



OPEN

# The genetic architecture of structural left-right asymmetry of the human brain

Zhiqiang Sha<sup>1</sup>, Dick Schijven<sup>1</sup>, Amaia Carrion-Castillo<sup>1</sup>, Marc Joliot<sup>2</sup>, Bernard Mazoyer<sup>2</sup>, Simon E. Fisher<sup>1,3</sup>, Fabrice Crivello<sup>2</sup> and Clyde Francks<sup>1,3</sup>✉

**Left-right hemispheric asymmetry is an important aspect of healthy brain organization for many functions including language, and it can be altered in cognitive and psychiatric disorders. No mechanism has yet been identified for establishing the human brain's left-right axis. We performed multivariate genome-wide association scanning of cortical regional surface area and thickness asymmetries, and subcortical volume asymmetries, using data from 32,256 participants from the UK Biobank. There were 21 significant loci associated with different aspects of brain asymmetry, with functional enrichment involving microtubule-related genes and embryonic brain expression. These findings are consistent with a known role of the cytoskeleton in left-right axis determination in other organs of invertebrates and frogs. Genetic variants associated with brain asymmetry overlapped with those associated with autism, educational attainment and schizophrenia. Comparably large datasets will likely be required in future studies, to replicate and further clarify the associations of microtubule-related genes with variation in brain asymmetry, behavioural and psychiatric traits.**

The human brain is characterized by various population-level asymmetries on its left–right axis<sup>1</sup>, including an overall ‘torque’ whereby the left hemisphere extends posteriorly and ventrally relative to the right, a left–right difference in frontal–occipital gradients of cortical thickness<sup>2</sup>, and hemispheric differences of morphology around the Sylvian fissure<sup>3</sup>. Many brain functions are also lateralized, including hand motor control and language which show left-hemisphere dominance in roughly 85% of people<sup>4–13</sup>. Altered brain or behavioural asymmetries have been reported in various cognitive and psychiatric disorders<sup>7,14–17</sup>, which suggests that population-typical asymmetries are linked to optimal human brain function.

Behavioural and anatomical brain asymmetries are already apparent in utero<sup>1,18–20</sup>, which indicates an early genetic-developmental programme of brain left–right axis formation<sup>21,22</sup>. Studies of visceral organ development (the heart, stomach, liver, etc.) have revealed that the generation of population-level asymmetry requires at least three important steps in the early embryo<sup>23,24</sup>: (1) the breaking of bilateral symmetry to create a left–right axis in a consistent orientation relative to the anterior–posterior and dorsal–ventral axes, (2) the triggering of different patterns of gene expression on the left and right sides of early embryonic structures and (3) the translation of asymmetric gene expression into lateralized morphology and organ placement.

In principle, establishing an embryonic left–right axis requires chirality at some level, that is, key biomolecules or cellular structures that exist in only one of two possible mirror forms. Life on Earth is based on L-form amino acids rather than the mirror D-form, and this chirality carries through to the macrostructure and movement of primary cilia<sup>25,26</sup>, which help to create the left–right axis of the visceral organs in embryos<sup>25</sup>. However, hemispheric dominances for language and hand motor control do not typically reverse in people with situs inversus of the viscera, that is, reversal of the visceral organs on the left–right axis, when caused by mutations

in genes that encode primary ciliary components<sup>27–30</sup>. This observation indicates that there are distinct and possibly organ-intrinsic mechanisms at play in brain development, but such mechanisms remain unidentified. Therefore the genetic origins of human brain asymmetry remain unknown.

Post mortem studies that contrasted gene expression between the left and right sides of the embryonic central nervous system have identified molecular pathways that may be involved, including that the two sides may transition through developmental stages slightly out of synchrony with each other<sup>31–33</sup>. However, these studies were necessarily based on data from just handfuls of samples, because of limited availability arising from appropriate ethical, legal and practical concerns. An alternative approach to identify genetic influences on brain asymmetry is to relate genomic variation in large population datasets to variation in adult brain asymmetry. Only three loci have previously been reported at a genome-wide significant level to affect variation in adult human brain asymmetries, in studies targeted at single features of temporal lobe anatomy, and they did not yield broader insights into biological pathways<sup>34,35</sup>. Here we made use of the unprecedented sample size available through the 2020 release of the UK Biobank magnetic resonance imaging (MRI) data, in combination with genome-wide genotype data, to perform the first multivariate, brain-wide genetic analysis of human brain anatomical asymmetry. The results from this analysis also allowed us to test whether genetic polymorphisms that are associated with variation in brain asymmetry are also associated with neurodevelopmental disorders or other behavioural and psychological traits, using publicly available genome-wide association scan (GWAS) summary statistics for these traits.

## Results

**Heritabilities and genetic correlations of brain regional asymmetry measures.** For each of 32,256 participants with post-quality-control MRI and genetic data (Methods), and each of 73 bilaterally

<sup>1</sup>Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands. <sup>2</sup>Groupe d’Imagerie Neurofonctionnelle, Institut des Maladies Neurodégénératives, Centre National de la Recherche Scientifique, Commissariat à l’Energie Atomique, et Université de Bordeaux, Bordeaux, France. <sup>3</sup>Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands. ✉e-mail: [Clyde.Francks@mpi.nl](mailto:Clyde.Francks@mpi.nl)

paired regional measures of brain structure, we calculated hemispheric asymmetry indexes (AI) as  $(\text{left} - \text{right}) / ((\text{left} + \text{right}) / 2)$  (for 33 cortical surface area AIs, 33 cortical thickness AIs and 7 subcortical volume AIs; Supplementary Table 1). The measures were derived from cortical parcellation and subcortical segmentation of structural brain images (Methods). All but one of the regional mean AIs were significantly different from zero, indicating population-level asymmetries (Bonferroni-corrected  $P < 0.05$ , Supplementary Fig. 1 and Supplementary Table 2), consistent with previous reports<sup>5,6</sup>. For example, some language-related regions showed greater average left than right surface areas, including superior temporal and supramarginal cortex, and pars opercularis.

GCTA<sup>36</sup> software was used to estimate the single-nucleotide polymorphism (SNP)-based heritability ( $h^2$ ) for each AI, that is, the extent to which variance in each AI was linked to common genetic variation over the entire genome (Methods). Forty-two AIs showed significant SNP-based heritabilities (false discovery rate (FDR)-corrected  $P < 0.05$ ), that is, 28 of the surface area AIs, 8 cortical thickness AIs and 6 subcortical volume AIs (Fig. 1a and Supplementary Table 3), ranging from 2.2% for the AI of entorhinal cortical thickness to 9.4% for the AI of superior temporal surface area. The overall pattern was consistent with previous, twin-based heritability analyses<sup>5,6</sup>.

SNP-based genetic correlation analysis (again using GCTA software) indicated overlapping genetic contributions to some of the AIs (Fig. 1b, Supplementary Fig. 2 and Supplementary Tables 4–10). Within a small number of cortical regions (Supplementary Table 10), surface area and thickness AIs had negative genetic correlations, which indicates that variants can have antagonistic effects on surface and thickness asymmetries of these regions.

**Multivariate genome-wide association analysis.** We performed a multivariate (mv)GWAS for 9,803,522 SNPs, using meta-canonical correlation analysis as implemented in MetaPhat<sup>37</sup>, with the 42 AIs that had significant SNP-based heritability. This analysis tested each SNP separately for its simultaneous associations with all 42 AIs (Methods). A multivariate approach had the dual advantages of achieving data reduction and increasing statistical power compared with running 42 separate univariate GWAS. FUMA<sup>38</sup> software was used to clump mvGWAS results on the basis of linkage disequilibrium (LD), and identify lead SNPs (maximally associated SNPs) at each associated locus. There were 21 distinct genomic loci at the  $5 \times 10^{-8}$  significance level associated with different aspects of brain asymmetry (Fig. 2, Table 1 and Supplementary Fig. 3), represented by 27 independent lead SNPs (with pairwise LD  $r^2 < 0.1$ , Table 1).

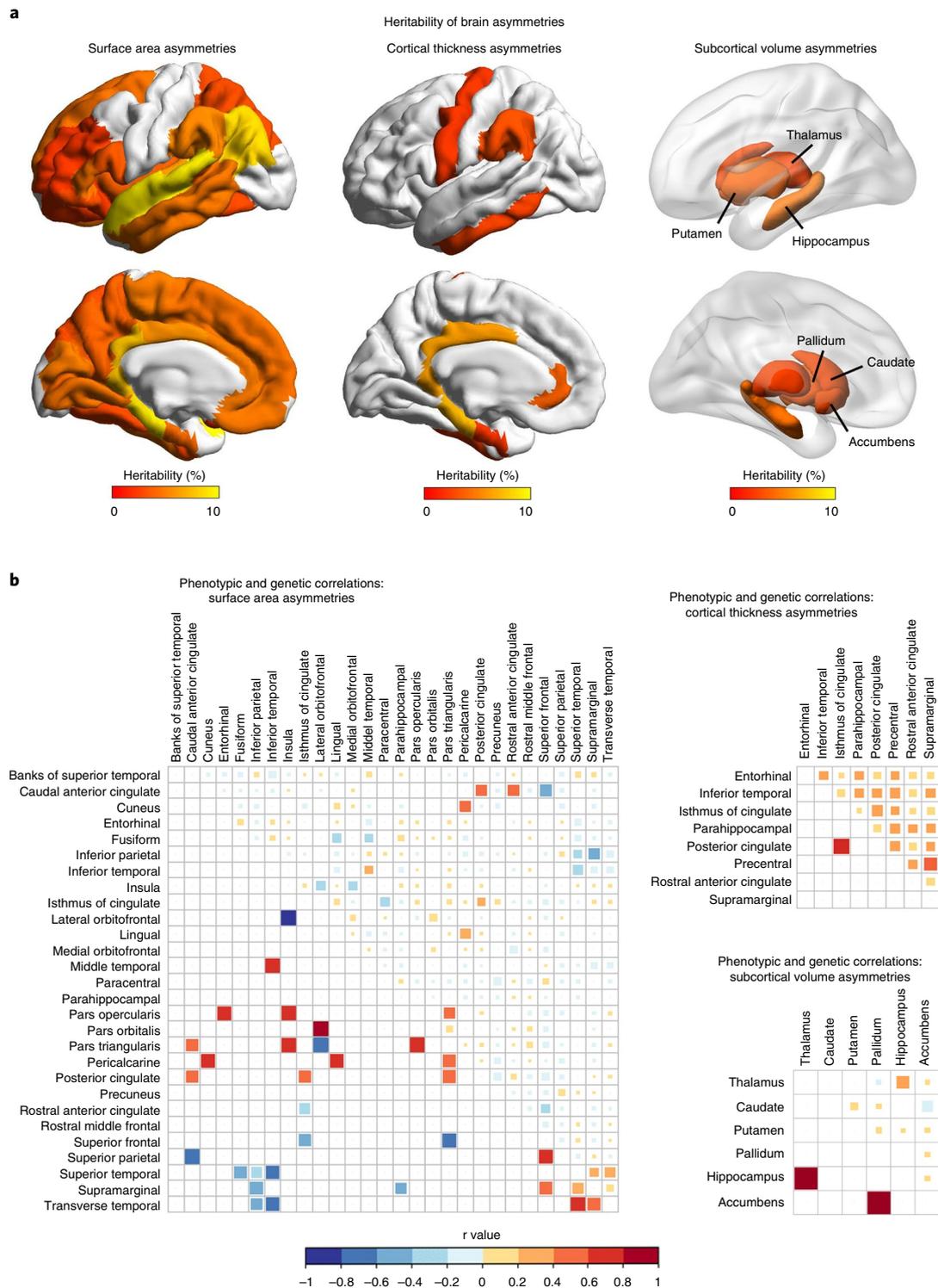
For each lead SNP, phenotype decomposition<sup>37</sup> identified the ‘central’ AIs that contributed to its multivariate association (Supplementary Table 11). Most central AIs associated with the 27 lead SNPs were distributed in core regions of the language system (for example, lateral temporal, pars opercularis and supramarginal) or limbic system (for example, cingulate, orbitofrontal and mesial temporal cortex; Fig. 3). For example, the most significant SNP, with multivariate  $r = 0.094$ ,  $P = 4.75 \times 10^{-38}$  (rs41298373 on 10p14) had five central AIs: the minor allele was associated with a leftward shift of surface area asymmetry for two lateral temporal regions, a rightward shift of surface area asymmetry for two medial temporal regions and a leftward shift of cortical thickness asymmetry in the inferior temporal gyrus (Supplementary Table 11). All lead SNPs were associated with at least one cortical regional surface area AI as one of their central traits. Thirteen lead SNPs were associated with cortical thickness AIs, and five lead SNPs were associated with subcortical volume AIs. Two lead SNPs, rs35853889 and rs6658111, were associated with the AIs of both surface area and cortical thickness within the same region, that is, parahippocampal area and thickness AIs were associated with both of these SNPs, and rostral anterior cingulate area and thickness AIs were associated with rs35853889.

In addition, five lead SNPs affecting cortical surface area AIs were also associated with AIs of subcortical volumes (Table 1 and Fig. 3), and the locus on 17q21 was associated with asymmetries of cortical surface area, thickness and subcortical volume (Table 1). The univariate associations of lead variants separately with left and right hemispheric measures are presented in Supplementary Table 11.

**Functional annotations of genomic loci associated with brain asymmetry.** FUMA<sup>38</sup> software applied three strategies to annotate candidate SNPs to genes at significantly associated loci (Methods): physical position, expression quantitative trait locus (eQTL) information and chromatin interactions (Supplementary Table 12 and Supplementary Figs. 4 and 5). Here we summarize and cite notable information on each of the lead SNPs:

Ten of the loci had annotations involving cytoskeleton-related genes: On 1p33, rs6658111 is close to AL356458.1, a pseudogene of *MTMR14* (myotubularin related protein 14). On 2p23.3, rs62130503 is intronic to *MAPRE3* (microtubule associated protein RP/EB family member 3a), and rs12617392 is a brain eQTL<sup>39</sup> of *MAPRE3*. Located between these two SNPs on 2p23.3 is also *AGBL5*, which is a post-translational modifier of tubulin<sup>40,41</sup>. On 2q34, rs368536282 is close to *MAP2* (microtubule associated protein 2), a well-known dendrite-specific marker of neurons<sup>42</sup>, previously implicated in left-handedness by GWAS analysis<sup>43–45</sup>. On 6p21.33, rs7781 is in the 3′ untranslated region (UTR) of *TUBB* (tubulin beta class I). On 12q13.12, rs10783306 is close to the alpha tubulin gene *TUBA1B*. This variant is also in high LD with a handedness-associated variant, rs11168884 ( $r^2 = 0.89$ )<sup>44</sup>. On 14q23.1, two lead variants for two independent genomic loci, rs160459 and rs201816193, showed evidence for cross-locus chromatin interaction via the promoters of nearby genes in foetal cortex<sup>46</sup> (Supplementary Fig. 4). The former is near to *DACT1*, a locus which has been reported to associate with superior temporal sulcus depth<sup>34</sup>, while the latter is close to *DAAMI*, which modulates the reorganization of the actin cytoskeleton and the stabilization of microtubules<sup>47,48</sup>. Two lead variants on 16q24.3, rs72813426 and rs11139892, are in introns of *SPIRE2* and the tubulin gene *TUBB3*, respectively, both of which are key proteins in cytoskeleton organization<sup>49,50</sup>. rs11139892 is also in high LD with a handedness-associated variant, rs4550447 ( $r^2 = 0.94$ )<sup>44</sup>. On 17q21.31 there were five independent lead SNPs: rs35908989 is intronic to *MAPT* which encodes microtubule-associated protein tau, and rs55938136, rs35853889 and rs568039055 are brain eQTLs<sup>39,51,52</sup> of *MAPT*, while rs80103986 is in high LD ( $r^2 = 0.91$ ) with handedness-associated variant rs55974014<sup>44</sup>. On 19p13.3, rs11672092 is intronic to the tubulin gene *TUBB4A*, and in high LD with rs66479618 ( $r^2 = 0.88$ ), another handedness-associated SNP<sup>44</sup>.

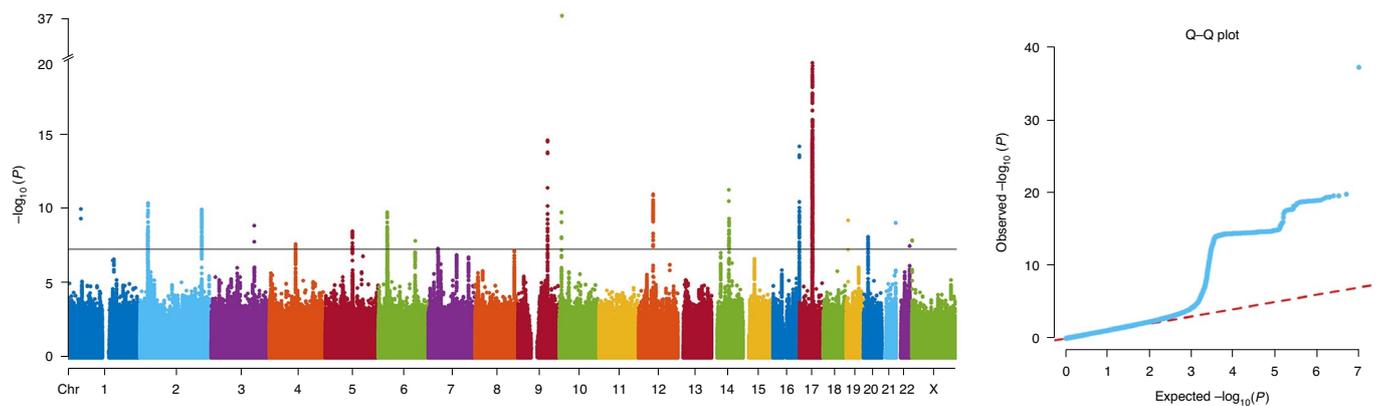
The 11 other loci did not have obvious microtubule-related annotations, but most had annotations related to brain phenotypes or development: On 3q24, rs2279829 is a cortical eQTL<sup>39</sup> of *ZIC4*, which is involved in visual and auditory pathway development<sup>53</sup>. rs9307052 on 4q22.1 is in high LD ( $r^2 = 0.99$ ) with the handedness-associated variant rs28658282<sup>44</sup>. On 5q15, rs869219775 is close to *NR2F1*, which is involved in neural activity during cortical patterning<sup>54</sup>. On 6q22.31–q22.32, rs9385385 is close to *NCOA7*, a nuclear receptor co-activator with its most abundant expression in the brain<sup>55</sup>. On 7p14.3, rs6947352 is intronic to *BBS9*, which causes Bardet–Biedl syndrome when mutated, involving retinopathy and intellectual disability<sup>56,57</sup>. On 9q22.33, rs911934 is located in a region having a chromatin interaction with *TRIM14* in adult cortex<sup>46</sup> (Supplementary Fig. 4), a gene which may activate Wnt/ $\beta$ -catenin signalling and affects mesodermal versus ectodermal differentiation of embryonic stem cells<sup>58</sup>. On 10p14, rs41298373 is a predicted deleterious missense coding variant in *ITIH5*, which was previously reported to affect planum temporale volumetric asymmetry<sup>35</sup>. Inter-alpha-trypsin inhibitor proteins are involved in extracellular matrix stabilization<sup>59</sup>. On 20p12.1, rs6135555 is in a region having a



**Fig. 1 | SNP-based heritability and correlation analysis of regional brain asymmetry measures. a**, SNP-based heritability estimates for brain asymmetry measures. Only regions for which AIs were significantly heritable are indicated in colour. **b**, Genetic and phenotypic correlations between AIs. Phenotypic (upper right triangle) and genetic (lower left triangle) correlations between each pair of AIs. Only significantly heritable AIs that also have at least one significant phenotypic or genetic correlation after FDR correction are shown. The colours of the squares indicate the correlation coefficients according to the colour key, and their areas are proportional to the correlation coefficients.

chromatin interaction with the *FLRT3* promoter in neural progenitor cells<sup>60</sup> (Supplementary Fig. 4), a gene which regulates axon guidance and excitatory synapse development<sup>61</sup>. On 21q22.3, *rs7283026* is intronic to *COL18A1*, involved in neural tube closure and mutated in Knobloch syndrome<sup>62</sup>, which can include skull abnor-

malities. On 22q13.31, *rs9615351* is an exonic variant of a gene involved in planar cell polarity, *CELSR1*<sup>63</sup>. On Xp22.33, *rs12400461* is close to pseudogene *ASS1P4* and upstream of *MXRA5*; the latter encodes a matrix remodelling-associated protein and is implicated in autism<sup>64</sup>.



**Fig. 2 | Multivariate GWAS analysis of regional brain asymmetries in 32,256 participants.** Manhattan plot for multivariate GWAS across asymmetries of surface area, cortical thickness and subcortical volumes. The red dashed line indicates the significance threshold  $P < 5 \times 10^{-8}$  (Methods). The Q-Q plot is also shown.

To further link asymmetry-associated SNPs to genes, the MAGMA<sup>65</sup> software was used to perform genome-wide gene-based association analysis<sup>65</sup> based on the results from mvGWAS. There were 57 significant genes at Bonferroni-corrected  $P < 0.05$  (Supplementary Fig. 6 and Supplementary Table 13). Five of these genes were previously associated with handedness<sup>44</sup>: *MAP2*, *FAM13A*, *TUBA1B*, *TUBB3* and *CRHR1*. Forty-three of the 57 genes have been reported to associate with educational attainment<sup>66</sup> and 15 with intelligence<sup>67</sup> (Supplementary Table 14). For the proteins encoded by the 57 genes, there were 80 known or putative pairwise interactions in the STRING database<sup>68</sup>, compared with 8 interactions expected for a random set of this size from the whole proteome ( $P < 1 \times 10^{-16}$ ). This observation supports the validity of our mvGWAS association findings, as random noise would not lead to such functional clustering. Microtubule-related genes (for example, *MAP2*, *MAPT*, *SPIRE2* and *TUBA1A*) linked different clusters together in the largest protein interaction network (Fig. 4a).

We also used the genome-wide, gene-based  $P$  values for functional enrichment analysis using MAGMA<sup>65</sup>, in relation to 7,343 Gene Ontology ‘biological process’ sets defined within the MSigDB<sup>69</sup> database. The gene set ‘regulation\_of\_microtubule\_binding’ ( $P = 3.73 \times 10^{-6}$ ) showed significant enrichment (adjusted  $P < 0.05$ , Bonferroni correction, Supplementary Table 15). Significant enrichment within various microtubule-related sets, such as ‘microtubule\_cytoskeleton\_organization’ ( $P = 2.19 \times 10^{-7}$ ) and ‘microtubule\_based\_process’ ( $P = 2.36 \times 10^{-6}$ ), was also found when using the list of single closest genes (Table 1) to the 27 lead SNPs (Supplementary Table 16). Enrichment in microtubule-related sets was not reported in a recent GWAS of bilaterally averaged cortical surface area and thickness measures in 51,665 individuals<sup>70</sup>, suggesting a particular involvement in hemispheric asymmetry rather than bilateral measures. We observed no statistically significant relation of our gene-based association  $P$  values with differential expression across cell types (Methods).

Testing our genome-wide, gene-based  $P$  values with respect to human gene expression data from the BrainSpan<sup>71</sup> database, from either 29 age groups or 11 defined developmental stages, we found relatively higher mRNA expression of brain-asymmetry-associated genes during early-prenatal ( $P = 4.27 \times 10^{-3}$ ) and early-mid-prenatal ( $P = 9.37 \times 10^{-4}$ ) stages, from 9 ( $P = 1.84 \times 10^{-3}$ ) to 24 ( $P = 7.36 \times 10^{-3}$ ) post-conceptional weeks (FDR-corrected  $P$  values  $< 0.05$ ) (Fig. 4b,c and Supplementary Table 17). This is consistent with the fact that various anatomical asymmetries of the brain are already visible in utero<sup>18,19</sup>, and supports the existence of an early developmental mechanism for establishing the brain’s left–right axis<sup>31,32,72</sup>.

**Genetic overlap of brain asymmetry with other traits.** We next used iSECA software<sup>73</sup> to perform genetic overlap analysis with our mvGWAS results in relation to GWAS summary statistics from neurodevelopmental disorders, behavioural and psychological traits which have been reported to associate phenotypically with aspects of structural and/or functional brain asymmetry: attention deficit hyperactivity disorder<sup>16,74–77</sup>, autism spectrum disorder<sup>15,78–82</sup>, educational attainment<sup>66,83,84</sup>, handedness<sup>2,4,45</sup>, intelligence<sup>67,85–88</sup> and schizophrenia<sup>17,89–93</sup>. There was evidence for genetic overlap between brain asymmetries and autism ( $P = 0.005$ ), educational attainment ( $P = 0.001$ ) and schizophrenia ( $P = 0.002$ ) which remained significant at Bonferroni-corrected  $P < 0.05$  (Fig. 5, Supplementary Figs. 7 and 8 and Supplementary Table 18). In other words, SNPs that showed lower (more significant) association  $P$  values in our mvGWAS for brain asymmetry showed a statistically significant tendency to also show lower  $P$  values in previous, large-scale GWAS of autism, educational attainment and schizophrenia. Although we did not observe genetic overlap of brain asymmetry with handedness at a genome-wide level, we did note individual loci in common between these traits (above). In addition, we found no overlap between our mvGWAS results and those from a previous GWAS of intracranial volume in 32,438 participants<sup>94</sup> (Supplementary Table 18, Supplementary Figs. 7 and 8), which again indicates that the genetic architecture of brain asymmetry is largely distinct from brain size.

**Validation of lead SNPs associated with brain asymmetry.** To achieve a reasonable level of genetic homogeneity in our mvGWAS, we had excluded any individuals not annotated as having ‘white British ancestry’ (through a combination of self-report and clustering based on principal components that capture major axes of ancestral diversity from the genotype data<sup>95</sup> (Methods)). The UK Biobank includes additional participants who self-identify as being ‘white’, but who did not self-identify as British, or did not cluster genetically with the bulk of the ‘white British’ ancestry participants (Methods). After applying the same quality control criteria to these additional participants as in our discovery mvGWAS (except with respect to ancestry), and imposing the extra criterion that relatedness coefficients should not be greater than 0.0442 with any participants from the discovery mvGWAS, data were available from 3,600 participants for an independent replication set. We tested each of the 27 lead SNPs from the mvGWAS in the replication set, using the same approach as the mvGWAS analysis, except that 40 genetic principal components were used as covariates to control for the greater degree of ancestral diversity in the replication set.

Ten of the 27 independent lead SNPs from the discovery mvGWAS showed association  $P$  values  $< 0.05$  on multivariate testing in

**Table 1 | Genomic loci associated with brain asymmetries on multivariate analysis. All lead SNPs are shown**

Genomic locus	Lead SNP	Position	Functional category	Effect allele	Effect allele frequency	mvGWAS <i>r</i> value	mvGWAS <i>P</i> value	Nearest gene	Central asymmetry indexes <sup>a</sup>
1	<a href="#">rs6658111</a>	1p33	Intergenic	G	0.37	0.0631	$9.75 \times 10^{-11}$	<a href="#">AL356458.1</a>	Parahippocampal (SA), superior frontal (SA), parahippocampal (CT)
2	<a href="#">rs62130503</a>	2p23.3	NcRNA_intronic <sup>b</sup>	T	0.05	0.0630	$1.22 \times 10^{-10}$	<a href="#">MAPRE3</a>	Thalamus (SUB), parahippocampal (SA)
2	<a href="#">rs12617392</a>	2p23.3	Intronic	A	0.44	0.0638	$4.02 \times 10^{-11}$	<a href="#">CGREF1</a>	Inferior temporal (SA), caudal anterior cingulate (SA), isthmus of cingulate (SA)
3	<a href="#">rs368536282<sup>c</sup></a>	2q34	Intergenic	T	0.03	0.0631	$1.07 \times 10^{-10}$	<a href="#">MAP2</a>	Superior frontal (SA), accumbens (SUB), posterior cingulate (CT)
4	<a href="#">rs2279829</a>	3q24	3' UTR <sup>d</sup>	T	0.22	0.0613	$1.26 \times 10^{-9}$	<a href="#">ZIC4</a>	Isthmus of cingulate (CT), precuneus (SA), posterior cingulate (CT), fusiform (SA)
5	<a href="#">rs9307052<sup>c</sup></a>	4q22.1	Intronic	T	0.11	0.0591	$2.27 \times 10^{-8}$	<a href="#">FAM13A</a>	Rostral anterior cingulate (CT), posterior cingulate (CT), medial orbitofrontal (SA)
6	<a href="#">rs869219775</a>	5q15	Intergenic	T	0.14	0.0606	$3.06 \times 10^{-9}$	<a href="#">NR2F1</a>	Inferior parietal (SA), transverse temporal (SA)
7	<a href="#">rs7781</a>	6p21.33	Downstream	G	0.24	0.0628	$1.62 \times 10^{-10}$	<a href="#">TUBB</a>	Isthmus of cingulate (CT), rostral anterior cingulate (CT), pars triangularis (SA)
8	<a href="#">rs9385385</a>	6q22.31	ncRNA_intronic <sup>b</sup>	T	0.45	0.0595	$1.37 \times 10^{-8}$	<a href="#">NCOA7</a>	Posterior cingulate (CT), pericalcarine (SA)
9	<a href="#">rs6947352</a>	7p14.3	Intronic	A	0.31	0.0585	$4.38 \times 10^{-8}$	<a href="#">BBS9</a>	Banks of the superior temporal sulcus (SA)
10	<a href="#">rs911934</a>	9q22.33	Intergenic	G	0.70	0.0699	$2.39 \times 10^{-15}$	<a href="#">GALNT12</a>	Inferior parietal (SA), isthmus of cingulate (SA), precuneus (SA), paracentral (SA), supramarginal (SA), entorhinal (CT)
11	<a href="#">rs41298373</a>	10p14	Exonic	A	0.10	0.0940	$4.75 \times 10^{-38}$	<a href="#">ITIH5</a>	Superior temporal (SA), parahippocampal (SA), fusiform (SA), inferior temporal (CT), transverse temporal (SA)
12	<a href="#">rs10783306<sup>c</sup></a>	12q13.12	Intergenic	C	0.33	0.0647	$9.99 \times 10^{-12}$	<a href="#">TUBA1B</a>	Superior frontal (SA), entorhinal (SA), medial orbitofrontal (SA), pars triangularis (SA)
13	<a href="#">rs160459</a>	14q23.1	Intergenic	C	0.46	0.0652	$4.98 \times 10^{-12}$	<a href="#">DACT1</a>	Banks of the superior temporal sulcus (SA), transverse temporal (SA), pericalcarine (SA)
14	<a href="#">rs201816193</a>	14q23.1	Intergenic	G	0.12	0.0621	$4.38 \times 10^{-10}$	<a href="#">DAAM1</a>	Isthmus of cingulate (SA), cuneus (SA)
15	<a href="#">rs72813426</a>	16q24.3	Intronic	G	0.24	0.0685	$2.45 \times 10^{-14}$	<a href="#">SPIRE2</a>	Paracentral (SA), isthmus of cingulate (SA), middle temporal (SA)
15	<a href="#">rs111398992<sup>c</sup></a>	16q24.3	Intronic	T	0.13	0.0694	$5.99 \times 10^{-15}$	<a href="#">TUBB3</a>	Isthmus of cingulate (CT), fusiform (SA), rostral anterior cingulate (CT), pericalcarine (SA)

Continued

**Table 1 | Genomic loci associated with brain asymmetries on multivariate analysis. All lead SNPs are shown (Continued)**

Genomic locus	Lead SNP	Position	Functional category	Effect allele	Effect allele frequency	mvGWAS <i>r</i> value	mvGWAS <i>P</i> value	Nearest gene	Central asymmetry indexes <sup>a</sup>
16	rs55938136	17q21.31	NcRNA_intronic <sup>b</sup>	G	0.22	0.0695	4.91 × 10 <sup>-15</sup>	<i>CRHR1</i>	Parahippocampal (SA), middle temporal (SA), pallidum (SUB), hippocampus (SUB), pars triangularis (SA)
16	rs35908989	17q21.31	Intronic	C	0.23	0.0595	1.34 × 10 <sup>-8</sup>	<i>MAPT</i>	Supramarginal (SA), caudate (SUB)
16	rs35853889	17q21.31	3' UTR <sup>d</sup>	TG	0.19	0.0765	1.43 × 10 <sup>-20</sup>	<i>MAPT</i>	Rostral anterior cingulate (CT), cuneus (SA), isthmus of cingulate (SA), parahippocampal (SA), rostral anterior cingulate (SA), parahippocampal (CT)
16	rs80103986 <sup>c</sup>	17q21.31	Intronic	T	0.20	0.0708	5.16 × 10 <sup>-16</sup>	<i>KANSL1</i>	Parahippocampal (SA), middle temporal (SA), pallidum (SUB), hippocampus (SUB), pars triangularis (SA)
16	rs568039055	17q21.31	3' UTR <sup>d</sup>	C	0.20	0.0692	7.87 × 10 <sup>-15</sup>	<i>LRRC37A2</i> <i>ARL17A</i>	Parahippocampal (SA), isthmus of cingulate (SA), rostral anterior cingulate (CT), cuneus (SA)
17	rs11672092 <sup>c</sup>	19p13.3	Intronic	T	0.22	0.0619	5.69 × 10 <sup>-10</sup>	<i>TUBB4A</i>	Isthmus of cingulate (CT), lateral orbitofrontal (SA), middle temporal (SA)
18	rs6135555	20p12.1	Intronic	A	0.39	0.0600	7.00 × 10 <sup>-9</sup>	<i>MACROD2</i>	Pericalcarine (SA), caudate (SUB)
19	rs7283026	21q22.3	Intronic	C	0.27	0.0616	8.42 × 10 <sup>-10</sup>	<i>COL18A1</i>	Supramarginal (SA), transverse temporal (SA)
20	rs9615351	22q13.31	Exonic	G	0.25	0.0588	3.02 × 10 <sup>-8</sup>	<i>CELSR1</i>	Isthmus of cingulate (CT), transverse temporal (SA)
21	rs12400461	Xp22.33	Intergenic	C	0.58	0.0595	1.24 × 10 <sup>-8</sup>	<i>ASSIP4</i>	Inferior temporal (SA), pars opercularis (SA)

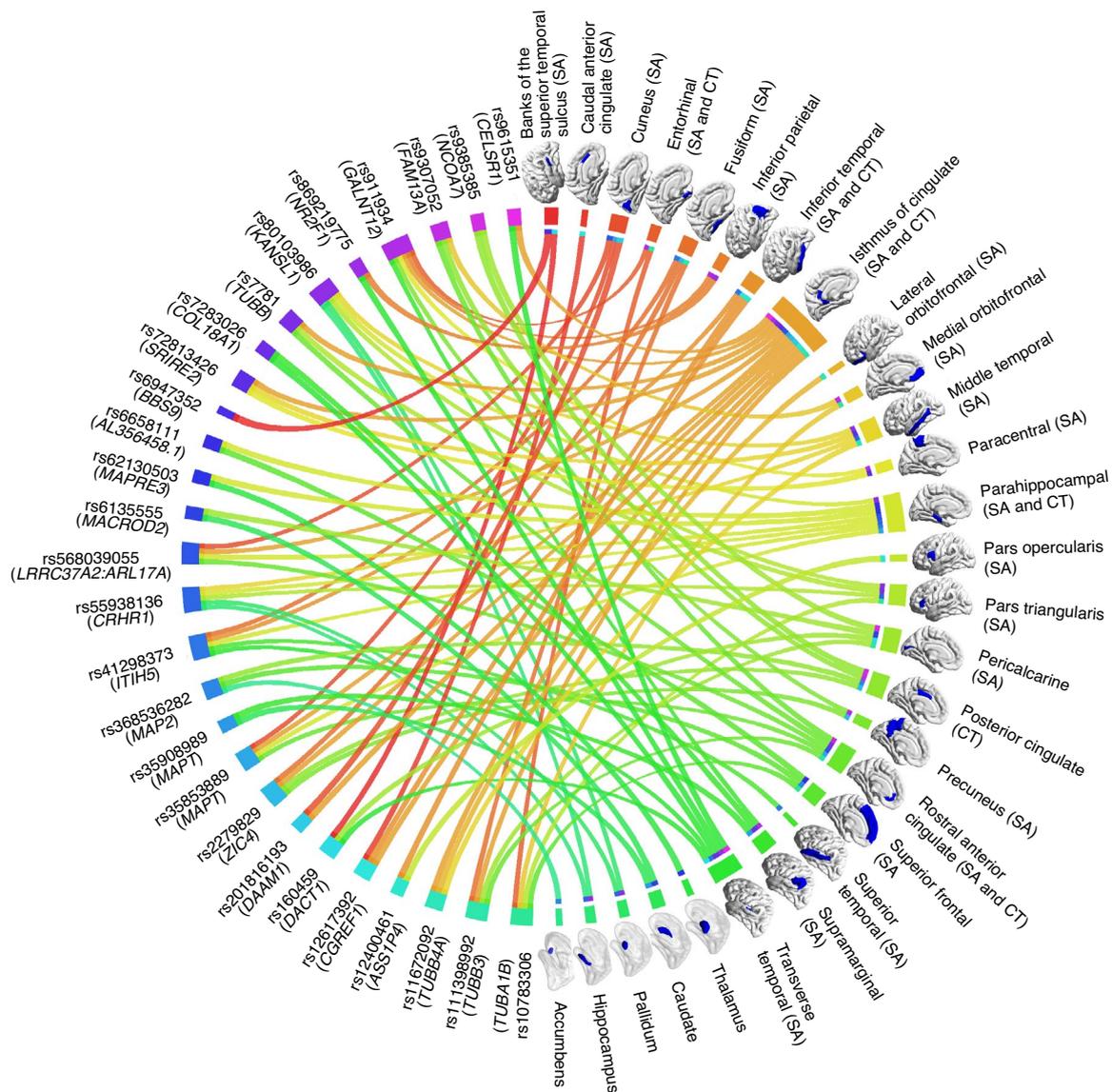
<sup>a</sup>Central traits for each SNP are those asymmetry indexes that contribute to its multivariate association (Methods). SA, surface area; CT, cortical thickness; SUB, subcortical volume. <sup>b</sup>Intronic to a gene for a non-coding RNA. <sup>c</sup>Lead variants are in high LD with handedness-associated variants. <sup>d</sup>Untranslated region.

the replication set (Supplementary Table 19). The combined *P* value of the remaining 17 lead SNPs was  $P = 3.3 \times 10^{-4}$  in the replication set (calculated by Stouffer's method), which we confirmed by permutation with respect to 10,000 repeat random samplings of 17 SNPs from the whole genome in the replication set (permutation-based  $P = 4 \times 10^{-4}$ ). This indicates that the limited sample size of the replication set, compared with the discovery set, did not provide adequate power to replicate at the level of some individual SNPs, but that in combination there was evidence for replication. Moreover, among the 17 SNPs that showed  $P > 0.05$  on multivariate testing in the replication set, some showed association  $P < 0.05$  on univariate testing of the specific central traits identified for those SNPs in the discovery mvGWAS (Supplementary Table 19). It is also worth noting that 4 of these 17 SNPs (or SNPs in high linkage disequilibrium with them) have been reported to associate with left-handedness at a genome-wide significant level<sup>44</sup> (see above for details), which is an additional form of validation with respect to a phenotype related to brain asymmetry. As also mentioned above, the high degree of functional clustering of genes identified through gene-based association testing is another form of support for the association results in the mvGWAS.

## Discussion

Despite the importance of asymmetry as an organizing feature of the human brain, the early developmental processes which establish its left–right axis are unknown. In this study, we carried out multivariate GWAS analysis which identified 21 genetic loci associated with different aspects of adult brain anatomical asymmetry. Functional annotation implicated genes particularly involved in microtubule organization and prenatal brain development. Our study therefore sheds new light on the molecular genetic foundations of human brain asymmetry. In addition, at the genome-wide level, SNPs associated with brain asymmetry overlapped with those associated with educational attainment, autism and schizophrenia, while five specific loci that are associated with both brain asymmetry and handedness were identified.

Previous studies in invertebrates and frog embryos have shown that the cytoskeleton plays a role in determining cellular chirality and asymmetrical patterning of other organs<sup>96–102</sup>. Cellular chirality refers to directional biases in cellular morphology, position, rotation or migration, which arise because of the inherent chirality of intracellular macromolecules such as those composing the cytoskeleton<sup>96</sup>, and also manifest in terms of the intracellular distributions



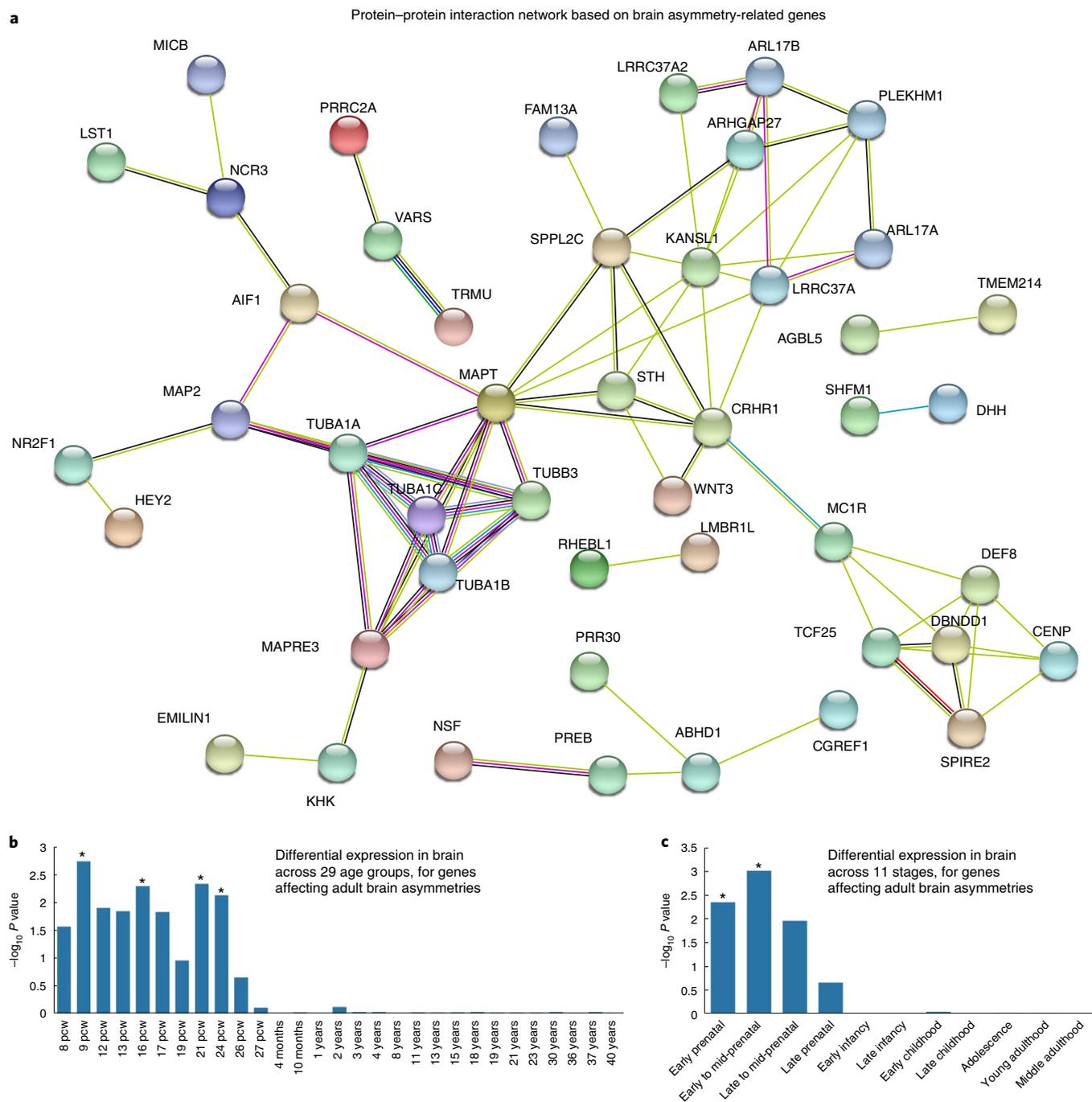
**Fig. 3 | Overview of 27 independent lead variants associated with different regional brain asymmetries.** Circle plot illustrating the 27 lead variants from mvGWAS (left) in relation to the central asymmetry indexes (right) underlying their specific multivariate associations. Different colours indicate different lead variants or regional asymmetries. Lines linking lead variants to regional asymmetries are coloured according to the regions. The closest genes to the lead variants are shown. Most central asymmetry indexes are of regional surface areas, and some variants affected multiple asymmetries of different types. SA, surface area; CT, cortical thickness; SUB, subcortical volume.

of organelles<sup>103</sup>. For example, during early cell divisions in *Xenopus* (frog) embryos, the cytoskeleton has been reported to mediate asymmetric intracellular protein localization, as protein transport molecules move along cytoskeletal tracks within cells<sup>104</sup>. Thus the cytoskeleton can provide a directionally consistent, organ-intrinsic bias during embryonic development that acts as a determinant of morphological asymmetry<sup>97,98,105</sup>, arising from fundamental aspects of molecular and cellular biology.

As our study associated cytoskeletal and embryonically expressed genes with variation in adult human brain asymmetry, it is possible that these genes are involved in the establishment of left–right asymmetry during early brain development, through a mechanism involving cellular chirality. As mentioned in Main, at least some aspects of human brain asymmetry appear to be uncoupled from the developmental pathway that leads to left–right organization of the visceral organs, which involves cilia and the nodal pathway.

A cytoskeletal-mediated mechanism of brain asymmetry may therefore be organ intrinsic<sup>97,98,105</sup>, that is, distinct from other pathways that establish broader aspects of body asymmetry.

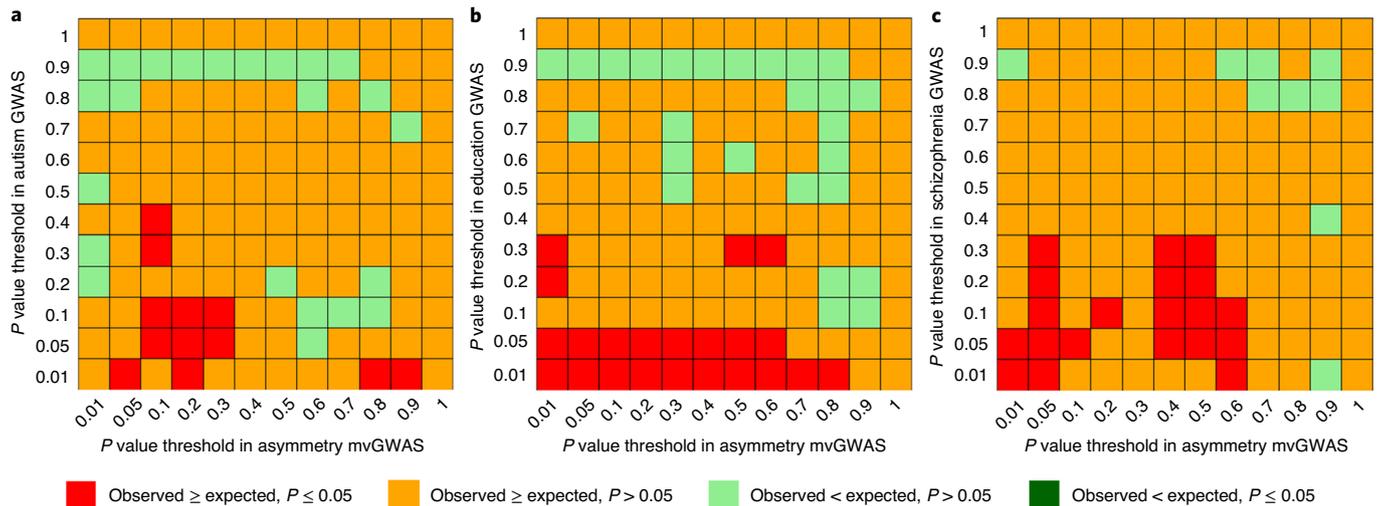
In this study, we identified genetic loci that are associated with 42 heritable aspects of brain asymmetry through a multivariate, brain-wide approach. A multivariate approach can boost statistical power while achieving data reduction, compared with separate univariate analyses of individual brain traits<sup>37</sup>. A single set of genome-wide association results, pertaining simultaneously to multiple aspects of brain asymmetry, was then taken forward into functional annotation and downstream analyses, such as testing for genetic overlaps with other traits. The multivariate approach therefore helped to detect and interpret key aspects of the genetic architecture of brain asymmetry, without the noise inherent in repeat univariate testing. An important challenge remained to identify the particular brain traits that drove the multivariate



**Fig. 4 | Functional annotations of variants associated with brain asymmetry.** **a**, Brain asymmetry-associated genes integrated into a protein–protein interaction network. Proteins are represented by nodes. Edges between nodes represent different types of protein–protein interactions according to the STRING database (Methods), including known interactions (turquoise and dark purple represent interactions identified by curated databases and biological experiments, respectively), predicted interactions (green, red and blue represent interactions predicted by gene neighbourhood, gene fusions and gene co-occurrence, respectively) and others (yellow, black and light purple represent interactions determined by text mining, co-expression and protein homology, respectively). Coloured nodes represent the queried proteins. Only medium-confidence (>0.4) links were retained, and disconnected proteins are not shown. **b**, Relation between gene-based association with brain asymmetries and relatively higher mRNA expression in the human brain at particular ages, using BrainSpan data from 29 age groups. Asterisks indicate significant age groups meeting  $P < 0.05$  with FDR correction. pcw, post-conceptual weeks. **c**, Relation between gene-based association with brain asymmetries and relatively higher mRNA expression in the human brain at particular ages, using BrainSpan data from 11 defined age groups. Asterisks indicate significant groups meeting  $P < 0.05$  with FDR correction.

associations at each locus, which was achieved in MetaPhat<sup>37</sup> by decomposing associations into sets of ‘central’ traits on the basis of the Bayesian information criterion and canonical correlation  $P$  values.

A consequence of the multivariate approach is that it does not yield univariate association effect sizes, and therefore mvGWAS results cannot be used for standard genetic correlation analyses, such as is performed with LD score regression<sup>106</sup>. Therefore, we used



**Fig. 5 | Genetic overlaps between brain asymmetries and other traits. a–c,** Heatmap plots illustrating pleiotropic effects between brain asymmetries and autism (a), educational attainment (b) and schizophrenia (c), based on per-SNP GWAS  $P$  values for these traits from previous studies (Methods), in relation to the mvGWAS  $P$  values from the present study of brain asymmetries.

iSECA<sup>73</sup> to explore the genetic overlap of brain asymmetry with neurodevelopmental disorders, behavioural and psychological traits. This analysis was based on SNP-wise  $P$  values of association over the whole genome, that is, a genetic overlap was found when the SNPs tending to show low  $P$  values in our mvGWAS for brain asymmetry also tended to show low  $P$  values in publicly available GWAS summary statistics for a given disorder or trait. We found significant genetic overlaps of brain asymmetry with autism, schizophrenia and educational attainment, which suggests that genes affecting brain asymmetry also influence these traits. This is in line with literature that has shown phenotypic associations between altered brain asymmetry and these traits (see Main), and indicates that such phenotypic associations are contributed to an extent by genetic factors. As we found that brain asymmetry-related genes tend to be especially highly expressed in the embryonic brain, it seems likely that the genetic overlap of brain asymmetry and disorders reflects a genetic susceptibility to alterations of early neurodevelopment away from the typical trajectory. However, brain asymmetry continues to develop throughout the lifespan<sup>107,108</sup>, and the UK Biobank consists of middle- to older-age adults, so that our mvGWAS may have also identified genetic factors that affect brain asymmetrical changes later in life.

Further research, for example, using Mendelian randomization<sup>109</sup>, will be needed to understand whether brain asymmetries mediate gene–disorder associations in a causal sense, or whether altered brain asymmetry and disorder susceptibility are two distinct consequences arising from a partly overlapping genetic basis. It will also be important to map, on a brain-regional basis, which aspects of asymmetry show the strongest genetic overlaps with disorder susceptibility. Larger imaging–genetic datasets may be needed to support causal mediation and mapping studies with respect to disorders, as the present genetic association analysis was based on brain-wide asymmetry (albeit in a multivariate context).

Many of the brain asymmetries were strong and directional at the population level, but their heritabilities were generally low, ranging up to 10%. This suggests that developmental mechanisms for brain asymmetry are tightly constrained and largely genetically invariant in the population, and that environmental factors and/or developmental randomness are responsible for most variability<sup>45,110–112</sup>. A cytoskeleton-based origin of brain asymmetry would fit this scenario, as the cytoskeleton is essential for various fundamental

functions in cellular biology, beyond axis formation<sup>113,114</sup>. Previous, twin- and family-based analyses<sup>5,6</sup> have reported heritabilities up to roughly 25% for some of the same asymmetry measures that we analysed in the present study, with an overall similar regional pattern. Higher heritabilities were found particularly for regions that are important in the language system (for example, superior temporal cortex) and limbic system (for example, medial temporal and cingulate cortex). Twin-based heritability is often measured to be higher than SNP-based heritability, which may be expected because SNPs are just one class of genetic variation, and also because twin studies can overestimate heritability when certain assumptions are not fully met<sup>115</sup>. As twin studies have not indicated effects of shared environment on brain asymmetries<sup>5,6</sup>, early developmental randomness is likely to cause most variation<sup>116</sup>.

We did not correct for handedness as a covariate in our genetic analyses, as it is generally not advisable to correct for covariates which are themselves partly heritable. This is because biased genetic effects can be measured with respect to the target trait<sup>117</sup> (in this case, brain asymmetry). Handedness cannot therefore be treated safely as a confound variable when analysing brain asymmetry. We were interested in any genetic effects associated with brain asymmetry, regardless of whether they might also be shared with other traits such as handedness. Having identified genetic variants associated with brain asymmetry, we then queried post hoc whether they have been reported as significant in previous GWAS of handedness in over 1 million people<sup>44</sup>. We did not observe a significant genetic overlap between structural brain asymmetry and handedness at the genome-wide level, which again may be due to the relatively low SNP-based heritabilities of these traits, in combination with limited statistical power in the present sample size for this kind of analysis. However, five individual SNPs associated with both brain asymmetry and handedness were identified, which suggests that a significant genome-wide overlap might be detected when using a larger dataset in the future.

The UK Biobank currently includes by far the largest single imaging–genetic dataset available to the scientific community. A limitation of the present study is the lack of a large, age-matched replication sample with comparable homogeneity of ancestry to the discovery mvGWAS analysis. We included an independent replication sample of 3,600 individuals from the UK Biobank, with greater diversity of ancestry than the 32,256 individuals of the primary

mvGWAS. Association was replicated for 10 of the 27 independent lead SNPs from the mvGWAS, and for the remaining 17, their combined  $P$  value in the replication set was  $3.3 \times 10^{-4}$ , indicating that a larger dataset would have supported replication at the level of more individual SNPs. Functional clustering of the closest genes to the 27 independent lead SNPs, according to microtubule-related biology and protein–protein interactions, also supported the validity of the mvGWAS findings, as this was unlikely to occur by chance. In addition, four of the SNPs that did not replicate individually at  $P < 0.05$  (or SNPs in high linkage disequilibrium with them; see Results) were previously associated at a genome-wide significant level with left-handedness, which is an additional form of validation with respect to a behavioural asymmetry trait.

Another limitation of the present study is that functional enrichment analysis with respect to post-mortem, developmental gene expression was only possible from roughly 8 weeks post conception and onwards. The Brainspan database does not contain sufficient numbers of samples from earlier embryonic stages than this. It may be that the left–right brain axis is established extremely early in development, for example during formation of the neural tube, which begins in the third week post conception<sup>118</sup>. As noted in Main, possibilities for research using human embryos from this developmental stage are highly restricted. Gene expression studies using animal models may therefore be necessary to understand the establishment of the mammalian brain's left–right axis.

In conclusion, our findings motivate genetic-developmental studies of left–right differentiation of the embryonic mammalian brain, focused on a possible cytoskeletal-mediated mechanism of axis formation. Our study also suggests that disruption of this mechanism may contribute to susceptibility to cognitive and psychiatric disorders, in line with asymmetry being an important aspect of healthy brain organization for many functions.

## Methods

**Participants.** This study was conducted under UK Biobank application 16066, with C.F. as principal investigator. The UK Biobank is a general adult population cohort. The UK Biobank received ethical approval from the National Research Ethics Service Committee North West–Haydock (reference 11/NW/0382), and all of their procedures were performed in accordance with the World Medical Association guidelines. Informed consent was obtained for all participants. We used the brain imaging data released in February 2020, and data availability and processing (described below) resulted in a final sample of 32,256 participants of white British ancestry for the primary GWAS analysis, together with the structural MRI data and genotype data from the same participants. The age range of these participants was 45 to 81 years (mean 63.77 years), 15,288 were male and 16,968 were female. An independent replication dataset of 3,600 individuals was also drawn from the UK Biobank, who self-identified as white, but not British, or did not cluster genetically with the bulk of the 'white British' ancestry participants (see below). The age range of these participants was 45 to 80 years (mean 62.89 years), 1,574 were male and 2,026 were female.

**Genetic quality control.** We downloaded imputed SNP genotype data from the UK Biobank data portal (bgen files; imputed data v3-release March 2018). We first excluded subjects with a mismatch of their self-reported and genetically inferred sex, with putative sex chromosome aneuploidies, or who were outliers on the basis of heterozygosity (principle-component-corrected heterozygosity  $> 0.19$ ) and genotype missingness (missing rate  $> 0.05$ ) as calculated by Bycroft et al.<sup>95</sup>. The primary analyses were restricted to participants with 'white British ancestry', which was defined by Bycroft et al. ('in.white.British.ancestry.subset')<sup>95</sup>, using a combination of self-report and cluster analysis on the basis of the first six principal components which capture genetic ancestry from the genome-wide genotype data. We randomly excluded one from each pair of remaining individuals who had a kinship coefficient  $> 0.0442$ , as defined by Bycroft et al.<sup>95</sup>. Next, QCTOOL (v.2.0.6) and PLINK<sup>119</sup> were used to perform genotype quality control: excluding SNPs with minor allele frequency  $< 1\%$ , Hardy–Weinberg equilibrium test  $P$  value  $< 1 \times 10^{-7}$  and imputation INFO score  $< 0.7$  (a measure of genotype imputation confidence). We also excluded multi-allelic SNPs because most of the downstream software (below) could not handle them. This resulted in 9,803,522 bi-allelic variants.

The same process was applied to derive the independent replication dataset of 3,600 individuals, except that these participants did not self-identify as British (although they did identify as 'white'), or they did not fall within the bounds of the 'white British ancestry' cluster as defined by Bycroft et al.<sup>95</sup>. We also imposed the

extra requirement that participants in the replication set should not show a kinship coefficient  $> 0.0442$  with any individual in the primary discovery dataset.

**Neuroimaging phenotypes and covariates.** Brain anatomical measures of regional cortical surface area, cortical thickness and subcortical volumes were derived from the structural scans (Siemens Skyra 3-T MRI with 32-channel radiofrequency receive head coil) released by the UK Biobank Imaging Study (for the full protocol, see <http://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=2367>). Briefly, in vivo whole-brain T1-weighted MRI scans were used to perform cortical parcellation into 34 regions per hemisphere with the Desikan–Killiany atlas<sup>120</sup>, and 7 subcortical structural segmentations. Surface area was measured at the grey–white matter boundary, and thickness was measured as the average distance in a region between the white matter and pial surfaces. Details of the image quality control and processing are described elsewhere<sup>121</sup>. Given that the data for the temporal pole were reported as unreliable<sup>121</sup>, we only used 33 surface area, 33 cortical thickness and 7 subcortical volume measures in each hemisphere (Supplementary Table 1). Per measure, we removed data points greater than six standard deviations from the mean. Then, we calculated the AI for each matching pair of left and right measures, in each participant, as  $(\text{left} - \text{right}) / ((\text{left} + \text{right}) / 2)$ . Given this definition, a positive AI reflects leftward asymmetry (greater left than right). The AI is a widely used measure in brain asymmetry studies<sup>5,122,123</sup>. The denominator ensures that the index does not simply scale with brain size, that is, the left–right difference is adjusted for the bilateral measure. For each AI, one-sample  $t$  testing was used to examine whether the population mean AI was significantly different from zero, with Bonferroni correction at 0.05 for multiple testing. Subsequently, the distributions of AIs were normalized by rank-based inverse normalization to minimize statistical artifacts. The normalized AIs were used as input for subsequent analysis.

The Desikan–Killiany atlas<sup>120</sup> was derived from manual segmentations of sets of reference brain images. The labelling system incorporates hemisphere-specific information on sulcal and gyral geometry with spatial information regarding the locations of brain structures, and shows a high accuracy when compared with manual labelling results<sup>120</sup>. Accordingly, the mean regional asymmetries in the UK Biobank might partly reflect left–right differences present in the reference dataset used to construct the atlas. However, our study was focused primarily on comparing relative asymmetry between genotypes, at the regional level. The use of an asymmetrical atlas based on healthy individuals had the advantage that regional identification was likely to be accurate for structures that are asymmetrical in the general population, while taking hemisphere-specific information into account.

We also made use of continuous variables as covariates in heritability estimation and genome-wide association analysis (below), which were: age when attended assessment centre (UK Biobank fields 21003-2.0), nonlinear age, that is  $(\text{age} - \text{mean\_age})^2$ , the first ten genetic principal components capturing population genetic diversity (fields 22009-0.1 to 22009-0.10) (or the first 40 principal components in the replication dataset with higher diversity of ancestry), scanner position parameters ( $X$ ,  $Y$  and  $Z$  position: fields 25756-2.0, 25757-2.0 and 25758-2.0), T1 signal-to-noise ratio (field 25734-2.0) and T1 contrast-to-noise ratio (field 25735-2.0), plus categorical covariates which were: assessment centre (field 54-2.0), genotype measurement batch (field 22000-0.0) and sex (field 31-0.0).

**SNP-based heritability and genetic correlation analysis within the UK Biobank data.** From the primary dataset, 9,516,074 autosomal variants with minor allele frequencies  $> 1\%$ , INFO score  $> 0.7$  and Hardy–Weinberg equilibrium  $P > 1 \times 10^{-7}$  were used to build a genetic relationship matrix using GCTA<sup>36</sup> (version 1.93.0beta). Specifically for analyses using GCTA, we further excluded one random participant from each pair having a kinship coefficient higher than 0.025 based on the calculated genetic relationship matrix (as this analysis is especially sensitive to higher levels of relatedness), resulting in 30,315 participants. Genome-based restricted maximum likelihood (GREML)<sup>36</sup> analyses were performed to estimate the SNP-based heritability for each AI, controlling for the above-mentioned covariates, and applying FDR 0.05 across the 73 AIs to define significantly heritable AIs. Bivariate GREML<sup>124</sup> analysis was used to estimate genetic correlations between pairs of AIs, separately for cortical surface area, cortical thickness and subcortical volume AIs, with FDR correction at 0.05 for multiple testing.

**Multivariate genome-wide association analysis.** In mvGWAS, a single association test is performed for each SNP in relation to multiple traits simultaneously. We used MetaPhat<sup>27</sup> software to perform mvGWAS analysis across asymmetries for cortical surface area, cortical thickness and subcortical volume, including only the 42 AIs that had shown significant SNP-based heritability. MetaPhat performs meta-canonical correlation analysis, and uses univariate GWAS summary statistics as input from each separate AI, which were derived under an additive genetic model while controlling for the above-mentioned covariates, using BGENIE software (v.1.2)<sup>95</sup>. Thus our mvGWAS tested effectively for association with 42 traits. This approach estimates the linear combination of traits that is maximally associated with genotype, which can differ for each SNP, while maintaining a correct false-positive rate. A total of 9,803,522 SNPs (see further above) were used for mvGWAS, spanning all autosomes and chromosome X. Statistically significant SNPs were considered as those with  $P < 5 \times 10^{-8}$  in mvGWAS, which is a widely

used threshold to account for multiple testing over the whole genome, in the context of LD in European-descent populations<sup>125,126</sup>.

MetaPhat also uses systematic criteria to define central traits which make the greatest contributions to significant multivariate associations, on the basis of an iterative process to optimize multivariate model properties with reference to canonical correlation analysis *P* values and the Bayesian information criterion<sup>37</sup>. For the lead SNPs at genome-wide significant loci (see below for how these were defined), we also performed post hoc analysis in which we examined their separate left and right hemispheric associations, using traits corresponding to the central AIs that were involved in the multivariate associations (Supplementary Table 11), again using BGENIE, an additive genetic model and covariates as described above.

As a sensitivity analysis, we re-ran the mvGWAS after excluding from the primary dataset 886 participants who had lifetime diagnoses of neurological conditions that could potentially disrupt brain structure (Supplementary Table 20). The significant mvGWAS loci were minimally affected by this exclusion (Supplementary Fig. 9).

**Identification of genomic risk loci and functional annotations.** FUMA (version v1.3.6)<sup>38</sup>, an online platform for functional annotation of GWAS results, was applied to the results from mvGWAS. A multi-step process, using default parameters, was used to identify distinct, significantly associated genomic loci, and independent lead SNPs within those loci. Briefly, on the basis of the pre-calculated LD structure from the 1000 Genomes European reference panel<sup>127</sup>, SNPs with genome-wide significant mvGWAS *P* values  $<5 \times 10^{-8}$  that had LD  $r^2 < 0.6$  with any others were identified. For each of these SNPs, other SNPs that had  $r^2 \geq 0.6$  with them were included for further annotation (see below), and independent lead SNPs were also defined among them as having low LD ( $r^2 < 0.1$ ) with any others. If LD blocks of significant SNPs are located within 250 kb of each other (default parameter), they are merged into one genomic locus. Therefore, some genomic loci could include one or more independent lead SNPs (Table 1). The major histocompatibility complex region on chromosome 6 was excluded from this process by default, because of its especially complex and long-range LD structure.

Functional annotations were applied by matching chromosome location, base-pair position, reference and alternate alleles to databases containing known functional annotations, which were ANNOVAR<sup>128</sup> categories, Combined Annotation-Dependent Depletion<sup>129</sup> scores, RegulomeDB<sup>130</sup> scores and chromatin state<sup>131,132</sup>:

1. ANNOVAR categories identify SNPs on the basis of their locations with respect to genes, such as exonic, intronic and intergenic, using Ensembl gene definitions.
2. Combined Annotation-Dependent Depletion scores predict deleteriousness, with scores higher than 12.37 suggesting potential pathogenicity<sup>133</sup>.
3. RegulomeDB scores integrate regulatory information from eQTL and chromatin marks, and range from 1a to 7, with lower scores representing more importance for regulatory function.
4. Chromatin states show the accessibility of genomic regions, and were labelled by 15 categorical states on the basis of five chromatin marks for 127 epigenomes in the Roadmap Epigenomics Project<sup>132</sup>, which were H3K4me3, H3K4me1, H3K36me3, H3K27me3 and H3K9me3. For each SNP, FUMA calculated the minimum chromatin state across 127 tissue/cell-type in the Roadmap Epigenomics Project<sup>132</sup>. Categories 1–7 are considered as open chromatin states.

We also used FUMA to annotate independent significant SNPs and their candidate SNPs according to previously reported phenotype associations ( $P < 5 \times 10^{-9}$ ) in the National Human Genome Research Institute–European Bioinformatics Institute catalogue<sup>134</sup>.

For a significant mvGWAS association in the major histocompatibility complex region (Table 1), we took the most significant individual SNP, rs7781 ( $P = 1.62 \times 10^{-10}$ ), as the single lead SNP to represent this locus, and annotated it manually.

**SNP-to-gene mapping.** SNP-to-gene mapping at significant mvGWAS loci was performed using the default FUMA processes for these three strategies:

1. Positional mapping was used to map SNPs to protein-coding genes on the basis of physical distance (within 10 kb) in the human reference assembly (GRCh37/hg19).
2. eQTL mapping was used to annotate SNPs to genes (that is, when SNP genotypes are associated with variation in gene mRNA expression levels). eQTL mapping was carried out in relation to genes up to 1 Mb away on the basis of four brain-expression data repositories: PsychENCORE<sup>52</sup>, CommonMind Consortium<sup>39</sup>, BRAINEAC<sup>51</sup> and GTEx v8 Brain<sup>135</sup>. FUMA applied a FDR of 0.05 within each analysis to identify significant eQTL associations.
3. Chromatin interaction mapping was performed to map SNPs to genes on the basis of seven brain-related Hi-C chromatin conformation capture datasets: PsychENCORE EP link (one way)<sup>52</sup>, PsychENCORE promoter anchored loops<sup>39</sup>, HiC adult cortex<sup>46</sup>, HiC foetal cortex<sup>46</sup>, HiC (GSE87112) dorsolateral prefrontal cortex<sup>60</sup>, HiC (GSE87112) hippocampus<sup>60</sup> and HiC (GSE87112)

neural progenitor cell<sup>60</sup>. We further selected only those genes for which one or both regions involved in the chromatin interaction overlapped with a predicted enhancer or promoter region (250 bp up- and 500 bp downstream of the transcription start site) in any of the brain-related repositories from the Roadmap Epigenomics Project<sup>132</sup>, that is, E053 (neurospheres) cortex, E054 (neurospheres) ganglion eminence, E067 (brain) angular gyrus, E068 (brain) anterior caudate, E069 (brain) cingulate gyrus, E070 (brain) germinal matrix, E071 (brain) hippocampus middle, E072 (brain) inferior temporal lobe, E073 (brain) dorsolateral prefrontal cortex, E074 (brain) substantia nigra, E081 (brain) foetal brain male, E082 (brain) foetal brain female, E003 embryonic stem (ES) H1 cells, E008 ES H9 cells, E007 (ES-derived) H1 derived neuronal progenitor cultured cells, E009 (ES-derived) H9 derived neuronal progenitor cultured cells and E010 (ES-derived) H9 derived neuron cultured cells. A FDR of  $1 \times 10^{-6}$  was applied to identify significant interactions (default parameter), separately for each analysis.

**Gene-based association analysis.** Genome-wide gene-based association analysis was performed using mvGWAS summary statistics as input into MAGMA (v1.08)<sup>65</sup>, using default parameters implemented in FUMA (SNP-wide mean model). This process examines the joint association signals of all SNPs within a given gene (including 50 kb upstream to 50 kb downstream of the gene), while considering the LD between the SNPs. SNPs were mapped to 20,146 protein-coding genes on the basis of National Center for Biotechnology Information build 37.3 gene definitions, and each gene was represented by at least one SNP. We applied a Bonferroni correction for the number of tested genes ( $P < 0.05/20,146$ ).

**Gene-set enrichment analysis.** We used MAGMA<sup>65</sup>, again with default settings as implemented in FUMA, to test for enrichment of association within predefined gene sets. This process tests whether gene-based *P* values among all 20,146 genes are lower for those genes within pre-defined functional sets than the rest of the genes in the genome, while correcting for other gene properties such as the number of SNPs. A total of 7,343 gene sets, defined according to Gene Ontology biological processes, were tested from MSigDB version 7.0<sup>66</sup>. In the main text we report the gene sets with *P* values that met Bonferroni correction for multiple testing ( $P < 0.05/7,343$ ).

In addition, we used the list of single closest genes to the 27 lead SNPs arising from mvGWAS (Table 1) as input for gene set enrichment analysis, using the same 7,343 Gene Ontology biological process gene sets, but now on the basis of the hypergeometric test as implemented in GENE2FUNC of FUMA<sup>38</sup>, which is appropriate for gene lists.

Finally, we used the CELL TYPE function (as implemented within FUMA) to test whether lower gene-based association *P* values for brain asymmetry were associated with differential expression levels across cell types, using Bonferroni correction within each separate analysis with respect to each cell-type expression dataset included in FUMA.

**Protein–protein interaction network.** We used the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING; <http://string-db.org>)<sup>68</sup> for protein network analysis, using as input the names of 57 genes identified through gene-based association analysis, as described above. The STRING dataset includes protein–protein interaction information from numerous sources, including experimental data, publications and computational prediction methods. Only links with medium confidence or higher (confidence score  $>0.4$ ; default parameter) were retained.

**Developmental stage analysis.** Using the gene-based association *P* values for all 20,146 genes genome wide, we used MAGMA (default settings as implemented in FUMA) to examine whether lower gene-based *P* values tended to be found for genes showing relatively higher expression in BrainSpan<sup>71</sup> gene expression data from any particular ages compared with all other ages, separately for 29 different age groups ranging from 8 postconceptional weeks to 40 years old, and 11 defined developmental stages from early prenatal to middle adulthood. We corrected for multiple testing through a FDR of 0.05 (separately for the two analyses).

**Genetic overlap of brain asymmetry with brain disorders, behavioural and cognitive traits.** We applied the iSECA<sup>73</sup> toolbox that can test for genetic overlap on the basis of per-SNP association *P* values only (mvGWAS does not produce univariate beta coefficient effect size estimates that can be used in standard genetic correlation analysis). We tested for genetic overlap in relation to traits previously reported to associate phenotypically with different aspects of brain structural asymmetry (Main), using GWAS *P* values from previously published, large-scale studies: educational attainment ( $n = 1,131,881$ )<sup>66</sup>, handedness ( $n = 331,037$ )<sup>65</sup>, intelligence ( $n = 269,867$ )<sup>67</sup>, autism spectrum disorder ( $n = 46,350$ )<sup>78</sup>, attention deficit hyperactivity disorder ( $n = 55,374$ )<sup>74</sup> and schizophrenia ( $n = 82,315$ )<sup>69</sup>. We also tested for genetic overlap in relation to brain intracranial volume ( $n = 32,438$ )<sup>94</sup>. After LD-based filtering and clumping using default parameters, iSECA tests for pleiotropy between two sets of GWAS results using an exact binomial statistical test at each of 12 *P* value levels:  $P \leq (0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1)$ . The analysis compares the expected and observed overlap

in the subsets of SNPs at these levels from two GWAS (144 combinations in total). In other words, iSECA iterates through each of the 12  $P$  value levels and counts the number of overlapping variants between two GWAS at each  $P$  value threshold, and compares that number with the number expected under the null hypothesis of no genetic overlap, using the exact binomial test. iSECA then counts up the number of comparisons with evidence of overlap at a nominally significant level of  $P \leq 0.05$ . To assess the significance level of overlap, we generated 1000 datasets through permutations (default parameter), which contained all the possible combinations for a pair of traits, and determined whether the number of levels with nominally significant genetic overlap was significantly more than expected by chance. Finally, Bonferroni correction  $< 0.05$  was applied for multiple testing of seven traits. Additionally, iSECA generated Q–Q plots for asymmetry mvGWAS  $P$  values conditioned on the other trait  $P$  values (for example,  $P \leq 0.1, 0.2, 0.3, 0.4, 0.5, 0.75, 1.0$ ) to visualize whether there is an excess of pleiotropic SNPs, which should be visible as a leftward shift of the curve as the  $P$  value threshold becomes tighter (Supplementary Fig. 8).

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data Availability

The primary data used in this study are available via the UK Biobank, <https://www.ukbiobank.ac.uk>. Other publicly available data sources and applications are cited in the Methods section. The GWAS summary statistics are made available online within the GWAS catalogue <https://www.ebi.ac.uk/gwas/>.

### Code availability

This study used openly available software and codes, specifically GCTA (<https://cns.genomics.com/software/gcta/#GREML>), BGENIE (<https://jmarchini.org/bgenie/>), MetaPhat (<https://sourceforge.net/projects/meta-pheno-association-tracer/>), FUMA (<https://fuma.ctglab.nl/>), MAGMA (<https://ctg.cncr.nl/software/magma>, also implemented in FUMA) and iSECA (<https://web.archive.org/web/20190326171900/https://neurogenetics.qimrberghofer.edu.au/iSECA/>).

Received: 7 July 2020; Accepted: 5 February 2021;

Published online: 15 March 2021

### References

- Duboc, V., Dufourcq, P., Blader, P. & Roussigne, M. Asymmetry of the brain: development and implications. *Annu. Rev. Genet.* **49**, 647–672 (2015).
- Kong, X.-Z. et al. Handedness and other variables associated with human brain asymmetrical skew. Preprint at *bioRxiv* <https://doi.org/10.1101/756395> (2019).
- Galaburda, A. M., Corsiglia, J., Rosen, G. D. & Sherman, G. F. Planum temporale asymmetry, reappraisal since Geschwind and Levitsky. *Neuropsychologia* **25**, 853–868 (1987).
- Herve, P. Y., Crivello, F., Perchev, G., Mazoyer, B. & Tzourio-Mazoyer, N. Handedness and cerebral anatomical asymmetries in young adult males. *Neuroimage* **29**, 1066–1079 (2006).
- Kong, X. Z. et al. Mapping cortical brain asymmetry in 17,141 healthy individuals worldwide via the ENIGMA Consortium. *Proc. Natl Acad. Sci. U. S. A.* **115**, E5154–E5163 (2018).
- Guadalupe, T. et al. Human subcortical brain asymmetries in 15,847 people worldwide reveal effects of age and sex. *Brain Imaging Behav.* **11**, 1497–1514 (2017).
- Toga, A. W. & Thompson, P. M. Mapping brain asymmetry. *Nat. Rev. Neurosci.* **4**, 37–48 (2003).
- Mazoyer, B. et al. Gaussian mixture modeling of hemispheric lateralization for language in a large sample of healthy individuals balanced for handedness. *PLoS ONE* **9**, e101165 (2014).
- Corballis, M. C. The evolution of lateralized brain circuits. *Front Psychol.* **8**, 1021 (2017).
- Güntürkün, O., Ströckens, F. & Ocklenburg, S. Brain lateralization: a comparative perspective. *Physiol. Rev.* **100**, 1019–1063 (2020).
- Ocklenburg, S., Hirnstein, M., Beste, C. & Güntürkün, O. Lateralization and cognitive systems. *Front. Psychol.* **5**, 1143 (2014).
- Rentería, M. E. Cerebral asymmetry: a quantitative, multifactorial, and plastic brain Phenotype. *Twin Res. Hum. Genet.* **15**, 401–413 (2012).
- Vingerhoets, G. Phenotypes in hemispheric functional segregation? Perspectives and challenges. *Phys. Life Rev.* **30**, 1–18 (2019).
- Karolis, V. R., Corbetta, M. & Thiebaut de Schotten, M. The architecture of functional lateralisation and its relationship to callosal connectivity in the human brain. *Nat. Commun.* **10**, 1417 (2019).
- Postema, M. C. et al. Altered structural brain asymmetry in autism spectrum disorder in a study of 54 datasets. *Nat. Commun.* **10**, 4958 (2019).
- Postema, M. C. et al. Analysis of structural brain asymmetries in attention-deficit/hyperactivity disorder in 39 datasets. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.03.03.974758> (2020).
- Okada, N. et al. Abnormal asymmetries in subcortical brain volume in schizophrenia. *Mol. Psychiatry* **21**, 1460–1466 (2016).
- Kasprian, G. et al. The prenatal origin of hemispheric asymmetry: an in utero neuroimaging study. *Cereb. Cortex* **21**, 1076–1083 (2011).
- Abu-Rustum, R. S., Ziade, M. F. & Abu-Rustum, S. E. Reference values for the right and left fetal choroid plexus at 11 to 13 weeks: an early sign of “developmental” laterality? *J. Ultrasound Med* **32**, 1623–1629 (2013).
- McCartney, G. & Hepper, P. Development of lateralized behaviour in the human fetus from 12 to 27 weeks’ gestation. *Dev. Med Child Neurol.* **41**, 83–86 (1999).
- Francks, C. In search of the biological roots of typical and atypical human brain asymmetry: Comment on “Phenotypes in hemispheric functional segregation? Perspectives and challenges” by Guy Vingerhoets. *Phys. Life Rev.* **30**, 22–24 (2019).
- Francks, C. Exploring human brain lateralization with molecular genetics and genomics. *Ann. N. Y. Acad. Sci.* **1359**, 1–13 (2015).
- Vandenberg, L. N., Lemire, J. M. & Levin, M. It’s never too early to get it right: a conserved role for the cytoskeleton in left–right asymmetry. *Commun. Integr. Biol.* **6**, e27155 (2013).
- Vandenberg, L. N. & Levin, M. Perspectives and open problems in the early phases of left–right patterning. *Semin. Cell Dev. Biol.* **20**, 456–463 (2009).
- Norris, D. P. Cilia, calcium and the basis of left–right asymmetry. *BMC Biol.* **10**, 102 (2012).
- Fliegeauf, M., Benzeng, T. & Omran, H. When cilia go bad: cilia defects and ciliopathies. *Nat. Rev. Mol. Cell Biol.* **8**, 880–893 (2007).
- Postema, M. C., Carrion-Castillo, A., Fisher, S. E., Vingerhoets, G. & Francks, C. The genetics of situs inversus without primary ciliary dyskinesia. *Sci. Rep.* **10**, 3677 (2020).
- McManus, I. C., Martin, N., Stubbings, G. F., Chung, E. M. & Mitchison, H. M. Handedness and situs inversus in primary ciliary dyskinesia. *Proc. Biol. Sci.* **271**, 2579–2582 (2004).
- Vingerhoets, G. et al. Brain structural and functional asymmetry in human situs inversus totalis. *Brain Struct. Funct.* **223**, 1937–1952 (2018).
- Tanaka, S., Kanzaki, R., Yoshiyoshi, M., Kamiya, T. & Sugishita, M. Dichotic listening in patients with situs inversus: brain asymmetry and situs asymmetry. *Neuropsychologia* **37**, 869–874 (1999).
- de Kovel, C. G. F. et al. Left–right asymmetry of maturation rates in human embryonic neural development. *Biol. Psychiatry* **82**, 204–212 (2017).
- Ocklenburg, S. et al. Epigenetic regulation of lateralized fetal spinal gene expression underlies hemispheric asymmetries. *eLife* **6**, e22784 (2017).
- Sun, T. et al. Early asymmetry of gene transcription in embryonic human left and right cerebral cortex. *Science* **308**, 1794–1798 (2005).
- Le Guen, Y. et al. Enhancer locus in *ch14q23.1* modulates brain asymmetric temporal regions involved in language processing. *Cereb. Cortex* **30**, 5322–5332 (2020).
- Carrion-Castillo, A. et al. Genetic effects on planum temporale asymmetry and their limited relevance to neurodevelopmental disorders, intelligence or educational attainment. *Cortex* **124**, 137–153 (2020).
- Yang, J. et al. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
- Lin, J., Tabassum, R., Ripatti, S. & Pirinen, M. MetaPhat: detecting and decomposing multivariate associations from univariate genome-wide association statistics. *Front. Genet.* **11**, 431 (2020).
- Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
- Fromer, M. et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat. Neurosci.* **19**, 1442–1453 (2016).
- Berezniuk, I. et al. Cytosolic carboxypeptidase 5 removes  $\alpha$ - and  $\gamma$ -linked glutamates from tubulin. *J. Biol. Chem.* **288**, 30445–30453 (2013).
- Janke, C. & Magiera, M. M. The tubulin code and its role in controlling microtubule properties and functions. *Nat. Rev. Mol. Cell Biol.* **21**, 307–326 (2020).
- Bernhardt, R. & Matus, A. Light and electron microscopic studies of the distribution of microtubule-associated protein 2 in rat brain: a difference between dendritic and axonal cytoskeletons. *J. Comp. Neurol.* **226**, 203–221 (1984).
- Wiberg, A. et al. Handedness, language areas and neuropsychiatric diseases: insights from brain imaging and genetics. *Brain* **142**, 2938–2947 (2019).
- Partida, G. C. et al. Genome-wide association study identifies 48 common genetic variants associated with handedness. *Nat. Hum. Behav.* **5**, 59–70 (2020).
- de Kovel, C. G. F. & Francks, C. The molecular genetics of hand preference revisited. *Sci. Rep.* **9**, 5986 (2019).
- Giusti-Rodríguez, P. M. & Sullivan, P. F. Using three-dimensional regulatory chromatin interactions from adult and fetal cortex to interpret genetic

- results for psychiatric disorders and cognitive traits. Preprint at *bioRxiv* <https://doi.org/10.1101/406330> (2019).
47. Lu, J. et al. Structure of the FH2 domain of Daam1: implications for formin regulation of actin assembly. *J. Mol. Biol.* **369**, 1258–1269 (2007).
  48. Ang, S.-F., Zhao, Z.-S., Lim, L. & Manser, E. DAAM1 is a formin required for centrosome re-orientation during cell migration. *PLoS ONE* **5**, e13064 (2010).
  49. Lancaster, O. M. & Baum, B. Shaping up to divide: coordinating actin and microtubule cytoskeletal remodelling during mitosis. *Semin. Cell Dev. Biol.* **34**, 109–115 (2014).
  50. Tischfield, M. A. et al. Human *TUBB3* mutations perturb microtubule dynamics, kinesin interactions, and axon guidance. *Cell* **140**, 74–87 (2010).
  51. Ramasamy, A. et al. Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat. Neurosci.* **17**, 1418–1428 (2014).
  52. Wang, D. et al. Comprehensive functional genomic resource and integrative model for the human brain. *Science* **362**, eaat8464 (2018).
  53. Horng, S. et al. Differential gene expression in the developing lateral geniculate nucleus and medial geniculate nucleus reveals novel roles for *Zic4* and *Foxp2* in visual and auditory pathway development. *J. Neurosci.* **29**, 13672–13683 (2009).
  54. Del Pino, I. et al. COUP-TFI/Nr2f1 orchestrates intrinsic neuronal activity during development of the somatosensory cortex. *Cereb. Cortex* **30**, 5667–5685 (2020).
  55. Shao, W., Halachmi, S. & Brown, M. ERAP140, a conserved tissue-specific nuclear receptor coactivator. *Mol. Cell. Biol.* **22**, 3358–3372 (2002).
  56. Beales, P. L., Elcioglu, N., Woolf, A. S., Parker, D. & Flintner, F. A. New criteria for improved diagnosis of Bardet–Biedl syndrome: results of a population survey. *J. Med. Genet.* **36**, 437–446 (1999).
  57. Zaghoul, N. A. & Katsanis, N. Mechanistic insights into Bardet–Biedl syndrome, a model ciliopathy. *J. Clin. Invest.* **119**, 428–437 (2009).
  58. Nenashva, V. V. & Tarantul, V. Z. Many faces of TRIM proteins on the road from pluripotency to neurogenesis. *Stem Cells Dev.* **29**, 1–14 (2020).
  59. Himmelfarb, M. et al. ITH5, a novel member of the inter-alpha-trypsin inhibitor heavy chain family is downregulated in breast cancer. *Cancer Lett.* **204**, 69–77 (2004).
  60. Schmitt, A. D. et al. A compendium of chromatin contact maps reveals spatially active regions in the human genome. *Cell Rep.* **17**, 2042–2059 (2016).
  61. O’Sullivan, M. L. et al. FLRT proteins are endogenous latrophilin ligands and regulate excitatory synapse development. *Neuron* **73**, 903–910 (2012).
  62. Sertie, A. L. et al. Collagen XVIII, containing an endogenous inhibitor of angiogenesis and tumor growth, plays a critical role in the maintenance of retinal structure and in neural tube closure (Knobloch syndrome). *Hum. Mol. Genet.* **9**, 2051–2058 (2000).
  63. Feng, J., Han, Q. & Zhou, L. Planar cell polarity genes, *Celsr1-3*, in neural development. *Neurosci. Bull.* **28**, 309–315 (2012).
  64. Al-Mubarak, B. et al. Whole exome sequencing reveals inherited and de novo variants in autism spectrum disorder: a trio study from Saudi families. *Sci. Rep.* **7**, 5679 (2017).
  65. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
  66. Lee, J. J. et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018).
  67. Savage, J. E. et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* **50**, 912–919 (2018).
  68. Szklarczyk, D. et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362–D368 (2017).
  69. Liberzon, A. et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
  70. Grasby, K. L. et al. The genetic architecture of the human cerebral cortex. *Science* **367**, eaay6690 (2020).
  71. Miller, J. A. et al. Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–206 (2014).
  72. de Kovel, C. G. F., Lisgo, S. N., Fisher, S. E. & Francks, C. Subtle left–right asymmetry of gene expression profiles in embryonic and foetal human brains. *Sci. Rep.* **8**, 12606 (2018).
  73. Nyholt, D. R. SECA: SNP effect concordance analysis using genome-wide association summary results. *Bioinformatics* **30**, 2086–2088 (2014).
  74. Demontis, D. et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* **51**, 63–75 (2019).
  75. Dang, L. C. et al. Caudate asymmetry is related to attentional impulsivity and an objective measure of ADHD-like attentional problems in healthy adults. *Brain Struct. Funct.* **221**, 277–286 (2016).
  76. Wu, Z. M. et al. Altered brain white matter microstructural asymmetry in children with ADHD. *Psychiatry Res* **285**, 112817 (2020).
  77. Zou, H. & Yang, J. Temporal variability-based functional brain lateralization study in ADHD. *J. Atten. Disord.* **25**, 839–847 (2019).
  78. Grove, J. et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **51**, 431–444 (2019).
  79. Carper, R. A., Treiber, J. M., DeJesus, S. Y. & Muller, R. A. Reduced hemispheric asymmetry of white matter microstructure in autism spectrum disorder. *J. Am. Acad. Child Adolesc. Psychiatry* **55**, 1073–1080 (2016).
  80. De Fosse, L. et al. Language-association cortex asymmetry in autism and specific language impairment. *Ann. Neurol.* **56**, 757–766 (2004).
  81. Floris, D. L. et al. Atypical brain asymmetry in autism—a candidate for clinically meaningful stratification. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* <https://doi.org/10.1016/j.bpsc.2020.08.008> (2020).
  82. Herbert, M. R. et al. Brain asymmetries in autism and developmental language disorder: a nested whole-brain analysis. *Brain* **128**, 213–226 (2005).
  83. Noroozian, M., Lotfi, J., Gassemzadeh, H., Emami, H. & Mehrabi, Y. Academic achievement and learning abilities in left-handers: guilt or gift? *Cortex* **38**, 779–785 (2002).
  84. Cheyney, C. P., Roberts, N., Crow, T. J., Leask, S. J. & Garcia-Finana, M. The effect of handedness on academic ability: a multivariate linear mixed model approach. *Laterality* **15**, 451–464 (2010).
  85. Mellet, E. et al. Weak language lateralization affects both verbal and spatial skills: an fMRI study in 297 subjects. *Neuropsychologia* **65**, 56–62 (2014).
  86. Papadatou-Pastou, M. & Tomprou, D. M. Intelligence and handedness: meta-analyses of studies on intellectually disabled, typically developing, and gifted individuals. *Neurosci. Biobehav. Rev.* **56**, 151–165 (2015).
  87. Prichard, E., Propper, R. E. & Christman, S. D. Degree of handedness, but not direction, is a systematic predictor of cognitive performance. *Front Psychol.* **4**, 9 (2013).
  88. Reio, T. G. Jr., Czarnolewski, M. & Eliot, J. Handedness and spatial ability: differential patterns of relationships. *Laterality* **9**, 339–358 (2004).
  89. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
  90. DeLisi, L. E. et al. Anomalous cerebral asymmetry and language processing in schizophrenia. *Schizophr. Bull.* **23**, 255–271 (1997).
  91. Shenton, M. E., Dickey, C. C., Frumin, M. & McCarley, R. W. A review of MRI findings in schizophrenia. *Schizophr. Res* **49**, 1–52 (2001).
  92. Kawasaki, Y. et al. Anomalous cerebral asymmetry in patients with schizophrenia demonstrated by voxel-based morphometry. *Biol. Psychiatry* **63**, 793–800 (2008).
  93. Sun, Y., Chen, Y., Collinson, S. L., Bezerianos, A. & Sim, K. Reduced hemispheric asymmetry of brain anatomical networks is linked to schizophrenia: a connectome study. *Cereb. Cortex* **27**, 602–615 (2017).
  94. Adams, H. H. et al. Novel genetic loci underlying human intracranial volume identified through genome-wide association. *Nat. Neurosci.* **19**, 1569–1582 (2016).
  95. Bycroft, C. et al. The UK biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
  96. Tee, Y. H. et al. Cellular chirality arising from the self-organization of the actin cytoskeleton. *Nat. Cell Biol.* **17**, 445–457 (2015).
  97. Inaki, M., Liu, J. & Matsuno, K. Cell chirality: its origin and roles in left–right asymmetric development. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **371**, 20150403 (2016).
  98. Okumura, T. et al. The development and evolution of left–right asymmetry in invertebrates: lessons from *Drosophila* and snails. *Dev. Dyn.* **237**, 3497–3515 (2008).
  99. Davison, A. et al. Formin is associated with left–right asymmetry in the pond snail and the frog. *Curr. Biol.* **26**, 654–660 (2016).
  100. Steinhauer, J. & Kalderon, D. Microtubule polarity and axis formation in the *Drosophila* oocyte. *Dev. Dyn.* **235**, 1455–1468 (2006).
  101. McNiven, M. A. & Porter, K. R. Organization of microtubules in centrosome-free cytoplasm. *J. Cell Biol.* **106**, 1593–1605 (1988).
  102. Lobikin, M. et al. Early, nonclialy role for microtubule proteins in left–right patterning is conserved across kingdoms. *Proc. Natl Acad. Sci. U. S. A.* **109**, 12586–12591 (2012).
  103. Fan, J., Zhang, H., Rahman, T., Stanton, D. N. & Wan, L. Q. Cell organelle-based analysis of cell chirality. *Commun. Integr. Biol.* **12**, 78–81 (2019).
  104. Levin, M. Is the early left–right axis like a plant, a kidney, or a neuron? The integration of physiological signals in embryonic asymmetry. *Birth Defects Res. Part C: Embryo Today: Rev.* **78**, 191–223 (2006).
  105. McDowell, G., Rajadurai, S. & Levin, M. From cytoskeletal dynamics to organ asymmetry: a nonlinear, regulative pathway underlies left–right patterning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **371**, 20150409 (2016).
  106. Bulik-Sullivan, B. K. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).

107. Zhou, D., Lebel, C., Evans, A. & Beaulieu, C. Cortical thickness asymmetry from childhood to older adulthood. *Neuroimage* **83**, 66–74 (2013).
108. Roe, J. M. et al. Asymmetric thinning of the cerebral cortex across the adult lifespan is accelerated in Alzheimer's disease. *Nat. Commun.* **12**, 721 (2021).
109. Davey Smith, G. & Ebrahim, S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* **32**, 1–22 (2003).
110. de Kovel, C. G. F., Carrion-Castillo, A. & Francks, C. A large-scale population study of early life factors influencing left-handedness. *Sci. Rep.* **9**, 584 (2019).
111. McManus, I. C. Handedness, language dominance and aphasia: a genetic model. *Psychol. Med Monogr. Suppl.* **8**, 1–40 (1985).
112. Bishop, D. V. M. & Bates, T. C. Heritability of language laterality assessed by functional transcranial doppler ultrasound: a twin study. *Wellcome Open Res* **4**, 161 (2019).
113. Janke, C. & Bulinski, J. C. Post-translational regulation of the microtubule cytoskeleton: mechanisms and functions. *Nat. Rev. Mol. Cell Biol.* **12**, 773–786 (2011).
114. Geiger, B., Bershady, A., Pankov, R. & Yamada, K. M. Transmembrane crosstalk between the extracellular matrix–cytoskeleton crosstalk. *Nat. Rev. Mol. Cell Biol.* **2**, 793–805 (2001).
115. Young, A. I. Solving the missing heritability problem. *PLoS Genet.* **15**, e1008222 (2019).
116. Mitchell, K. J. *Innate: How the Wiring of Our Brains Shapes Who We Are* (Princeton Univ. Press, 2018).
117. Aschard, H., Vilhjalmsson, 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).
118. Sadler, T. W. Embryology of neural tube development. *Am. J. Med. Genet. C. Semin. Med. Genet.* **135C**, 2–8 (2005).
119. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
120. Desikan, R. S. et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968–980 (2006).
121. Alfaro-Almagro, F. et al. Image processing and quality control for the first 10,000 brain imaging datasets from UK biobank. *Neuroimage* **166**, 400–424 (2018).
122. Kurth, F., Gaser, C. & Luders, E. A 12-step user guide for analyzing voxel-wise gray matter asymmetries in statistical parametric mapping (SPM). *Nat. Protoc.* **10**, 293–304 (2015).
123. Leroy, F. et al. New human-specific brain landmark: the depth asymmetry of superior temporal sulcus. *Proc. Natl Acad. Sci. U. S. A.* **112**, 1208–1213 (2015).
124. Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M. & Wray, N. R. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* **28**, 2540–2542 (2012).
125. Hoggart, C. J., Clark, T. G., De Iorio, M., Whittaker, J. C. & Balding, D. J. Genome-wide significance for dense SNP and resequencing data. *Genet. Epidemiol.* **32**, 179–185 (2008).
126. Panagiotou, O. A., Ioannidis, J. P. & Project, G.-W. S. What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations. *Int. J. Epidemiol.* **41**, 273–286 (2012).
127. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
128. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
129. Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J. & Kircher, M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* **47**, D886–D894 (2019).
130. Boyle, A. P. et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* **22**, 1790–1797 (2012).
131. Ernst, J. & Kellis, M. ChromHMM: automating chromatin-state discovery and characterization. *Nat. Methods* **9**, 215–216 (2012).
132. Roadmap Epigenomics Consortium. et al. Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
133. Kircher, M. et al. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310–315 (2014).
134. MacArthur, J. et al. The new NHGRI-EBI catalog of published genome-wide association studies (GWAS catalog). *Nucleic Acids Res.* **45**, D896–D901 (2017).
135. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).

## Acknowledgements

This research was funded by the Max Planck Society (Germany) and grants from the Netherlands Organization for Scientific Research (NWO) (054-15-101) and French National Research Agency (ANR, grant No. 15-HBPR-0001-03) as part of the FLAG-ERA consortium project 'MULTI-LATERAL', a Partner Project to the European Union's Flagship Human Brain Project. The authors thank N. Tzourio-Mazoyer and A. Pepe for contributions to the MULTI-LATERAL project. This research was conducted using the UK Biobank resource under application no. 16066, with C.F. as the principal applicant. Our study made use of imaging-derived phenotypes generated by an image-processing pipeline developed and run on behalf of UK Biobank. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

## Author contributions

Z.S.: conceptualization, methodology, analysis, visualization, original draft writing, review and editing. D.S.: methodology, analysis, bioinformatics, review and editing. A.C.-C.: conceptualization, methodology, analysis, visualization, review and editing. M.J.: conceptualization, funding acquisition, review and editing. B.M.: conceptualization, funding acquisition, review and editing. S.E.E.: conceptualization, funding acquisition, review and editing. F.C.: conceptualization, funding acquisition, review and editing. C.F.: conceptualization, direction, funding acquisition, supervision, original draft writing, review and editing.

## Competing interests

The authors declare that they have no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41562-021-01069-w>.

**Correspondence and requests for materials** should be addressed to C.F.

**Peer review information** *Nature Human Behaviour* thanks Philip Jansen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor(s): Jamie Horder; Charlotte Payne.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection Data were downloaded from the sources stated in the manuscript.

Data analysis Publicly available software and versions are stated in the Methods section together with the relevant citations.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The primary data used in this study are available via the UK Biobank website [www.ukbiobank.ac.uk](http://www.ukbiobank.ac.uk). Other publicly available data sources and applications are cited in the Methods section.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences  Behavioural & social sciences  Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size of >30,000 participants was determined by data availability. This range of sample size has been successfully used in many genome-wide association scan studies of diverse human phenotypes.
Data exclusions	Brain imaging phenotypes were excluded at > 6 SD from the mean to reduce the chance of including spurious datapoints or extreme points that could bias statistical testing. Single nucleotide polymorphisms were excluded when they had population frequencies below 1%, as statistical testing can be unreliable below this frequency.
Replication	We included a replication sample of >3000 individuals as described in the paper.
Randomization	Not relevant, this was an observational study.
Blinding	Not relevant.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

### Methods

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input type="checkbox"/>	<input checked="" type="checkbox"/> MRI-based neuroimaging

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	The Uk Biobank dataset has been extensively described before. We include demographic information in the paper for the specific subset used in our analysis (i.e. those with post-quality-control brain imaging and genetic data).
Recruitment	The Uk Biobank dataset has been extensively described before.
Ethics oversight	National Research Ethics Service Committee North West-Haydock (reference 11/NW/0382)

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Magnetic resonance imaging

### Experimental design

Design type	Structural T1 MRI
Design specifications	1 structural scan per subject
Behavioral performance measures	Not relevant

## Acquisition

Imaging type(s)

Field strength

Sequence & imaging parameters http://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=2367

Area of acquisition

Diffusion MRI  Used  Not used

## Preprocessing

Preprocessing software

Normalization

Normalization template

Noise and artifact removal

Volume censoring

## Statistical modeling & inference

Model type and settings https://www.frontiersin.org/articles/10.3389/fgene.2020.00431/full

Effect(s) tested

Specify type of analysis:  Whole brain  ROI-based  Both

Anatomical location(s)

Statistic type for inference  
(See [Eklund et al. 2016](#))

Correction

## Models & analysis

n/a | Involved in the study

Functional and/or effective connectivity

Graph analysis

Multivariate modeling or predictive analysis

Multivariate modeling and predictive analysis https://www.frontiersin.org/articles/10.3389/fgene.2020.00431/full  
Covariates are defined in the Methods section."/>