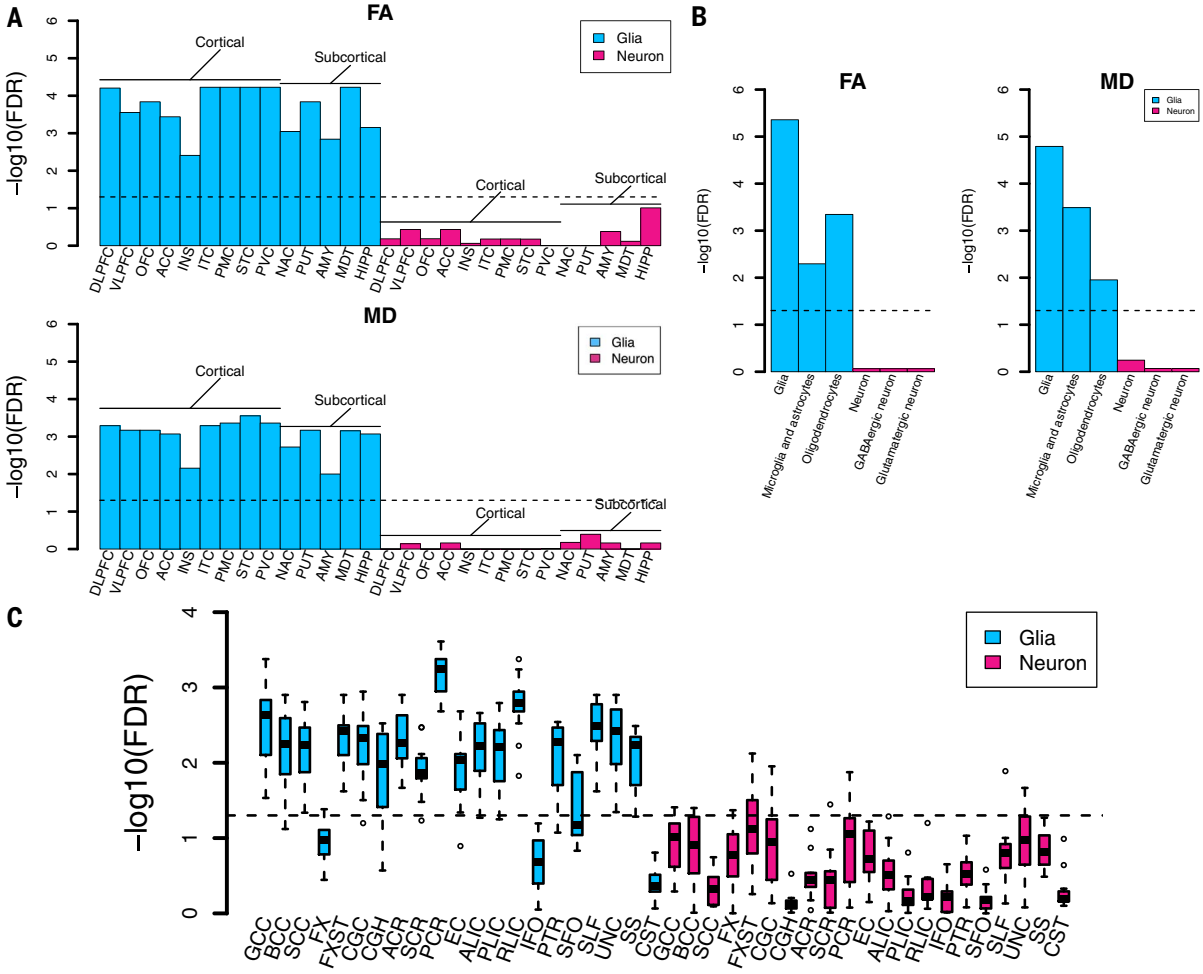


Fig. 6. Partitioned heritability enrichment analysis in brain cell types. (**A**) Heritability enrichment of global mean FA and MD in regulatory elements of two brain cell types (neuron and glia) sampled from 14 brain cortical and subcortical regions. DLFPC, dorsolateral prefrontal cortex; VLPFC, ventrolateral prefrontal cortex; OFC, orbitofrontal cortex; ACC, anterior cingulate cortex; INS, insular cortex; ITC, inferior temporal cortex; STC, superior temporal cortex; PMC, primary motor cortex; PVC, primary visual cortex; AMY, amygdala; HIPP, hippocampus; MDT, mediodorsal thalamus; NAC, nucleus accumbens; PUT, putamen. (**B**) Heritability enrichment of global mean FA and MD in regulatory elements of glial cell subtypes (glia, including oligodendrocyte and microglia or astrocyte) and neuronal cell subtypes (neurons, including GABAergic and glutamatergic neurons). (**C**) Heritability enrichment of mean FA of 21 white tracts in regulatory elements of two brain cell types (neuron and glia) sampled from 14 brain cortical and subcortical regions. Full names of the 21 white matter tracts can be found in Fig. 1B. The dashed lines indicate the significance level after controlling the FDR at 5% level.

stratifications and heterogeneities across studies (*17*, *114*), but it may raise questions as to the degree to which the research findings can be generalized and applied to global populations (*115*). In our analysis, the UKB British–derived PRS still showed prediction power in Hispanic, Asian, and Black or African American testing cohorts but had reduced performance. This may indicate that the genetic architecture of white matter is similar but not the same across different populations. Identifying the cross-population and population-specific components of genetic factors for the human brain could be an interesting future topic. As more non-European neuroimaging data become available [for example, the ongoing CHIMGEN project (*116*) in the Chinese population], global integration efforts will be needed to study the comparative genetic architectures and to explore the multiethnic genetic relationships among the brain and other human complex traits. Specifically, the anticipated non-European data will be relevant to non-European ancestry groups, who often endure worse outcomes with respect to brain diseases that implicate white matter, such as stroke (*117*) and Alzheimer's disease (*118*). Additionally, the present study focused on DTI parameters. Previous studies have found that some beyond-DTI parameters, such as white matter tract integrity (WMTI) (*119*), neurite orientation dispersion and density imaging (NODDI) (*120*), and diffusional kurtosis imaging (DKI) (*121*) may provide additional information about microstructural properties (*122*). The g-ratio measure (*123*, *124*) is probably more directly correlated to neuronal conduction velocity and thus is of interest in assessing interhemispheric transfer time and information processing speed. Future work is needed to explore these beyond-DTI parameters in biobank-scale neuroimaging datasets. Finally, although we have harmonized the image processing pipelines to minimize the influences of imaging artifacts, confounding factors may still exist and limit the interpretation of our results. For example, the partial-volume effect (*125*) in DTI data could be a confounding factor that drives part of the observed genetic correlations between ventricle volumes and the FA of white matter tracts. More analysis is required to model and examine potential neuroimaging confounders (*126*).

Methods summary

We conducted a GWAS of white matter microstructure using dMRI data from 43,802 individuals. Measures of white matter microstructure

Downloaded from http://science.sciencemag.org/ on June 17, 2021

were derived from DTI models using the ENIGMA-DTI pipeline (*30*, *31*). We analyzed five major DTI measurements along 21 cerebral white matter tracts. For each individual, 215 DTI phenotypes were generated. We analyzed these phenotypes in subsequent genetic analyses.

Our discovery GWAS used data from UKB individuals of British ancestry. To replicate our results in the UKB British GWAS, we performed additional GWAS with data from nine independent European and non-European datasets. We checked whether the associations detected in the UKB British GWAS can be replicated in these independent GWAS. Additionally, we used polygenic risk scores to predict the proportion of variance of DTI phenotypes that can be explained by genetic variants in European and non-European cohorts.

SNP heritability explained by all autosomal variants was estimated using GCTA-GREML analysis (*34*). We used LDSC (*36*) to estimate genetic correlations between DTI parameters and other complex traits. Partitioned LDSC (*106*) was used to estimate the proportion of heritability explained by annotated regions of the genome. Heritability enrichment was estimated in multiple adult and fetal tissues (*107*), brain cell types (glia and neurons) (*108*), and glial and neuronal cell subtypes (*109*).

Gene-based association analysis was performed using MAGMA (*95*). We applied fastENLOC (*96*) to perform colocalization analysis with GTEx eQTL annotations. DEPICT (*110*) was used to explore the implicated biological pathways. Functional impact was investigated using FUMA (*38*), in which variants were linked to genes by a combination of positional, eQTL, and 3D chromatin interaction mappings.

REFERENCES AND NOTES

1. M. P. van den Heuvel, O. Sporns, A cross-disorder connectome landscape of brain dysconnectivity. *Nat. Rev. Neurosci.* **20**, 435–446 (2019). doi: 10.1038/s41583-019-0177-6; pmid: 31127193
2. P. Hagmann *et al.*, Mapping the structural core of human cerebral cortex. *PLOS Biol.* **6**, e159 (2008). doi: 10.1371/journal.pbio.0060159; pmid: 18597554
3. B. A. Zielinski, E. D. Gennatas, J. Zhou, W. W. Seeley, Network-level structural covariance in the developing brain. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 18191–18196 (2010). doi: 10.1073/pnas.1003109107; pmid: 20921389
4. C. M. Filley, R. D. Fields, White matter and cognition: Making the connection. *J. Neurophysiol.* **116**, 2093–2104 (2016). doi: 10.1152/jn.00221.2016; pmid: 27512019
5. L. F. Koziol, D. E. Budding, D. Chidekel, Adaptation, expertise, and giftedness: Towards an understanding of cortical, subcortical, and cerebellar network contributions. *Cerebellum* **9**, 499–529 (2010). doi: 10.1007/s12311-010-0192-7; pmid: 20680539
6. J. D. Schmahmann, E. E. Smith, F. S. Eichler, C. M. Filley, Cerebral white matter: Neuroanatomy, clinical neurology, and neurobehavioral correlates. *Ann. N. Y. Acad. Sci.* **1142**, 266–309 (2008). doi: 10.1196/annals.1444.017; pmid: 18990132
7. E. H. Chang *et al.*, The role of myelination in measures of white matter integrity: Combination of diffusion tensor imaging and two-photon microscopy of CLARITY intact brains. *Neuroimage* **147**, 253–261 (2017). doi: 10.1016/j.neuroimage.2016.11.068; pmid: 27986605
8. Z.-M. Wu *et al.*, White matter microstructural alterations in children with ADHD: Categorical and dimensional perspectives. *Neuropsychopharmacology* **42**, 572–580 (2017). doi: 10.1038/npp.2016.223; pmid: 27681441
9. K. Zou *et al.*, Alterations of white matter integrity in adults with major depressive disorder: A magnetic resonance imaging study. *J. Psychiatry Neurosci.* **33**, 525–530 (2008). pmid: 18982175
10. S. Cetin-Karayumak *et al.*, White matter abnormalities across the lifespan of schizophrenia: A harmonized multi-site diffusion MRI study. *Mol. Psychiatry* **25**, 3208–3219 (2020). doi: 10.1038/s41380-019-0509-y; pmid: 31511636
11. S. Lee *et al.*, White matter hyperintensities are a core feature of Alzheimer's disease: Evidence from the dominantly inherited Alzheimer network. *Ann. Neurol.* **79**, 929–939 (2016). doi: 10.1002/ana.24647; pmid: 27016429
12. P. J. Basser, J. Mattiello, D. LeBihan, Estimation of the effective self-diffusion tensor from the NMR spin echo. *J. Magn. Reson. B.* **103**, 247–254 (1994). doi: 10.1006/jmrb.1994.1037; pmid: 8019776
13. S. M. Grieve, L. M. Williams, R. H. Paul, C. R. Clark, E. Gordon, Cognitive aging, executive function, and fractional anisotropy: A diffusion tensor MR imaging study. *AJNR Am. J. Neuroradiol.* **28**, 226–235 (2007). pmid: 17296985
14. W.-S. Tae, B.-J. Ham, S.-B. Pyun, S.-H. Kang, B.-J. Kim, Current clinical applications of diffusion-tensor imaging in neurological disorders. *J. Clin. Neurol.* **14**, 129–140 (2018). doi: 10.3988/jcn.2018.14.2.129; pmid: 29504292
15. D. B. Ennis, G. Kindlmann, Orthogonal tensor invariants and the analysis of diffusion tensor magnetic resonance images. *Magn. Reson. Med.* **55**, 136–146 (2006). doi: 10.1002/mrm.20741; pmid: 16342267
16. E. Vuoksimaa *et al.*, Heritability of white matter microstructure in late middle age: A twin study of tract-based fractional anisotropy and absolute diffusivity indices. *Hum. Brain Mapp.* **38**, 2026–2036 (2017). doi: 10.1002/hbm.23502; pmid: 28032374
17. L. T. Elliott *et al.*, Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature* **562**, 210–216 (2018). doi: 10.1038/s41586-018-0571-7; pmid: 30305740
18. P. Kochunov *et al.*, Heritability of fractional anisotropy in human white matter: A comparison of Human Connectome Project and ENIGMA-DTI data. *Neuroimage* **111**, 300–311 (2015). doi: 10.1016/j.neuroimage.2015.02.050; pmid: 25747917
19. B. Zhao *et al.*, Large-scale GWAS reveals genetic architecture of brain white matter microstructure and genetic overlap with cognitive and mental health traits (n = 17,706). *Mol. Psychiatry* 10.1038/s41380-019-0569-z (2019). doi: 10.1038/s41380-019-0569-z; pmid: 31666681
20. L. C. A. Rutten-Jacobs *et al.*, Genetic study of white matter integrity in UK Biobank (N=8448) and the overlap with stroke, depression, and dementia. *Stroke* **49**, 1340–1347 (2018). doi: 10.1161/STROKEAHA.118.020811; pmid: 29752348
21. M. Traylor *et al.*, Genetic variation in *PLEKHG1* is associated with white matter hyperintensities (n = 11,226). *Neurology* **92**, e749–e757 (2019). doi: 10.1212/WNL.0000000000006952; pmid: 30659137
22. M. Fornage *et al.*, Genome-wide association studies of cerebral white matter lesion burden. *Ann. Neurol.* **69**, 928–939 (2011). doi: 10.1002/ana.22403; pmid: 21681796
23. P. Adib-Samii *et al.*, 17q25 Locus is associated with white matter hyperintensity volume in ischemic stroke, but not with lacunar stroke status. *Stroke* **44**, 1609–1615 (2013). doi: 10.1161/STROKEAHA.113.679936; pmid: 23674528
24. Z. Safadi *et al.*, Functional segmentation of the anterior limb of the internal capsule: Linking white matter abnormalities to specific connections. *J. Neurosci.* **38**, 2106–2117 (2018). doi: 10.1523/JNEUROSCI.2335-17.2017; pmid: 29358360
25. C. Sudlow *et al.*, UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLOS Med.* **12**, e1001779 (2015). doi: 10.1371/journal.pmed.1001779; pmid: 25826379
26. B. J. Casey *et al.*, The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* **32**, 43–54 (2018). doi: 10.1016/j.dcn.2018.03.001; pmid: 29567376
27. L. H. Somerville *et al.*, The Lifespan Human Connectome Project in Development: A large-scale study of brain connectivity development in 5-21 year olds. *Neuroimage* **183**, 456–468 (2018). doi: 10.1016/j.neuroimage.2018.08.050; pmid: 30142446
28. T. L. Jernigan *et al.*, The pediatric imaging, neurocognition, and genetics (PING) data repository. *Neuroimage* **124**, 1149–1154 (2016). doi: 10.1016/j.neuroimage.2015.04.057; pmid: 25937488
29. T. D. Satterthwaite *et al.*, Neuroimaging of the Philadelphia neurodevelopmental cohort. *Neuroimage* **86**, 544–553 (2014). doi: 10.1016/j.neuroimage.2013.07.064; pmid: 23921101
30. N. Jahanshad *et al.*, Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: A pilot project of the ENIGMA-DTI working group. *Neuroimage* **81**, 455–469 (2013). doi: 10.1016/j.neuroimage.2013.04.061; pmid: 23629049
31. P. Kochunov *et al.*, Multi-site study of additive genetic effects on fractional anisotropy of cerebral white matter: Comparing meta and megaanalytical approaches for data pooling. *Neuroimage* **95**, 136–150 (2014). doi: 10.1016/j.neuroimage.2014.03.033; pmid: 24657781
32. P. Hall, H.-G. Müller, J.-L. Wang, Properties of principal component methods for functional and longitudinal data analysis. *Ann. Stat.* **34**, 1493–1517 (2006). doi: 10.1214/009053606000000272
33. A. Srivastava, E. P. Klassen, *Functional and Shape Data Analysis* (Springer, 2016).
34. J. Yang, S. H. Lee, M. E. Goddard, P. M. Visscher, GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011). doi: 10.1016/j.ajhg.2010.11.011; pmid: 21167468
35. G. Kichaev *et al.*, Leveraging polygenic functional enrichment to improve GWAS power. *Am. J. Hum. Genet.* **104**, 65–75 (2019). doi: 10.1016/j.ajhg.2018.11.008; pmid: 30595370
36. B. Bulik-Sullivan *et al.*, An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015). doi: 10.1038/ng.3406; pmid: 26414676
37. B. K. Bulik-Sullivan *et al.*, LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015). doi: 10.1038/ng.3211; pmid: 25642630
38. K. Watanabe, E. Taskesen, A. van Bochoven, D. Posthuma, Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017). doi: 10.1038/s41467-017-01261-5; pmid: 29184056
39. N. Matoba, M. I. Love, J. L. Stein, Evaluating brain structure traits as endophenotypes using polygenicity and discoverability. *Hum. Brain Mapp.* 10.1002/hbm.25257 (2020). doi: 10.1002/hbm.25257; pmid: 33098356
40. I. E. Jansen *et al.*, Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat. Genet.* **51**, 404–413 (2019). doi: 10.1038/s41588-018-0311-9; pmid: 30617256
41. A. Buniello *et al.*, The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019). doi: 10.1093/nar/gky1120; pmid: 30445434
42. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium, Early Growth Genetics (EGG) Consortium, Common variants at 6q22 and 17q21 are associated with intracranial volume. *Nat. Genet.* **44**, 539–544 (2012). doi: 10.1038/ng0612-732c; pmid: 22504418
43. D. Vojinovic *et al.*, Genome-wide association study of 23,500 individuals identifies 7 loci associated with brain ventricular volume. *Nat. Commun.* **9**, 3945 (2018). doi: 10.1038/s41467-018-06234-w; pmid: 30258056
44. B. F. Verhaaren *et al.*, Multiethnic genome-wide association study of cerebral white matter hyperintensities on MRI. *Circ. Cardiovasc. Genet.* **8**, 398–409 (2015). doi: 10.1161/CIRCGENETICS.114.000858; pmid: 25663218
45. D. van der Meer *et al.*, Brain scans from 21,297 individuals reveal the genetic architecture of hippocampal subfield volumes. *Mol. Psychiatry* **25**, 3053–3065 (2020). doi: 10.1038/s41380-018-0262-7; pmid: 30279459
46. B. S. Melin *et al.*, Genome-wide association study of glioma subtypes identifies specific differences in genetic susceptibility to glioblastoma and non-glioblastoma tumors. *Nat. Genet.* **49**, 789–794 (2017). doi: 10.1038/ng.3823; pmid: 28346443
47. D. Liu *et al.*, Alterations of white matter integrity associated with cognitive deficits in patients with glioma. *Brain Behav.* **10**, e01639 (2020). doi: 10.1002/brb3.1639; pmid: 32415731
48. J. Wang *et al.*, Invasion of white matter tracts by glioma stem cells is regulated by a *NOTCH1-SOX2* positive-feedback loop.

Downloaded from http://science.sciencemag.org/ on June 17, 2021

Nat. Neurosci. **22**, 91–105 (2019). doi: 10.1038/s41593-018-0285-z; pmid: 30559479
49. G. Sharifi, A. M. Pajavand, S. Nateghinia, T. E. Meybodi, H. Hasooni, Glioma Migration Through the Corpus Callosum and the Brainstem Detected by Diffusion and Magnetic Resonance Imaging: Initial Findings. *Front. Hum. Neurosci.* **13**, 472 (2020). doi: 10.3389/fnhum.2019.00472; pmid: 32161524
50. R. Malik *et al*., Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat. Genet.* **50**, 524–537 (2018). doi: 10.1038/s41588-018-0058-3; pmid: 29531354
51. M. L. Hackett, S. Köhler, J. T. O'Brien, G. E. Mead, Neuropsychiatric outcomes of stroke. *Lancet Neurol.* **13**, 525–534 (2014). doi: 10.1016/S1474-4422(14)70016-X; pmid: 24685278
52. G. Schlaug, B. Siewert, A. Benfield, R. R. Edelman, S. Warach, Time course of the apparent diffusion coefficient (ADC) abnormality in human stroke. *Neurology* **49**, 113–119 (1997). doi: 10.1212/WNL.49.1.113; pmid: 9222178
53. M. Ikeda *et al*., Genome-wide association study detected novel susceptibility genes for schizophrenia and shared trans-populations/diseases genetic effect. *Schizophr. Bull.* **45**, 824–834 (2019). doi: 10.1093/schbul/sby140; pmid: 30285260
54. B. W. Kunkle *et al*., Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nat. Genet.* **51**, 414–430 (2019). doi: 10.1038/s41588-019-0358-2; pmid: 30820047
55. N. Kouri *et al*., Genome-wide association study of corticobasal degeneration identifies risk variants shared with progressive supranuclear palsy. *Nat. Commun.* **6**, 7247 (2015). doi: 10.1038/ncomms8247; pmid: 26077951
56. J. Simón-Sánchez *et al*., Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat. Genet.* **41**, 1308–1312 (2009). doi: 10.1038/ng.487; pmid: 19915575
57. J. Grove *et al*., Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **51**, 431–444 (2019). doi: 10.1038/s41588-019-0344-8; pmid: 30804558
58. S. Sanchez-Roige *et al*., Genome-wide association study meta-analysis of the Alcohol Use Disorders Identification Test (AUDIT) in two population-based cohorts. *Am. J. Psychiatry* **176**, 107–118 (2019). doi: 10.1176/appi.ajp.2018.18040369; pmid: 30336701
59. C. L. Hyde *et al*., Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat. Genet.* **48**, 1031–1036 (2016). doi: 10.1038/ng.3623; pmid: 27479909
60. J. J. Lee *et al*., Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018). doi: 10.1038/s41588-018-0147-3; pmid: 30038396
61. G. Davies *et al*., Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat. Commun.* **9**, 2098 (2018). doi: 10.1038/s41467-018-04362-x; pmid: 29844566
62. M. Luciano *et al*., Association analysis in over 329,000 individuals identifies 116 independent variants influencing neuroticism. *Nat. Genet.* **50**, 6–11 (2018). doi: 10.1038/s41588-017-0013-8; pmid: 29255261
63. B. M. L. Baselmans *et al*., Multivariate genome-wide analyses of the well-being spectrum. *Nat. Genet.* **51**, 445–451 (2019). doi: 10.1038/s41588-018-0320-8; pmid: 30643256
64. R. Karlsson Linnér *et al*., Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nat. Genet.* **51**, 245–257 (2019). doi: 10.1038/s41588-018-0309-3; pmid: 30643258
65. M. Liu *et al*., Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat. Genet.* **51**, 237–244 (2019). doi: 10.1038/s41588-018-0307-5; pmid: 30643251
66. H. S. Dashti *et al*., Genome-wide association study identifies genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates. *Nat. Commun.* **10**, 1100 (2019). doi: 10.1038/s41467-019-08917-4; pmid: 30846698
67. S. E. Jones *et al*., Genome-wide association analyses of chronotype in 697,828 individuals provides insights into circadian rhythms. *Nat. Commun.* **10**, 343 (2019). doi: 10.1038/s41467-018-08259-7; pmid: 30696823
68. F. Yasuno *et al*., Microstructural abnormality in white matter, regulatory T lymphocytes, and depressive symptoms after stroke. *Psychogeriatrics* **14**, 213–221 (2014). doi: 10.1111/psyg.12084; pmid: 25495082
69. Z. Liu *et al*., Association between white matter impairment and cognitive dysfunction in patients with ischemic Moyamoya disease. *BMC Neurol.* **20**, 302 (2020). doi: 10.1186/s12883-020-01876-0; pmid: 32799829
70. L. Li *et al*., Population-based case-control study of white matter changes on brain imaging in transient ischemic attack and ischemic stroke. *Stroke* **44**, 3063–3070 (2013). doi: 10.1161/STROKEAHA.113.002775; pmid: 24021688
71. S. Ran *et al*., Atrophic Corpus Callosum Associated with Altered Functional Asymmetry in Major Depressive Disorder. *Neuropsychiatr. Dis. Treat.* **16**, 1473–1482 (2020). doi: 10.2147/NDT.S245078; pmid: 32606700
72. J. Xiao, Y. He, C. M. McWhinnie, S. Yao, Altered white matter integrity in individuals with cognitive vulnerability to depression: A tract-based spatial statistics study. *Sci. Rep.* **5**, 9738 (2015). doi: 10.1038/srep09738; pmid: 25984712
73. L. Penke *et al*., Brain white matter tract integrity as a neural foundation for general intelligence. *Mol. Psychiatry* **17**, 1026–1030 (2012). doi: 10.1038/mp.2012.66; pmid: 22614288
74. B. G. Travers *et al*., Diffusion tensor imaging in autism spectrum disorder: A review. *Autism Res.* **5**, 289–313 (2012). doi: 10.1002/aur.1243; pmid: 22786754
75. J.-G. Gehricke *et al*., The brain anatomy of attention-deficit/hyperactivity disorder in young adults - a magnetic resonance imaging study. *PLOS ONE* **12**, e0175433 (2017). doi: 10.1371/journal.pone.0175433; pmid: 28406942
76. X. Li *et al*., Impaired white matter connections of the limbic system networks associated with impaired emotional memory in Alzheimer's disease. *Front. Aging Neurosci.* **8**, 250 (2016). doi: 10.3389/fnagi.2016.00250; pmid: 27833549
77. K. H. Karlsgodt *et al*., Diffusion tensor imaging of the superior longitudinal fasciculus and working memory in recent-onset schizophrenia. *Biol. Psychiatry* **63**, 512–518 (2008). doi: 10.1016/j.biopsych.2007.06.017; pmid: 17720147
78. Y. Hannawi *et al*., Hypertension is associated with white matter disruption in apparently healthy middle-aged individuals. *AJNR Am. J. Neuroradiol.* **39**, 2243–2248 (2018). doi: 10.3174/ajnr.A5871; pmid: 30442693
79. N. J. Armstrong *et al*., Common genetic variation indicates separate causes for periventricular and deep white matter hyperintensities. *Stroke* **51**, 2111–2121 (2020). doi: 10.1161/STROKEAHA.119.027544; pmid: 32517579
80. S. Maraka *et al*., Degree of corticospinal tract damage correlates with motor function after stroke. *Ann. Clin. Transl. Neurol.* **1**, 891–899 (2014). doi: 10.1002/acn3.132; pmid: 25540803
81. Y. Xiong *et al*., A diffusion tensor imaging study on white matter abnormalities in patients with type 2 diabetes using tract-based spatial statistics. *AJNR Am. J. Neuroradiol.* **37**, 1462–1469 (2016). doi: 10.3174/ajnr.A4740; pmid: 26988810
82. T. M. Wassenaar, K. Yaffe, Y. D. van der Werf, C. E. Sexton, Associations between modifiable risk factors and white matter of the aging brain: Insights from diffusion tensor imaging studies. *Neurobiol. Aging* **80**, 56–70 (2019). doi: 10.1016/j.neurobiolaging.2019.04.006; pmid: 31103633
83. Y. K. Minn, S. H. Suk, S. Y. Do, Osteoporosis as an independent risk factor for silent brain infarction and white matter changes in men and women: The PRESENT project. *Osteoporos. Int.* **25**, 2465–2469 (2014). doi: 10.1007/s00198-014-2785-3; pmid: 25011984
84. M. A. Kashem, G. James, C. Harper, P. Wilce, I. Matsumoto, Differential protein expression in the corpus callosum (splenium) of human alcoholics: A proteomics study. *Neurochem. Int.* **50**, 450–459 (2007). doi: 10.1016/j.neuint.2006.10.009; pmid: 17141922
85. B. Zhao *et al*., Genome-wide association analysis of 19,629 individuals identifies variants influencing regional brain volumes and refines their genetic co-architecture with cognitive and mental health traits. *Nat. Genet.* **51**, 1637–1644 (2019). doi: 10.1038/s41588-019-0516-6; pmid: 31676860
86. C. E. Hugenschmidt *et al*., Relating imaging indices of white matter integrity and volume in healthy older adults. *Cereb. Cortex* **18**, 433–442 (2008). doi: 10.1093/cercor/bhm080; pmid: 17575289
87. L. G. Apostolova *et al*., Hippocampal atrophy and ventricular enlargement in normal aging, mild cognitive impairment (MCI), and Alzheimer Disease. *Alzheimer Dis. Assoc. Disord.* **26**, 17–27 (2012). doi: 10.1097/WAD.0b013e3182163b62; pmid: 22343374
88. B. P. Damasceno, Neuroimaging in normal pressure hydrocephalus. *Dement. Neuropsychol.* **9**, 350–355 (2015). doi: 10.1590/1980-57642015DN94000350; pmid: 29213984
89. S. J. Teipel *et al*., Fractional anisotropy changes in Alzheimer's disease depend on the underlying fiber tract architecture: A multiparametric DTI study using joint independent component analysis. *J. Alzheimers Dis.* **41**, 69–83 (2014). doi: 10.3233/JAD-131829; pmid: 24577476
90. S. M. Nestor *et al*., Ventricular enlargement as a possible measure of Alzheimer's disease progression validated using the Alzheimer's disease neuroimaging initiative database. *Brain* **131**, 2443–2454 (2008). doi: 10.1093/brain/awn146; pmid: 18669512
91. T. Koyama, K. Marumoto, K. Domen, T. Ohmura, H. Miyake, Diffusion tensor imaging of idiopathic normal pressure hydrocephalus: A voxel-based fractional anisotropy study. *Neurol. Med. Chir.* **52**, 68–74 (2012). doi: 10.2176/nmc.52.68; pmid: 22362286
92. C. Atkinson-Clement, S. Pinto, A. Eusebio, O. Coulon, Diffusion tensor imaging in Parkinson's disease: Review and meta-analysis. *Neuroimage Clin.* **16**, 98–110 (2017). doi: 10.1016/j.nicl.2017.07.011; pmid: 28765809
93. E. T. Rolls, The cingulate cortex and limbic systems for emotion, action, and memory. *Brain Struct. Funct.* **224**, 3001–3018 (2019). doi: 10.1007/s00429-019-01945-2; pmid: 31451898
94. A. Stocco, C. Lebiere, J. R. Anderson, Conditional routing of information to the cortex: A model of the basal ganglia's role in cognitive coordination. *Psychol. Rev.* **117**, 541–574 (2010). doi: 10.1037/a0019077; pmid: 20438237
95. C. A. de Leeuw, J. M. Mooij, T. Heskes, D. Posthuma, MAGMA: Generalized gene-set analysis of GWAS data. *PLOS Comput. Biol.* **11**, e1004219 (2015). doi: 10.1371/journal.pcbi.1004219; pmid: 25885710
96. X. Wen, R. Pique-Regi, F. Luca, Integrating molecular QTL data into genome-wide genetic association analysis: Probabilistic assessment of enrichment and colocalization. *PLOS Genet.* **13**, e1006646 (2017). doi: 10.1371/journal.pgen.1006646; pmid: 28278150
97. The GTEx Consortium, The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020). doi: 10.1126/science.aaz1776; pmid: 32913098
98. J. Scafidi *et al*., Intranasal epidermal growth factor treatment rescues neonatal brain injury. *Nature* **506**, 230–234 (2014). doi: 10.1038/nature12880; pmid: 24390343
99. Y.-C. Cheng *et al*., Genome-wide association analysis of young-onset stroke identifies a locus on chromosome 10q25 near *HABP2*. *Stroke* **47**, 307–316 (2016). doi: 10.1161/STROKEAHA.115.011328; pmid: 26732560
100. E. Hofer *et al*., White matter lesion progression: Genome-wide search for genetic influences. *Stroke* **46**, 3048–3057 (2015). doi: 10.1161/STROKEAHA.115.009252; pmid: 26451028
101. H. Mure *et al*., Akt2 and Akt3 play a pivotal role in malignant gliomas. *Neuro-Oncol.* **12**, 221–232 (2010). doi: 10.1093/neuonc/nop026; pmid: 20167810
102. D. Alcantara *et al*., Mutations of AKT3 are associated with a wide spectrum of developmental disorders including extreme megalencephaly. *Brain* **140**, 2610–2622 (2017). doi: 10.1093/brain/awx203; pmid: 28969385
103. K. R. Howell, K. Floyd, A. J. Law, PKBγ/AKT3 loss-of-function causes learning and memory deficits and deregulation of AKT/mTORC2 signaling: Relevance for schizophrenia. *PLOS ONE* **12**, e0175993 (2017). doi: 10.1371/journal.pone.0175993; pmid: 28467426
104. H. Wang *et al*., Impaired spatial learning is associated with disrupted integrity of the white matter in Akt3 knockout mice. *CNS Neurosci. Ther.* **23**, 99–102 (2017). doi: 10.1111/cns.12647; pmid: 27671373
105. Q. Wang *et al*., A Bayesian framework that integrates multi-omics data and gene networks predicts risk genes from schizophrenia GWAS data. *Nat. Neurosci.* **22**, 691–699 (2019). doi: 10.1038/s41593-019-0382-7; pmid: 30988527
106. H. K. Finucane *et al*., Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015). doi: 10.1038/ng.3404; pmid: 26414678
107. A. Kundaje *et al*., Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015). doi: 10.1038/nature14248; pmid: 25693563

Downloaded from http://science.sciencemag.org/ on June 17, 2021

108. J. F. Fullard *et al*., An atlas of chromatin accessibility in the adult human brain. *Genome Res.* **28**, 1243–1252 (2018). doi: 10.1101/gr.232488.117; pmid: 29945882
109. M. E. Hauberg *et al*., Common schizophrenia risk variants are enriched in open chromatin regions of human glutamatergic neurons. *Nat. Commun.* **11**, 5581 (2020). doi: 10.1038/s41467-020-19319-2; pmid: 33149216
110. T. H. Pers *et al*., Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6**, 5890 (2015). doi: 10.1038/ncomms6890; pmid: 25597830
111. M. Kang, Y.-S. Lee, The impact of RASopathy-associated mutations on CNS development in mice and humans. *Mol. Brain* **12**, 96 (2019). doi: 10.1186/s13041-019-0517-5; pmid: 31752929
112. L. Qu *et al*., The Ras Superfamily of Small GTPases in Non-neoplastic Cerebral Diseases. *Front. Mol. Neurosci.* **12**, 121 (2019). doi: 10.3389/fnmol.2019.00121; pmid: 31213978
113. E.-E. Govek, S. E. Newey, L. Van Aelst, The role of the Rho GTPases in neuronal development. *Genes Dev.* **19**, 1–49 (2005). doi: 10.1101/gad.1256405; pmid: 15630019
114. S. M. Smith, T. E. Nichols, Statistical challenges in "big data" human neuroimaging. *Neuron* **97**, 263–268 (2018). doi: 10.1016/j.neuron.2017.12.018; pmid: 29346749
115. A. R. Martin *et al*., Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **51**, 584–591 (2019). doi: 10.1038/s41588-019-0379-x; pmid: 30926966
116. Q. Xu *et al*., CHIMGEN: A Chinese imaging genetics cohort to enhance cross-ethnic and cross-geographic brain research. *Mol. Psychiatry* **25**, 517–529 (2020). doi: 10.1038/s41380-019-0627-6; pmid: 31827248
117. H. Gardener *et al*., Race and ethnic disparities in stroke incidence in the Northern Manhattan Study. *Stroke* **51**, 1064–1069 (2020). doi: 10.1161/STROKEAHA.119.028806; pmid: 32078475
118. C. Chen, J. M. Zissimopoulos, Racial and ethnic differences in trends in dementia prevalence and risk factors in the United States. *Alzheimers Dement.* **4**, 510–520 (2018). doi: 10.1016/j.trci.2018.08.009; pmid: 30364652
119. E. Fieremans, J. H. Jensen, J. A. Helpern, White matter characterization with diffusional kurtosis imaging. *Neuroimage* **58**, 177–188 (2011). doi: 10.1016/j.neuroimage.2011.06.006; pmid: 21699989
120. H. Zhang, T. Schneider, C. A. Wheeler-Kingshott, D. C. Alexander, NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage* **61**, 1000–1016 (2012). doi: 10.1016/j.neuroimage.2012.03.072; pmid: 22484410
121. J. H. Jensen, J. A. Helpern, A. Ramani, H. Lu, K. Kaczynski, Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magn. Reson. Med.* **53**, 1432–1440 (2005). doi: 10.1002/mrm.20508; pmid: 15906300
122. I. O. Jelescu, M. D. Budde, Design and validation of diffusion MRI models of white matter. *Front. Phys.* **28**, 61 (2017). doi: 10.3389/fphy.2017.00061; pmid: 29755979
123. J. S. W. Campbell *et al*., Promise and pitfalls of g-ratio estimation with MRI. *Neuroimage* **182**, 80–96 (2018). doi: 10.1016/j.neuroimage.2017.08.038; pmid: 28822750
124. S. Mohammadi, M. F. Callaghan, Towards in vivo g-ratio mapping using MRI: Unifying myelin and diffusion imaging. *J. Neurosci. Methods* **348**, 108990 (2021). doi: 10.1016/j.jneumeth.2020.108990; pmid: 33129894
125. S. B. Vos, D. K. Jones, M. A. Viergever, A. Leemans, Partial volume effect as a hidden covariate in DTI analyses. *Neuroimage* **55**, 1566–1576 (2011). doi: 10.1016/j.neuroimage.2011.01.048; pmid: 21262366
126. F. Alfaro-Almagro *et al*., Confound modelling in UK Biobank brain imaging. *Neuroimage* **224**, 117002 (2021). doi: 10.1016/j.neuroimage.2020.117002; pmid: 32502668
127. B. Zhao, T. Li, Analysis codes for Common genetic variation influencing human white matter microstructure, Zenodo (2021); http://doi.org/10.5281/zenodo.4549465.
128. B. Zhao, Summary statistics dataset for Common genetic variation influencing human white matter microstructure (mean), Zenodo (2021); http://doi.org/10.5281/zenodo.4549730.
129. B. Zhao, Summary statistics dataset for Common genetic variation influencing human white matter microstructure (FA PCs), Zenodo (2021); http://doi.org/10.5281/zenodo.4549725.

ACKNOWLEDGMENTS

We thank J. Lin for making the bottom right panel of the summary figure. We thank S. Cui, X. Zong, and P. Vandehaar for helpful conversations. We thank the individuals represented in the UKB, ABCD, HCP, PING, and PNC studies for their participation and the research teams for their work in collecting, processing, and disseminating these datasets for analysis. We gratefully acknowledge all of the studies and databases that made GWAS summary data available. This research has been conducted using the UK Biobank resource (application no. 22783) and is subject to a data transfer agreement. **Funding:** This research was partially supported by NIH grants MH086633 (to H.Z.), MH116527 (to T.Li), and HD079124 (to Y.L.). This work was also supported by the National Institutes of Health (RO1AG050986, R01MH109677, U01MH116442, and R01MH110921 to P.R.) and the US Department of Veterans Affairs (merit grant BX002395 to P.R.). Part of the data collection and sharing for this project was funded by the Pediatric Imaging, Neurocognition and Genetics Study (PING) (US National Institutes of Health grant RC2DA029475). PING is funded by the National Institute on Drug Abuse and the Eunice Kennedy Shriver National Institute of Child Health and Human Development. PING data are disseminated by the PING Coordinating Center at the Center for Human Development, University of California, San Diego. We thank the research team of PING, including C. McCabe, L. Chang, N. Akshoomoff, E. Newman, T. Ernst, P. Van Zijl, J. Kuperman, S. Murray, C. Bloss, M. Appelbaum, A. Gamst, W. Thompson, and H. Bartsch. Support for the collection of the PNC datasets was provided by grant RC2MH089983 awarded to R. Gur and RC2MH089924 awarded to H. Hakonarson. All PNC subjects were recruited through the Center for Applied Genomics at The Children's Hospital in Philadelphia. Part of the data used in the preparation of this article was obtained from the Adolescent Brain Cognitive Development (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, ages 9 to 10, and follow them over the course of 10 years into early adulthood. The ABCD Study is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041022, U01DA041028, U01DA041048, U01DA041089, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, U24DA041147, U01DA041093, and U01DA041025. A full list of supporters is available at https://abcdstudy.org/federal-partners/. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/scientists/workgroups/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in analysis or writing of this Research Article. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. HCP data were provided by the Human Connectome Project; by the WU-Minn Consortium (principal investigators: D. Van Essen and K. Ugurbil; 1U54MH091657), funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. **Author contributions:** B.Z., H.Z., J.L.S., and Y.L. designed the study. B.Z., T.Li, Y.Y., X.W., and T.Lu. analyzed the data. T.Li, Y.S., Z.Z., Y.Y., X.W., T.Lu., and D.X. downloaded the datasets, preprocessed dMRI data, and undertook the quantity controls. P.R., M.E.H., J.B., and J.F.F. analyzed brain cell chromatin accessibility data. B.Z. and H.Z. wrote the manuscript with feedback from all authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** We made use of publicly available software and tools. The analysis code is freely available at Zenodo (*127*). Our GWAS summary statistics are publicly available at Zenodo (*128*, *129*). Our results can also be browsed through our BIG-KP knowledge portal at https://bigkp.org/. The individual-level raw data used in this study can be obtained from five publicly accessible data resources.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6548/eabf3736/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S22
Tables S1 to S29
References (*130*–*148*)
MDAR Reproducibility Checklist

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

26 October 2020; accepted 23 April 2021
10.1126/science.abf3736

Downloaded from http://science.sciencemag.org/ on June 17, 2021

Science

Common genetic variation influencing human white matter microstructure

Bingxin Zhao, Tengfei Li, Yue Yang, Xifeng Wang, Tianyou Luo, Yue Shan, Ziliang Zhu, Di Xiong, Mads E. Hauberg, Jaroslav Bendl, John F. Fullard, Panagiotis Roussos, Yun Li, Jason L. Stein and Hongtu Zhu

Science **372** (6548), eabf3736.
DOI: 10.1126/science.abf3736

Connecting the dots on white matter

The white matter of the brain, which is composed of axonal tracts connecting different brain regions, plays key roles in both normal brain function and a variety of neurological disorders. Zhao *et al.* combined detailed magnetic resonance imaging–based assessment of brain structures with genetic data on nearly 44,000 individuals (see the Perspective by Filley). On the basis of this comprehensive analysis, the authors identified structural and genetic abnormalities associated with neurological and psychiatric disorders, as well as some nondisease traits, thus creating a valuable resource and providing some insights into the underlying neurobiology.

Science, abf3736, this issue p. eabf3736; see also abj1881, p. 1265

ARTICLE TOOLS http://science.sciencemag.org/content/372/6548/eabf3736

SUPPLEMENTARY MATERIALS http://science.sciencemag.org/content/suppl/2021/06/16/372.6548.eabf3736.DC1

RELATED CONTENT http://science.sciencemag.org/content/sci/372/6548/1265.full

REFERENCES This article cites 143 articles, 14 of which you can access for free
http://science.sciencemag.org/content/372/6548/eabf3736#BIBL

PERMISSIONS http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Science (print ISSN 0036-8075; online ISSN 1095-9203) 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.

Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works

Downloaded from http://science.sciencemag.org/ on June 17, 2021