nature genetics

Article

https://doi.org/10.1038/s41588-023-01475-y

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

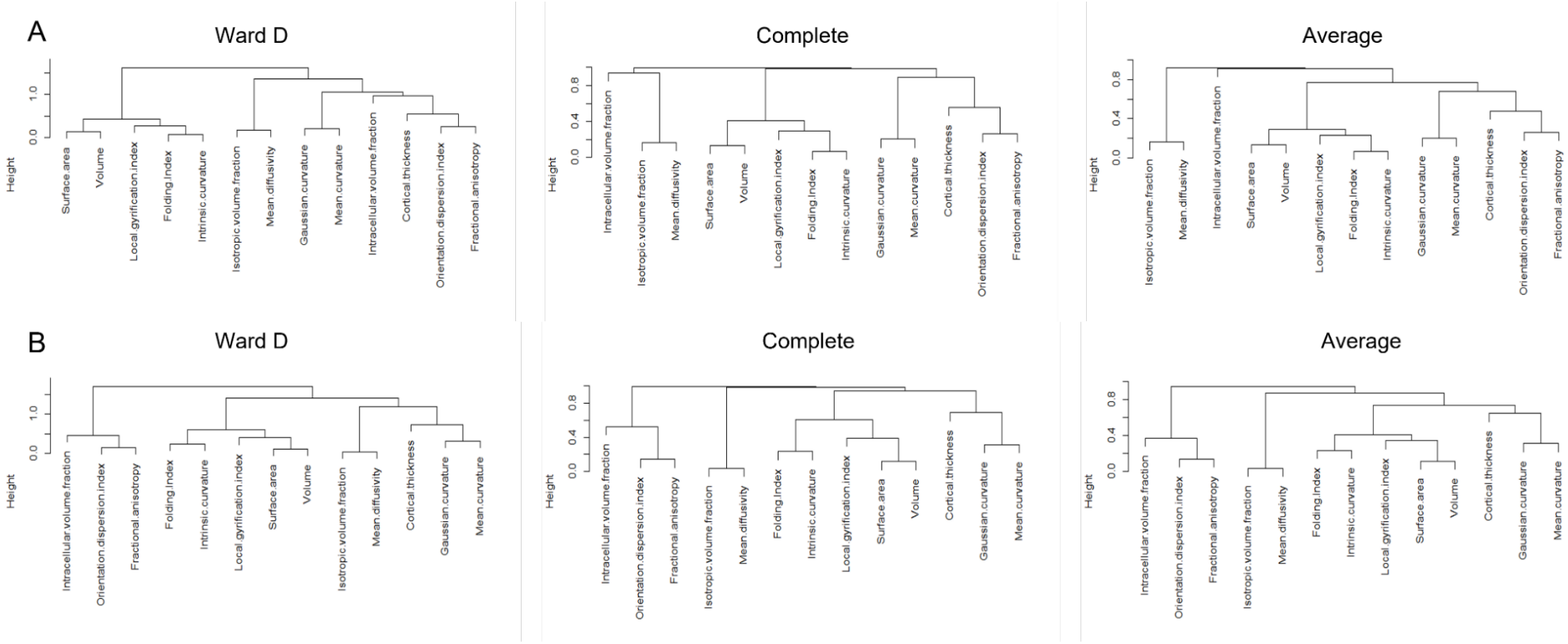
In the format provided by the authors and unedited

Supplementary information for:

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

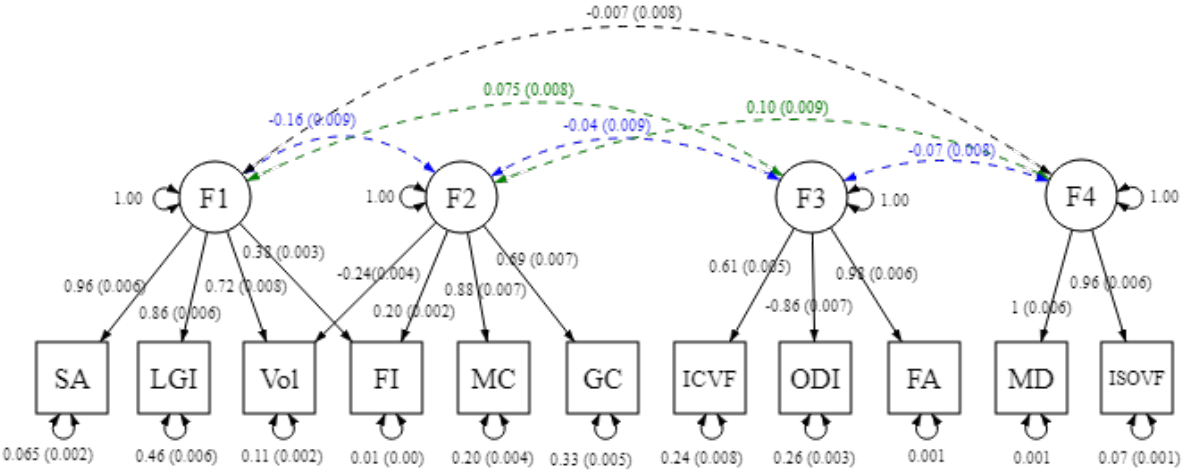
Supplementary Figures

Supplementary Figure 1: Clustering of global phenotypes using phenotypic (A) and genetic (B) correlation matrices.



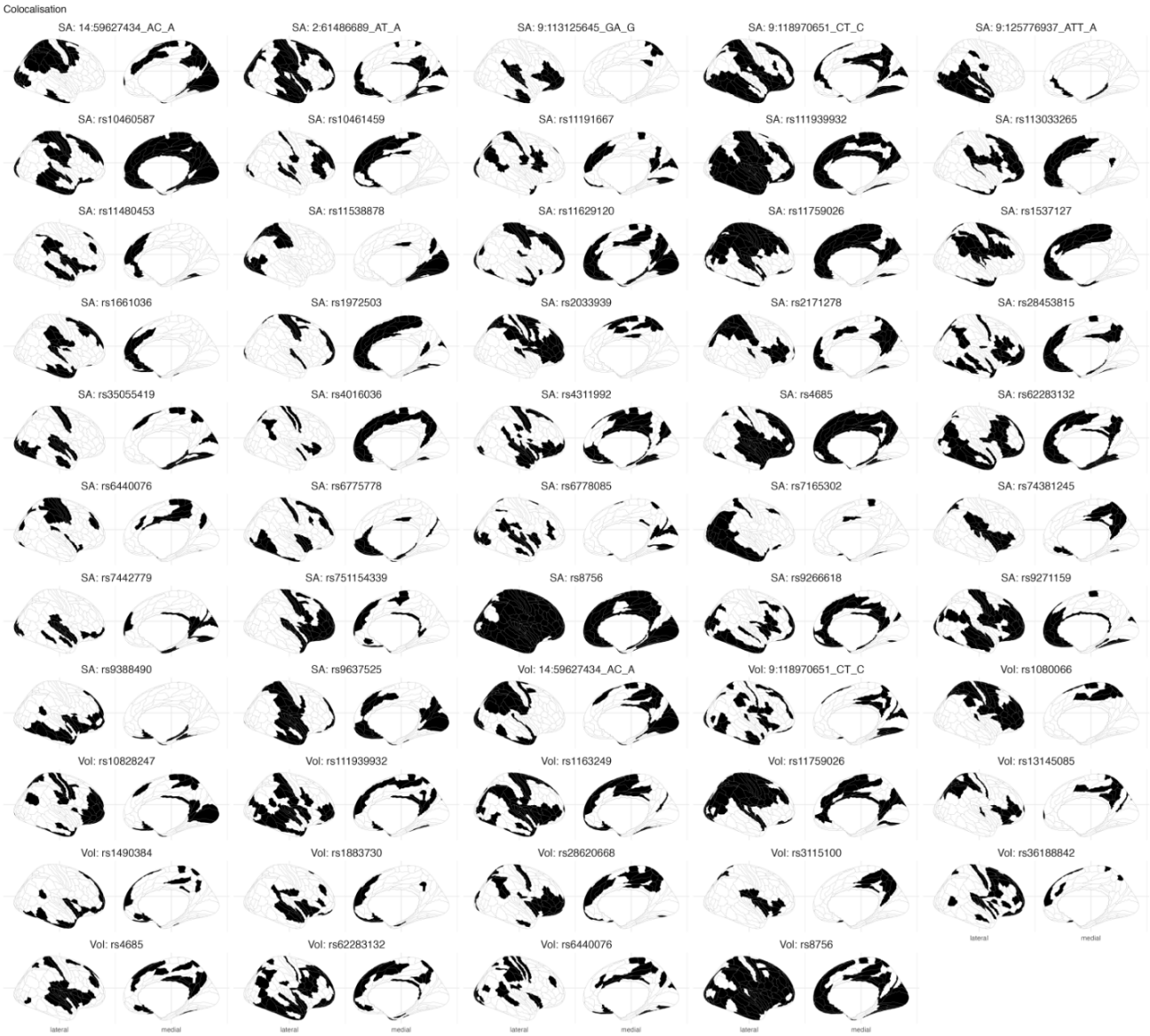
Consistency of clustering approaches was assessed using 3 methods for hierarchical clustering: clustering using Ward's criterion, complete linkage, and average linkage.

Supplementary Figure 2: Phenotypic structural equation model



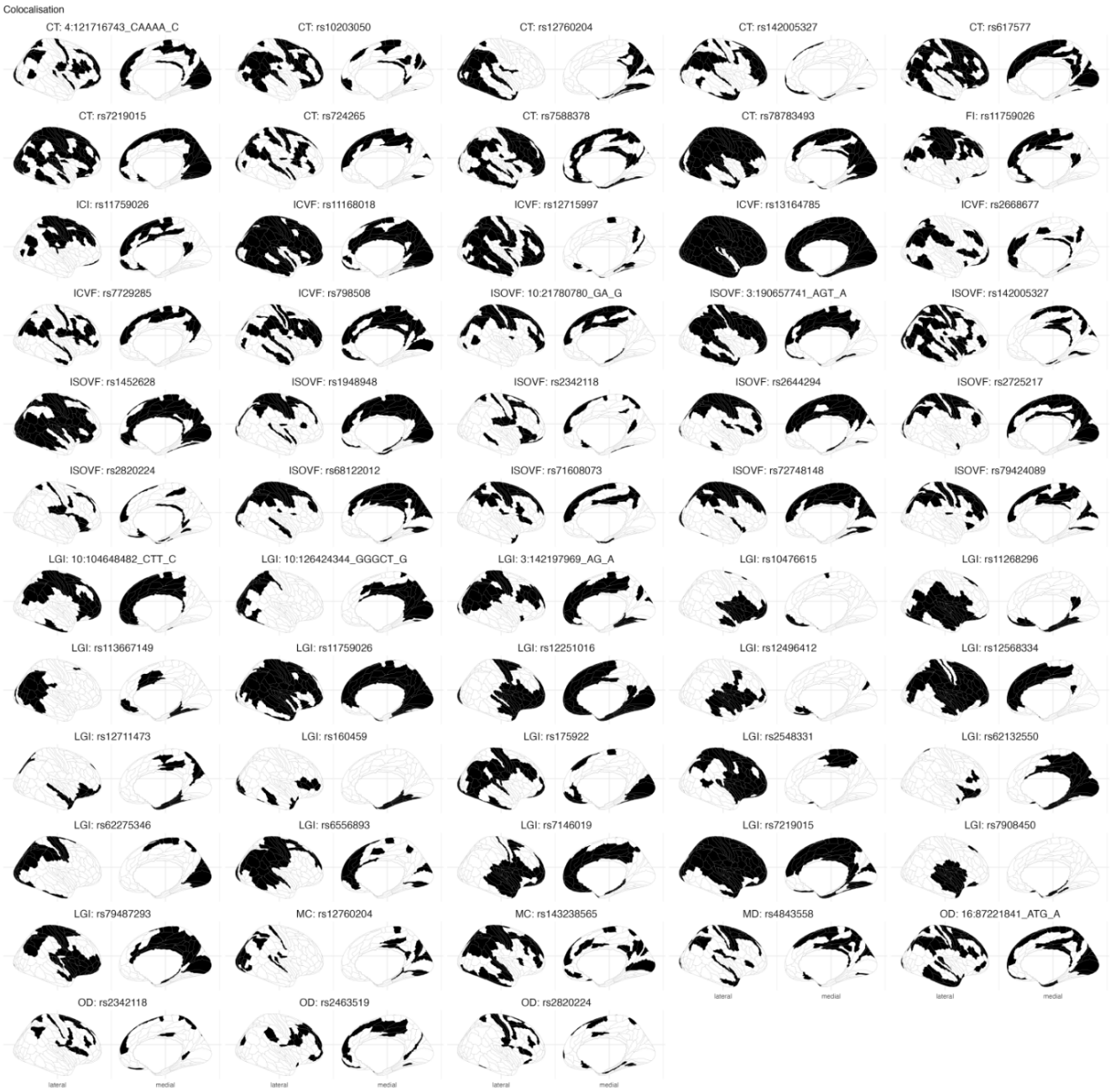
Phenotypic SEM path diagram demonstrating the underlying latent structure of 12 of the 13 global phenotypes and the interfactor genetic correlations. Covariance relationships = double-headed arrows connecting two variables, variance estimates = double-headed arrows connecting variable to itself, regression relationships = one-headed arrows pointing from independent variable to dependent variable. Circles indicate latent variables, squares indicate measured phenotypes. Abbreviations: cortical surface area (SA), grey matter volume (Vol), folding index (FI), local gyrification index (LGI), mean curvature (MC), gaussian curvature (GC), fractional anisotropy (FA), mean diffusivity (MD), intracellular volume fraction (ICVF), isotropic volume fraction (ISOVF), and orientation dispersion index (ODI).

Supplementary Figure 3: Co-localisation lots of all clusters with cluster size > 30



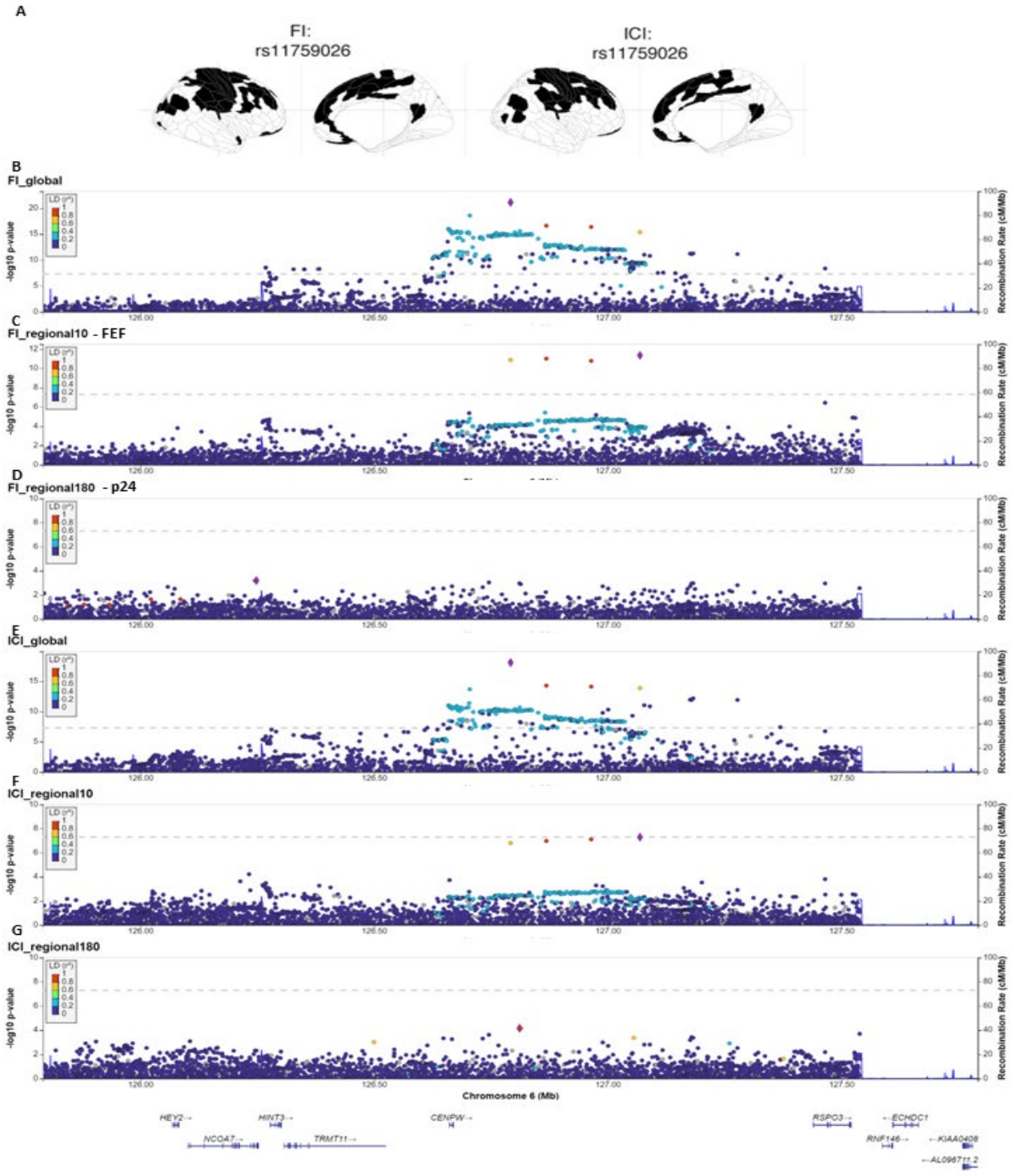
Cortical topographical plots demonstrating topographical distribution of co-localisation clusters with cluster size 30 for volume and surface area. Clusters are coloured in black. For each plot, the relevant phenotype and the SNP identified as the candidate causal variant for the cluster is provided.

Supplementary Figure 4: Co-localisation lots of all clusters with cluster size > 30



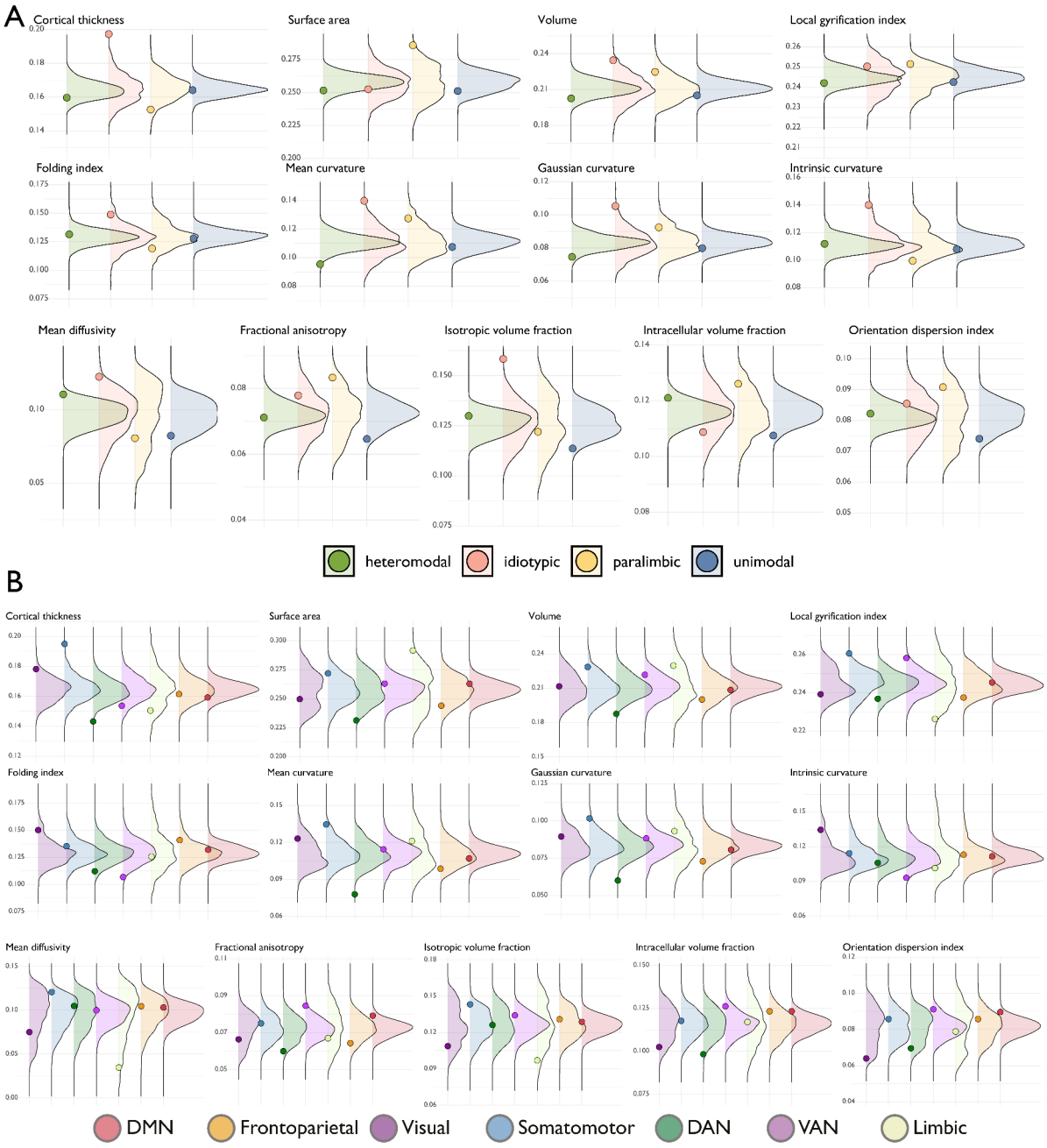
Cortical topographical plots demonstrating topographical distribution of co-localisation clusters with cluster size 30 for all phenotypes except volume and surface area. Clusters are coloured in black. For each plot, the relevant phenotype and the SNP identified as the candidate causal variant for the cluster is provided.

Supplementary Figure 5: Co-localisation and regional association plot for FI and ICI-6:125424383-127540461



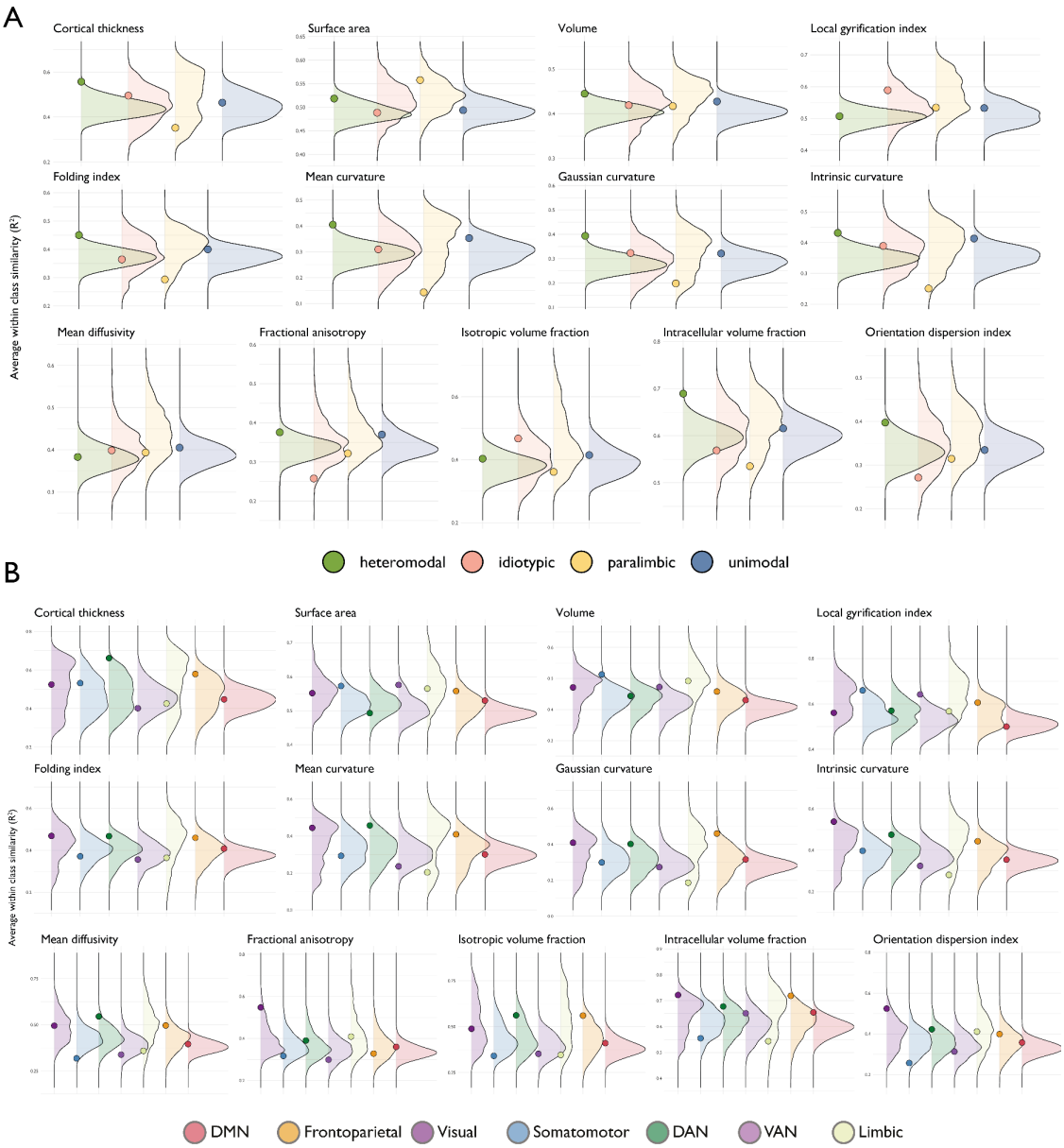
A: Cortical topographic plots of two co-localised loci with ICI and FI. B and E: Global GWAS for FI and ICI respectively. C and F regional GWAS (FEF region) that was part of the co-localised cluster in FI and ICI respectively. D and G: regional GWAS (p24 region) that was outside the co-localised cluster in FI and ICI.

Supplementary Figure 6: Spin permutation based enrichment of mean SNP heritability across mesulam classes and Yeo & Krienen communities.



A. The average SNP heritability for each Mesulam class (dots) overlaid onto a null distribution of the heritability in that class obtained by spinning the parcellation 1000 times (and computing the average heritability within each permutation). Thus, if the real average heritability (dots) are within the tail ends (5%) of the null distribution it could be concluded that this heritability is higher in that class and for that feature than would be expected from a spatially random distribution of the heritability across the cortex. Only idiotypic regions for some phenotypes (cortical thickness, volume, mean, gaussian and intrinsic curvature, mean diffusivity and isotropic volume fraction) show this relative spatial enrichment. B. The same type of spatial enrichment analysis for the Yeo and Krienen communities.

Supplementary Figure 7: Spin permutation based enrichment of mean genetic correlations within mesulam classes and Yeo & Krienen communities.



A. The average within class genetic similarity (i.e., the average of all edges in the genetic correlation matrix for regions belonging to the same class) for each Mesulam class (dots) overlaid onto a null distribution obtained by spinning the parcellation 1000 times (and computing the average similarity within each permutation). Thus, if the real genetic correlations for a given class (dots) are within the tail ends (5%) of the null distribution it could be concluded that this genetic correlation within regions belonging to the same class is higher in that class and for that feature than would be expected from a spatially random distribution across the cortex. B. The same type of spatial enrichment analysis for the Yeo and Krienen communities.

Supplementary Figure 8: Genetic-ancestry based clustering

Figure 17A: PCA plot of the first and second genetic principal components of ABCD (Self-reported ancestry) and the 1000 genomes ancestry.

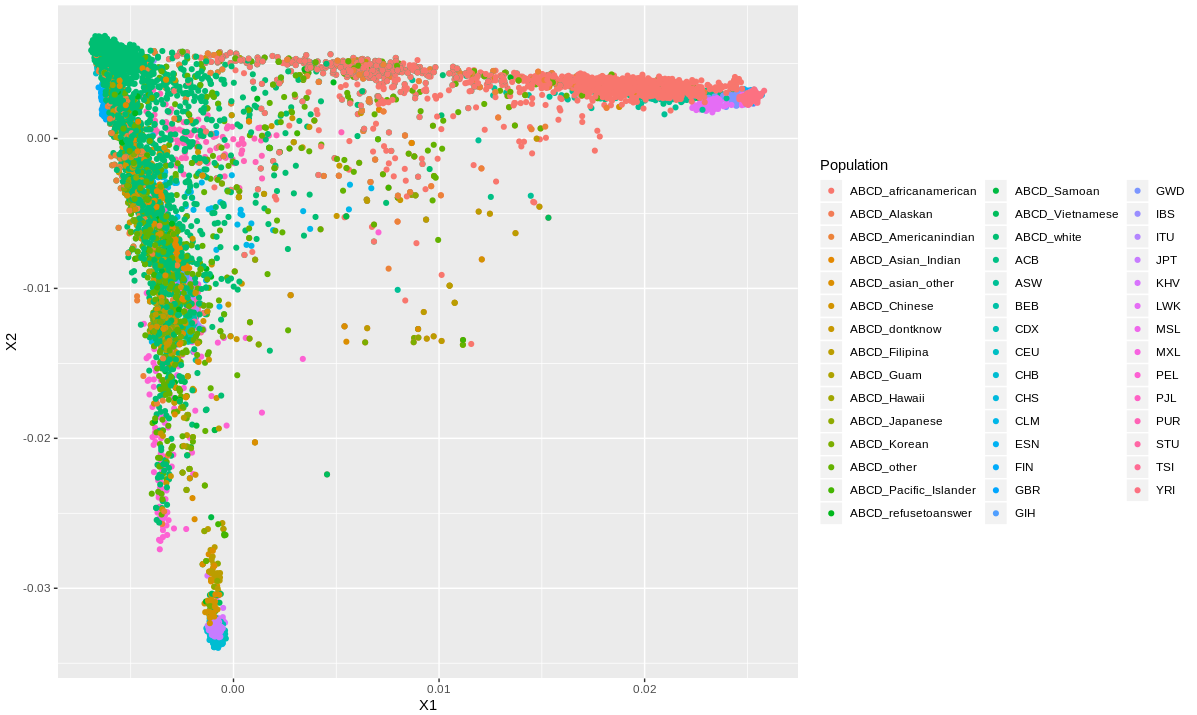
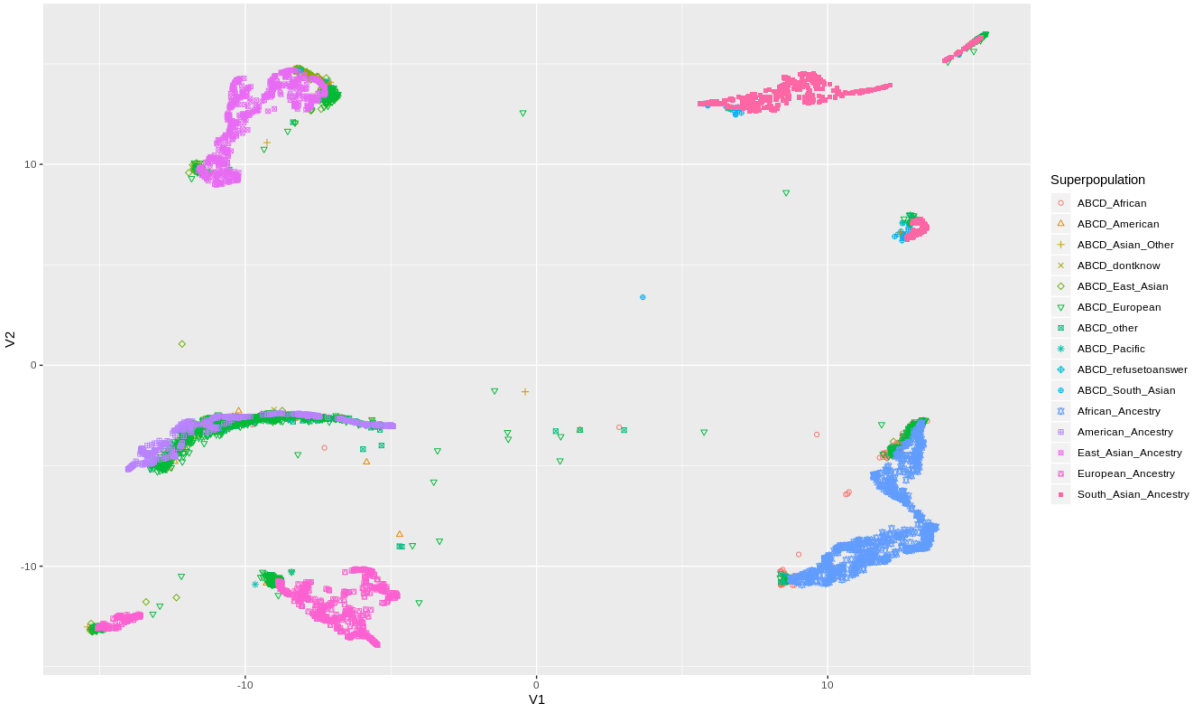


Figure 18B: UMAP of the 1000 genomes populations and ABCD using the first five genetic PCs



Supplementary notes and associated figures

Supplementary Note 1: Phenotypic factor analysis and structural equation modelling

Using the quality controlled phenotypic dataset which was included in the GWAS, we conducted exploratory factor analysis on half of the sample (N_{total} = 28,794, N_{half} = 14,397), randomly selected. We conducted confirmatory factor analysis in the other half of the sample. All phenotypes were scaled. Examination of the phenotypic correlation matrix indicated that, with the exception of CT, most phenotypes exhibited a pattern of high absolute correlation with at least one other phenotype and low correlations with others. In contrast, CT exhibited moderate correlation with almost all phenotypes, in line with the idea that multiple cell types underlie CT.

We next investigated if our data is amenable to factor analysis using Bartlett's test of sphericity Kaiser-Meyer-Olkin measure of sampling adequacy. The overall measure of sampling adequacy was 0.5, which is the minimum acceptable for factor analysis. Inspecting individual measures of sampling adequacy identified very low sampling adequacy for CT (0.29), consistent with earlier findings that CT does not cluster very well with other phenotypes. Excluding CT produced a better overall measure of sampling adequacy (0.57). Bartlett's test confidence was significant both before and after excluding CT.

Given the above points, we excluded CT from factor analysis and conducted exploratory factor analysis on 12 of the 13 global phenotypes. Scree plot and parallel analysis indicated four factors, in line with the number of clusters identified from hierarchical clustering. Exploratory factor analysis identified four factors: 1: Cortical expansion; 2: Cortical curvature; 3: Neurite density; and 4: Water diffusion (**SF 3**). There was significant and high (> 0.5) cross-loading of ICI onto two factors (cortical expansion and cortical curvature). Multiple iterations of the confirmatory factor analyses when including ICI failed to produce satisfactory models. After removing ICI, we identified a similar four factor model relating to cortical expansion, cortical curvature, neurite density and water diffusion with acceptable fit indices (CFI: 0.85, SRMR: 0.79).

Supplementary Note 2: Identifying causal factors for folding using Mendelian Randomisation

An overview of MR and the methods used

Understanding the processes that underlie cortical folding has proven elusive. Several theories have tried to explain folding including skull constraint, axonal tension, and differential tangential growth[1]. Skull constraint and axonal tension do not account for the uniform nature of folding across among humans. The cellular model for differential tangential expansion posits that folding occurs through two mechanisms: first, the outer layers expand more than the inner layers, causing the cortex to fold and second, there is heterogeneous distribution of progenitors across the brains, leading to differences in neurogenesis[1–3]. In humans, one way to test this is to investigate if genetic variants underlying surface area increase measures of curvature. We used multiple Mendelian Randomisation (MR) methods to investigate this.

MR analyses require three key assumptions: (1) the genetic variant is associated with the exposure; (2) the genetic variant is not associated with confounders; and (3) the genetic variant will only influence the outcome through the exposure.

To address the first rule, we use only genome-wide significant SNPs, or SNPs that are near genome-wide significant. However, addressing the second two rules are challenging. With phenotypes such as cortical macrostructure, it is theoretically impossible to say if that the genetic variant is not associated with a confounder or that the genetic variant influences the outcome solely through the exposure. This is because we cannot completely rule out pleiotropy. Therefore, to strengthen the evidence substantially, we use a suite of sensitivity analyses to ensure that the results are robust to some violations of the MR assumptions. First we use the gold-standard inverse-variance weighted Mendelian Randomisation to identify significant potential causal effects. As this is prone to bias if genetic variants do not meet the three rules above (i.e., invalid instrument), we use median-weighted MR which provides valid estimates if a large number of (up to 50%) the genetic variants are invalid. We then use two additional methods to account for or remove SNPs that may be pleiotropic - MR-egger and MR-PRESSO. Consistent results across all four indicates that our findings are reasonably robust to invalid instruments due to pleiotropy.

However, none of these methods account for reverse causality as, due to difference in statistical power between the GWAS for the exposure and outcomes, some genetic variants that are significant in the exposure may actually explain greater variance in the outcome. We use Steiger filtering to identify these SNPs and remove them. We conduct all the above analyses again, to ensure that the results are robust to reverse causal instruments. This is particularly important in our analyses due to the high pleiotropy and shared genetics between the cortical expansion phenotypes.

Next, we use a relatively new method (CAUSE) to model pleiotropy as correlated (i.e., mediated by a confounder) and uncorrelated pleiotropy. Whilst uncorrelated pleiotropy adds noise to the estimate, correlated pleiotropy can lead to bias and false positives. MR-Egger and MR-PRESSO rely on the assumption that all pleiotropy is uncorrelated. CAUSE models both correlated and uncorrelated pleiotropy. Significant results in CAUSE suggest that the causal estimates are robust to even correlated pleiotropy.

In addition to all of these methods, to ensure that our results are replicable across datasets, we conduct a validation by using two datasets. In the first analyses we split the UKB into two random approximate halves and conduct MR between the two datasets. Next, we use summary statistics from UKB as the exposure and summary statistics from ABCD as outcome.

An overview of the results

We split the UK Biobank into two datasets of roughly equal sample size to ensure that there is limited participant overlap and conducted MR. Inverse-variance weighted MR demonstrated that genetically predicted surface area increased genetically predicted local gyrification index (LGI), intrinsic curvature index (ICI), and folding index (FI) after Bonferroni correction. These results were statistically significant after applying methods that are robust to various assumptions: median-weighted[4] MR and MR-Presso[5], and after removing outliers by Steiger filtering[6]. These results also had consistent effect direction when using MR-Egger[7], which is statistically underpowered compared to other methods. However, in the reverse direction, we did not obtain consistent results to indicate that genetic variants associated with these measures increased genetically predicted surface area (**SF 4**). Visual inspection of the forest plots and leave-one-out plots did not indicate that the results were driven by one or two genetic variants (**SF 5**).

We confirmed our findings first by using a different method that can model both correlated and uncorrelated pleiotropy (CAUSE)[8] using two sets of instruments - one created from SNPs with $p < 5x10^{-8}$, and another with SNPs with $p < 1x10^{-3}$, as previously demonstrated by the developers of the method. CAUSE suggested that SA causally increased LGI, FI and ICI. In the reverse direction, CAUSE identified a shared model between LGI, FI, and SA, and a causal model between SA and ICI. Consistent results suggesting that genetically predicted surface area increased genetically predicted measures of folding were obtained when MR was run using ABCD and UKB. Altogether, MR provides support for the differential tangential growth theory for some measures of folding, but suggests that other mechanisms may contribute to other measures of folding.

Supplementary Note 3: Adjusting for global phenotypes can bias regional GWAS

Following Aschard and colleagues[9], we consider four scenarios in which a genetic variant may be associated with regional phenotypes. Acyclic graphs are provided below (**SF 13**).

Scenario 1: A genetic variant is associated with a global phenotype but not associated with a regional phenotype.

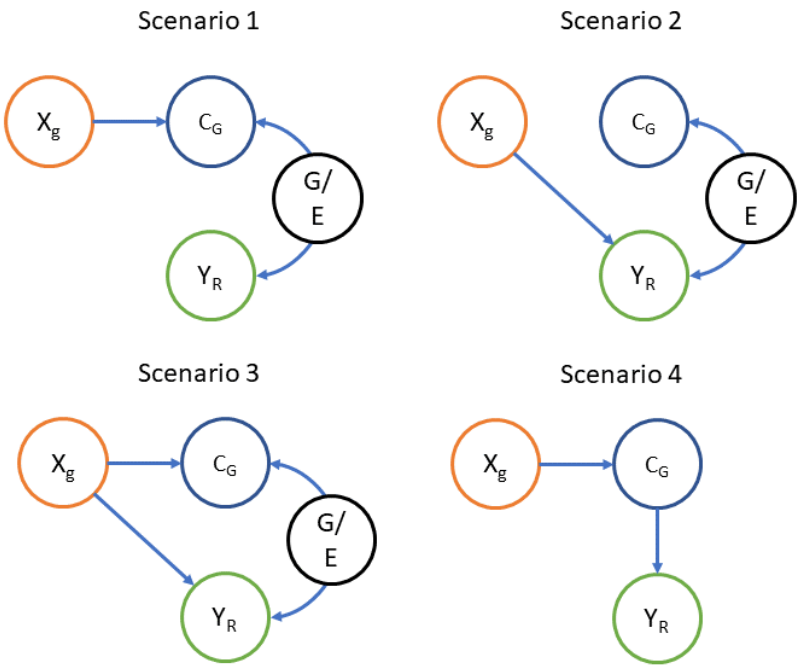
Scenario 2: A genetic variant is associated with a regional phenotype but not associated with global phenotypes.

Scenario 3: A genetic variant is independently associated with both regional and global phenotypes.

Scenario 4: The genetic effect of a variant on a regional phenotype is mediated partly or completely by global phenotypes.

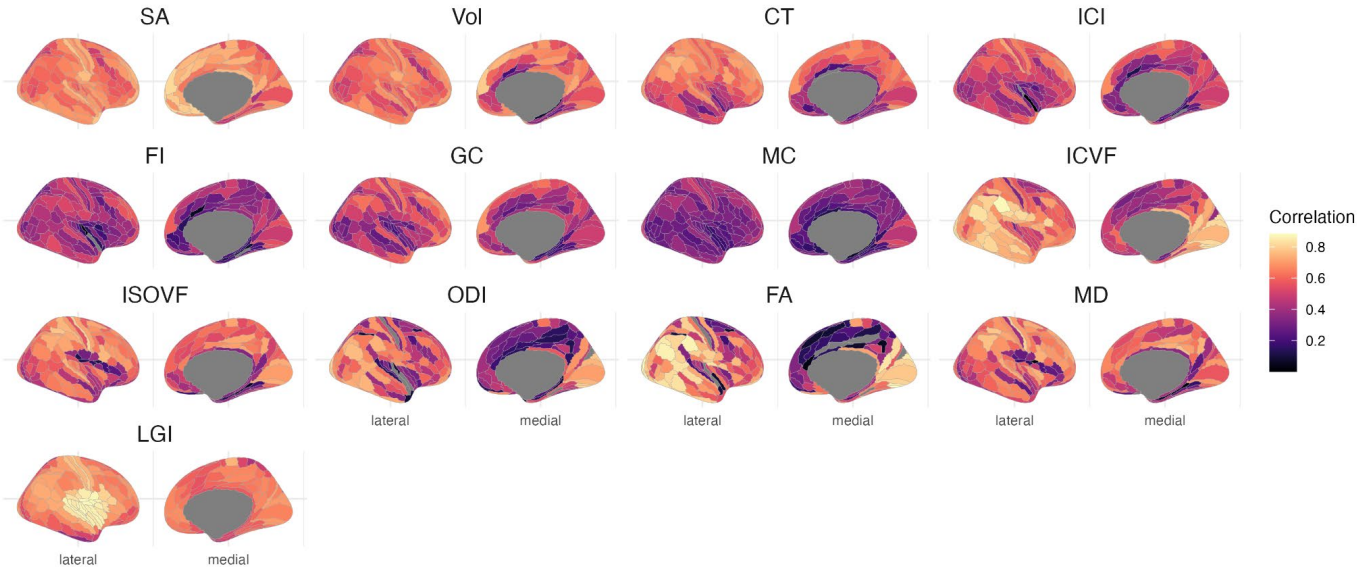
For Scenarios 2 and 4, controlling for the global phenotype will not affect the association between the genetic variant and the regional phenotype. However, for scenarios 1 and 3, correcting for the global phenotype will induce a correlation between the genetic variant and the regional phenotype. The estimate will be biased by $\beta_C\rho_{CY}$ where β_C is the effect of the genetic variant on the covariate (in this case the global phenotype), and ρ_{CY} is the correlation between the covariate (the global phenotype) and outcome. In our study, the β_C for the standardised global phenotypes range from (-1.47) to (1.16) for the genome-wide significant loci, and the correlations between the global phenotypes and regional phenotypes are high and displayed in (**SF 14**). Therefore, correcting for the global phenotypes will bias the regional estimates.

Supplementary Note 3, Figure 1: Acyclic graphs for the association between genetic variant and regional phenotypes



Acyclic graphs demonstrating causal relationship between genetic variant (X_g) and regional phenotype (Y_R), in the presence of the global phenotype which is the covariate (C_G) and other genetic and environmental factors (G/E) which contribute to the phenotypic correlation between the regional phenotype and the global phenotype.

Supplementary Note 3, Figure 2: Correlation between regional phenotypes and global phenotypes

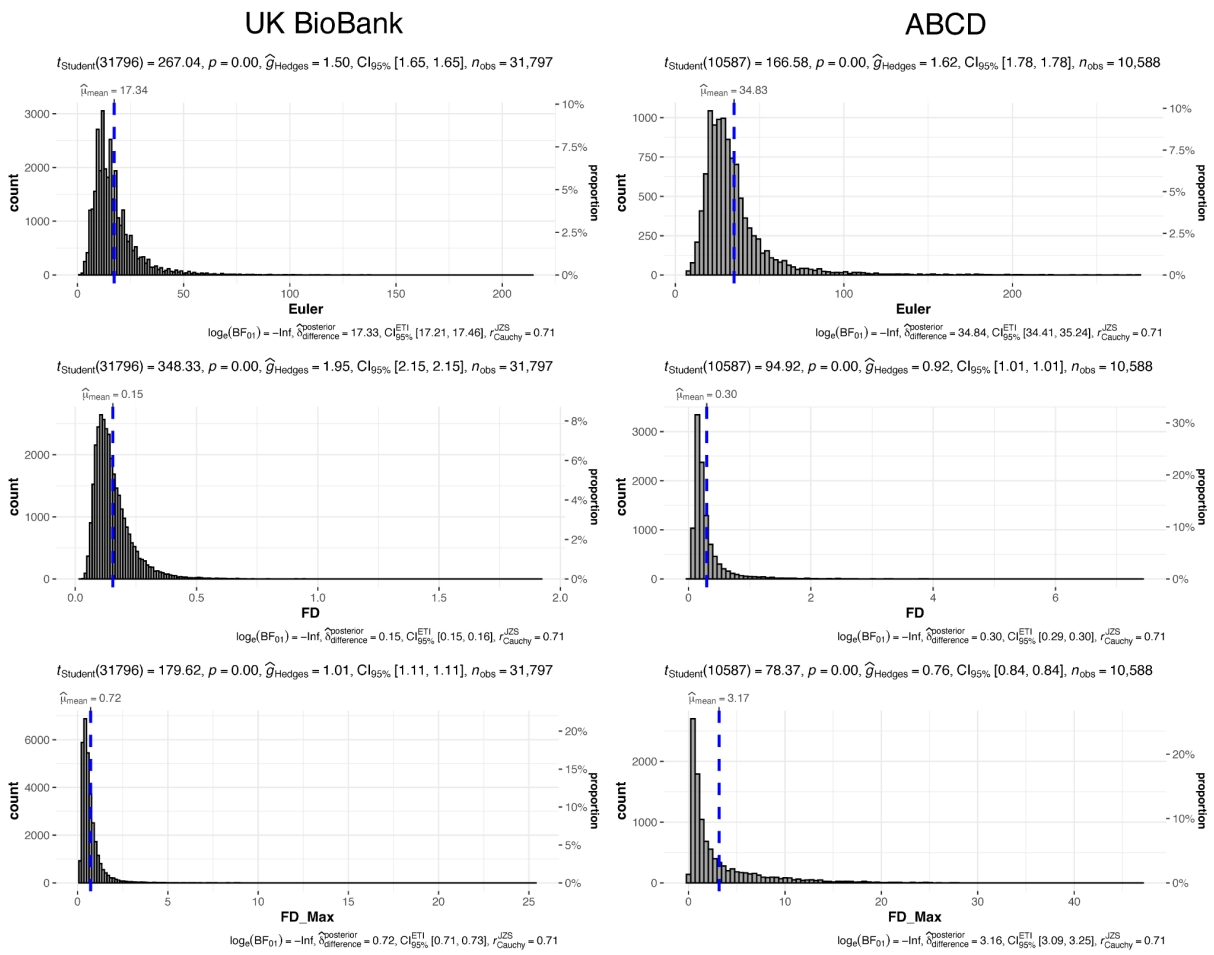


Pearson's correlation coefficient between regional phenotypes and global phenotypes.

Supplementary Note 4: Impact of processing pipeline and image quality

4a: Image quality

Motion and image reconstruction quality could potentially impact data quality and especially paediatric cohorts might be more susceptible to move more in the scanner. We note that both the UK BioBank and the ABCD datasets are of generally high-quality, i.e., less than 0.2% of individuals have an Euler index above the threshold of 216 used by Rosen et al. 2016 [10]. Similarly we do indeed find that the ABCD younger dataset exhibits relatively more motion than the older UK BioBank cohort, but analogous to the Euler index less than 0.1% of the sample showed an average framewise displacement above 2mm. Extreme head motion would likely result in poor image quality which will have been reflected in either unusable scans (which are generally not uploaded to the respective repositories or marked as unusable). In addition scans with high Euler indices or high levels of motion would likely show extreme deviations in any of the included phenotypes. To ensure these would not be included in any downstream analyses individuals were removed from analyses if they were more than 5 median absolute deviations removed from the group median on any given phenotype. Furthermore, an additional GWAS of all global phenotypes in the UK BioBank cohorts with individuals who were more than 5 median absolute deviations away from the group median on either the Euler Index, average framewise displacement or maximum framewise excluded yielded highly similar genome wide associations ($r_g > 0.90$ for all).



Histogram plots generated using gghistostat [11]. Includes computed Bayes Factors to quantify the likelihood of each variable being significantly different from 0 (BF10) and the null hypothesis (BF01), as well a one-sample, two-sided student's T-test to indicate the same. In our current example, the Bayes Factor value and T-test provide very strong evidence that all scores are above zero. Plots also indicate a clear leftward skew towards lower scores (indicating higher quality.

4b: Strategies for parcellation alignment

There is considerable variability in processing and parcellation alignment strategies in literature. While it is beyond the scope of the present work to provide an exhaustive evaluation of all possible approaches to the whole processing pipeline and how they might impact downstream genetic associations, we did evaluate the impact of direct surface based registration versus indirect surface based registration as these may impact spatial accuracy of assigned cortical regions [12]. Specifically, to evaluate robustness of the regional diffusion-weighted imaging phenotypes, we compared surface-based and volume-based approaches for individual subject registration. The 'bbregister' command in FreeSurfer was used to affine transform the diffusion-weighted images to the T1-weighted MRI for each subject using the transformation files from the original processing method. In contrast to the volumetric approach, where the brain parcellation was aligned to the diffusion-weighted image before averaging across voxels (using the native resolution of the diffusion-weighted image), the surface approach aligned the diffusion-weighted image to the T1-weighted MRI before projection to the surface. Surface projection was performed using the 'myelin-style' method in the 'volume-to-surface-

mapping' command in Connectome Workbench and using the computed surfaces from the FreeSurfer pipeline. Regional values were then derived by averaging across vertices using the original output parcellation files from FreeSurfer.

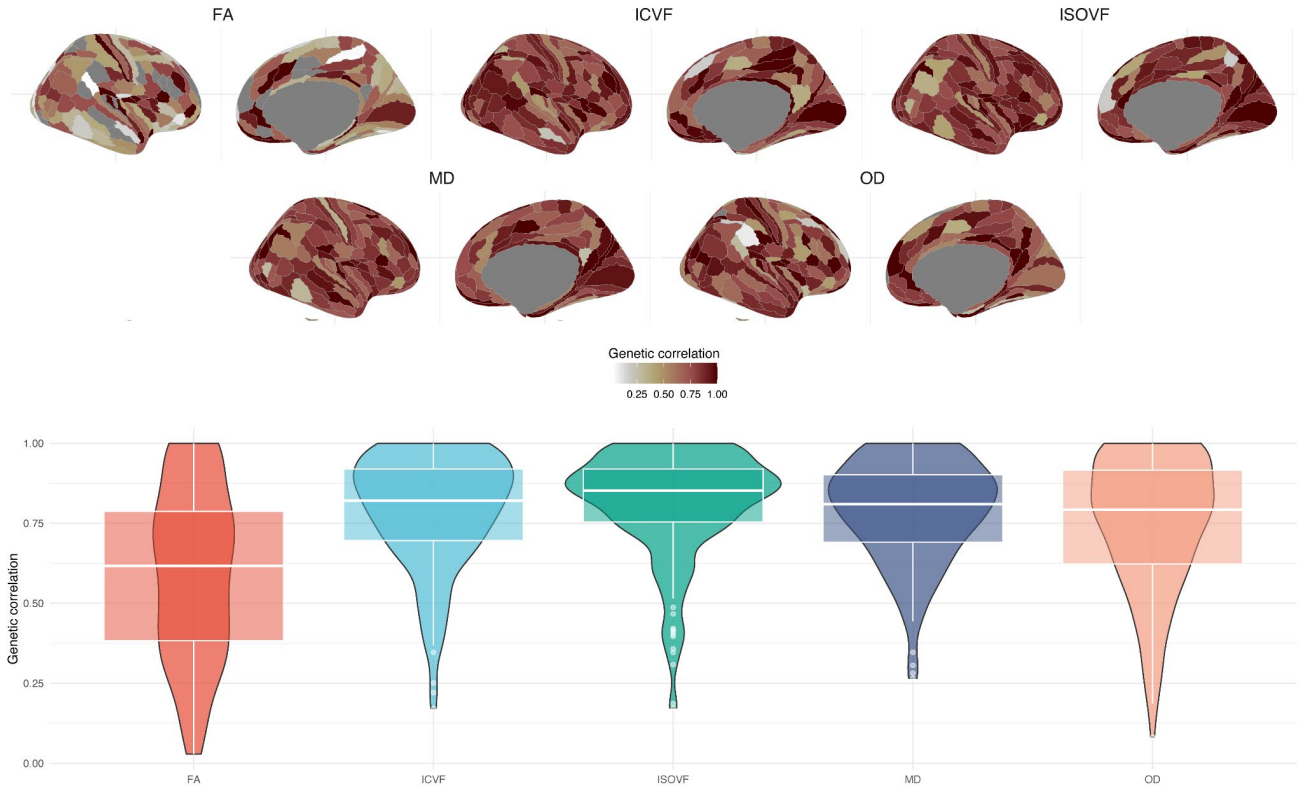
Subsequently, for the five diffusion-weighted phenotypes, we calculated genetic correlations between the indirect surface based registration and the direct surface based registration in approximately 9,650 individuals from the UK Biobank using GCTA-GREML after including the same covariates as in the GWAS. We observed very high genetic correlations for the global diffusion phenotypes (**Supplementary Note 4b Table 1**). We also observed relatively high median genetic correlations for the regional phenotypes, although, particularly for FA, the genetic correlations had very high standard errors due to the low heritability of

Supplementary Note 4b Table 1: Genetic correlations for the global phenotypes.

	Genetic correlation	SE	N
FA	0.829	0.21	9661
MD	0.903	0.03	9649
NODDI ICVF	0.937	0.05	9654
NODDI ISOVF	0.893	0.03	9664
NODDI OD	0.966	0.08	9663

Genetic correlations were calculated between the indirect surface based registration and the direct surface based registration for 9,650 individuals using GCTA-GREML.

Supplementary Note 4, Figure 2: Genetic correlation of parcellation approaches.



Genetic correlation between the direct and indirect surface-based registration for the five diffusion MRI phenotypes. N_{max} = 36,663.

Additional Acknowledgements

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

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

References

1. Garcia, K. E., Kroenke, C. D. & Bayly, P. V. Mechanics of cortical folding: stress, growth and stability. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **373**, (2018).
2. Kriegstein, A., Noctor, S. & Martínez-Cerdeño, V. Patterns of neural stem and progenitor cell division may underlie evolutionary cortical expansion. *Nat. Rev. Neurosci.* **7**, 883–890 (2006).
3. Reillo, I., de Juan Romero, C., García-Cabezas, M. Á. & Borrell, V. A role for intermediate radial glia in the tangential expansion of the mammalian cerebral cortex. *Cereb. Cortex* **21**, 1674–1694 (2011).
4. Bowden, J., Smith, G. D., Haycock, P. C. & Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genetic Epidemiology* vol. 40 304–314 Preprint at https://doi.org/10.1002/gepi.21965 (2016).
5. Verbanck, M., Chen, C.-Y., Neale, B. & Do, R. Publisher Correction: Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* **50**, 1196 (2018).
6. Hemani, G., Tilling, K. & Smith, G. D. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLOS Genetics* vol. 13 e1007081 Preprint at https://doi.org/10.1371/journal.pgen.1007081 (2017).
7. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**, 512–525 (2015).
8. Morrison, J., Knoblauch, N., Marcus, J. H., Stephens, M. & He, X. Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using

genome-wide summary statistics. *Nat. Genet.* **52**, 740–747 (2020).

9. Aschard, H., Vilhjálmsson, B. J., Joshi, A. D., Price, A. L. & Kraft, P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *Am. J. Hum. Genet.* **96**, 329–339 (2015).
10. Rosen, A. F. G. *et al.* Quantitative assessment of structural image quality. *Neuroimage* **169**, 407–418 (2018).
11. Patil, I. Visualizations with statistical details: The 'ggstatsplot' approach. *J. Open Source Softw.* **6**, 3167 (2021).
12. Coalson, T. S., Van Essen, D. C. & Glasser, M. F. The impact of traditional neuroimaging methods on the spatial localization of cortical areas. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E6356–E6365 (2018).