Genetic, transcriptomic, metabolic, and neuropsychiatric underpinnings of cortical functional gradients

Bin Wan [1,2,3], Yuankai He [4], Varun Warrier [4], Alexandra John [1,2],
Matthias Kirschner [3,5], Simon B. Eickhoff [2,6], Richard A.I. Bethlehem [7], Sofie L. Valk [1,2,6]

1. *Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany.*
2. *Institute of Neuroscience and Medicine (INM-7: Brain and Behavior), Research Center Jülich, Jülich, Germany.*
3. *Department of Psychiatry, University Hospitals of Genève, Thonex, Switzerland.*
4. *Department of Psychiatry, University of Cambridge, Cambridge, UK.*
5. *Synapsy Center for Neuroscience and Mental Health Research, University of Geneva, Geneva, Switzerland.*
6. *Institute of Systems Neuroscience, Heinrich Heine University Düsseldorf, Düsseldorf, Germany.*
7. *Department of Psychology, University of Cambridge, Cambridge, UK.*

Abstract

Functional gradients capture the organization of functional activity in the cerebral cortex, delineating transitions from sensory to higher-order association areas. While group-level gradient patterns are well-characterized, the biological mechanisms underlying individual variability remain poorly understood. Here, we integrate genetic, transcriptomic, and metabolic data across large-scale cohorts to investigate the biological basis of functional gradients and their relevance to neuropsychiatric conditions. Using twin-based heritability analyses and genome-wide association studies (GWAS) in over 30,000 individuals based on three datasets, we identified consistent heritability patterns and five genetic loci associated with gradient organization. These loci are linked to sixteen genes involved in metabolic pathways, with gene expression patterns spatially correlating with functional gradients. Furthermore, we observed significant associations between cardiovascular metabolic biomarkers and gradient architecture. Polygenic risk scores for neuropsychiatric conditions, including schizophrenia, bipolar disorder, and post-traumatic stress disorder, were significantly associated with gradient loadings, suggesting shared genetic influences on functional organization and mental health risk. Our findings highlight a complex interplay between genetic variation, metabolic processes, and brain function, offering new insights into the biological foundations of functional brain organization and its implications for neuropsychiatric vulnerability.

Main

Functional gradients capture transitions between sensory and association areas across the cerebral cortex, reflecting the inter-regional integration and segregation [1–3]. The sensory-association differentiation is a broad and major axis that differentiates sensory/motor and more abstract cognitive functions, supported by underlying neurobiological processes [4–8]. Moreover, gene expression in cortical regions follows a similar axis, which underpins the organized distribution of cell types in the cortex [8–10]. Generally, as stable features for the low-dimensional axis of functional connectivity, gradients represent a high-quality framework for the reproducibility and reliability of functional brain organization [11,12]. However, while gradients are robust at the group level, individual variability exists. For example, gradients vary over the lifespan, either in dominance of their axes, for example sensory-association and somatomotor-visual axes are shifted during childhood and adolescence [13,14] as well as integration and segregation of networks along gradient axes [15]. Regional alterations along these axes also are observed in various psychiatric conditions such as autism [16,17], depression [18], and schizophrenia [19,20]. This inter-individual variability raises critical questions about the biological mechanisms that shape these patterns. Recent studies suggest that this variability is heritable [21,22], yet the specific genetic architecture and its broader biological correlates have been poorly understood. As well, whether the genetic risks of those neuropsychiatric conditions contribute to functional gradients remains to be studied.

Moreover, to strengthen the conceptual link between these domains, we propose that metabolism may serve as an intermediary mechanism through which genetic factors exert their influence on brain function. Energy metabolism, which supports synaptic activity and neural plasticity, may modulate the development and maintenance of functional gradients. For example, brain structure and function are affected by a myriad of metabolic markers including cardiovascular factors [23–25], inflammatory processes [26–28], and lipids [29]. Also, energy metabolism relates to functional organization [30–32] in the brain. The current framework allows us to examine the interplay between genetic architecture, metabolic processes, and neuropsychiatric risk, offering a comprehensive view of the biological basis of cortical functional organization.

Therefore, our central research question is: How do genetic, transcriptomic, and metabolic factors contribute to cortical functional organization gradients, and how are these gradients

linked to neuropsychiatric conditions? To address this, we integrate multiple lines of evidence, including twin-based heritability analyses, genome-wide association studies (GWAS), transcriptomic mapping, and metabolic profiling, across diverse datasets spanning adolescence to late adulthood. We pursue four interrelated aims. First, we investigate the heritability of functional gradients across the lifespan using twin and SNP-based approaches. This allows us to quantify the genetic contribution to individual variability in gradient organization. Second, we identify specific genetic loci associated with functional gradients and examine their expression patterns in the cortex using transcriptomic data. This helps to elucidate the molecular pathways through which genetic variation may influence brain organization. Third, we assess the relationship between functional gradients and polygenic scores (PGSs) for neuropsychiatric conditions to understand how genetic predispositions for mental health disorders intersect with brain functional architecture. Finally, we explore how metabolic factors relate to gradient architecture, providing a potential biological bridge between genetic variation and functional brain organization.

Results

Heritability of regional gradients across three datasets

To quantify the genetic contribution to individual variability in functional gradients, we analyzed data from twin-based and SNP-based cohorts. The group-level gradient template generated by Human Connectome Project (HCP) forms a typical unimodal-transmodal pattern [3] (**Fig. 1a**). We aligned all individual gradients to the group-level HCP template because HCP provides high-quality multi-session long-time scans of resting-state functional magnetic resources imaging (fMRI) in young adults. To do so, we computed the individual functional connectome and then reduced the dimensionality of its affinity matrix using diffusion map embedding. Procrustes alignment was used to rotate all individuals to the template. All the analyses were conducted based on a multimodal parcellation (MMP) that contains 180 pairs of bilateral symmetric brain cortical regions [33]. The first three gradients accounted for 23.8%, 19.1%, and 15.8% of the total variance respectively.

Utilizing the genotype information from UKB and pedigree information in HCP and Queensland Twin Adolescent Brain (QTAB), we calculated the heritability (h^2) in each sample (**Fig. 1b**, detailed methods see **Methods**). Among G1, G2, and G3, dorsolateral prefrontal and posterior temporal regions showed a relatively moderate heritability ($h^2_{SNP} > 0.2$) and sensory areas showed a relatively low heritability ($h^2_{SNP} < 0.01$) in the UKB (**Fig. 1c**). It also showed similar results in twin-based HCP (**Fig. 1d**) and QTAB (**Fig. 1e**). Detailed heritability scores, standard errors, and *P*-values were shown in **Source Data**. For the G1, spatial correlation (Pearson) was $r = 0.209$ between UKB and HCP, $r = 0.293$ between UKB and QTAB, and $r = 0.330$ between HCP and QTAB, all $P_{variogram} < 0.03$. For the G2, the correlation was $r = 0.237$ between UKB and HCP, $r = 0.334$ between UKB and QTAB, and $r = 0.247$ between HCP and QTAB with all $P_{variogram} < 0.02$. For the G3, the correlation was $r = 0.220$ ($P_{variogram} = 0.013$) between UKB and HCP, $r = 0.138$ ($P_{variogram} = 0.159$) between UKB and QTAB, $r = 0.185$ ($P_{variogram} = 0.016$) between HCP and QTAB, **Fig. 1f**. It suggests consistent heritability maps using diverse datasets with different ages and genetic information.

We also observed correlations of heritability across gradients (**Fig. 1f**). In the UKB, spatial correlation (Pearson) was $r = 0.453$ between G1 and G2, $r = 0.543$ between G1 and G3, and $r =$

0.399 between G2 and G3 with all $P_{variogram} < 0.001$. In the HCP, spatial correlation (Pearson) was $r = 0.359$ between G1 and G2, $r = 0.335$ between G1 and G3, and $r = 0.217$ between G2 and G3 with all $P_{variogram} < 0.02$. In the QTAB, the correlation was $r = 0.502$ ($P_{variogram} < 0.001$) between G1 and G2, $r = 0.302$ ($P_{variogram} < 0.001$) between G1 and G3, and $r = 0.126$ ($P_{variogram} = 0.197$) between G2 and G3. Overall, our findings demonstrate consistent heritability patterns across datasets spanning different age groups, suggesting that genetic factors robustly influence gradient organization throughout the lifespan.

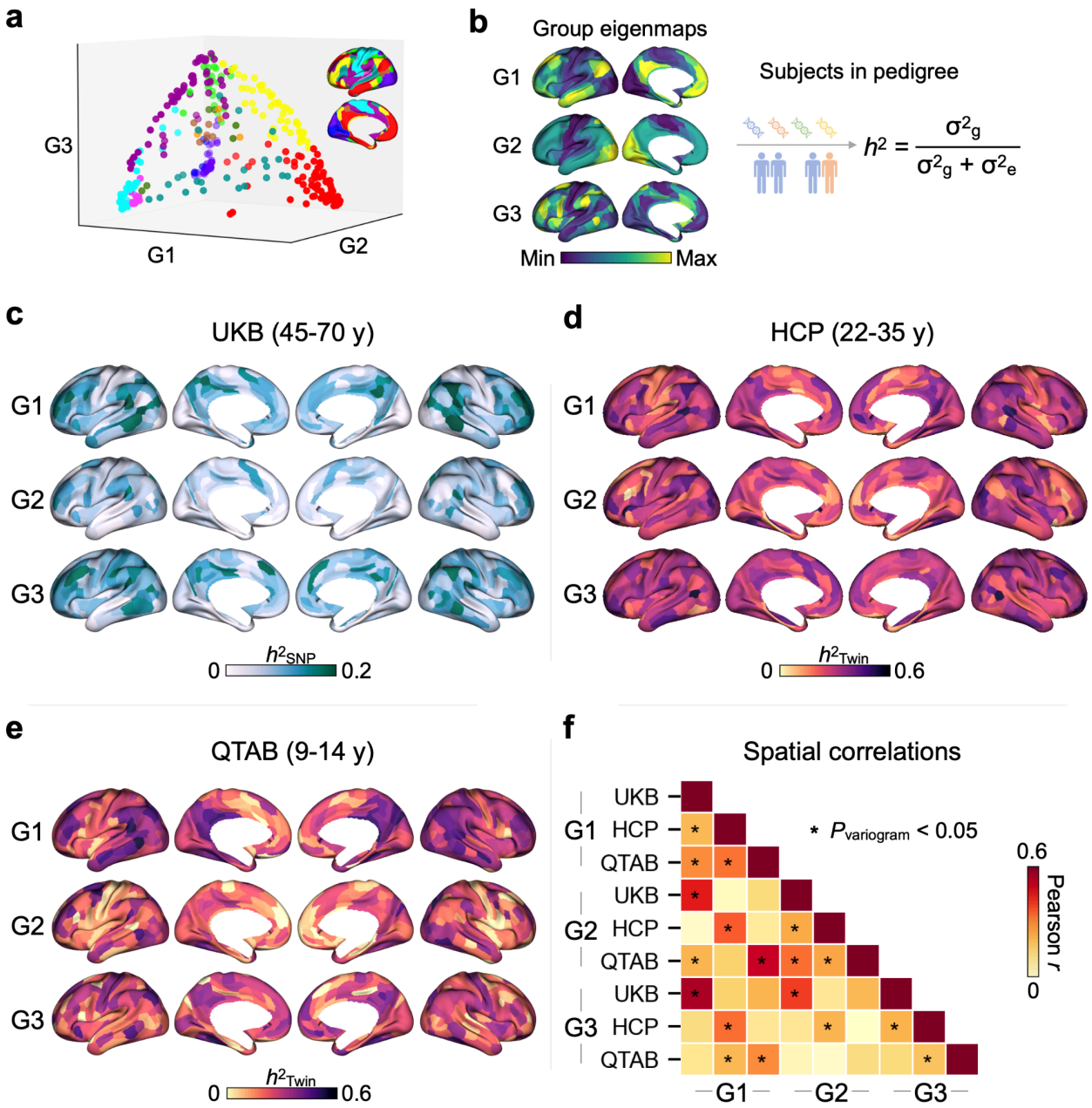

Fig. 1: Heritability (h^2) maps in different datasets. a. shows the low-dimensional space with colors coded by functional networks [34]. **b**. Shows eigenmaps for the three organization axes, calculated from the function connectome template [22] of the Human Connectome Project (HCP) [35]. All individuals were aligned to this group-level template. We used individual gradients and pedigree/genotype information to calculate the heritability (h^2) per region per gradient for the single-nucleotide polymorphisms (SNPs)-based UKB (**c**), twin-based HCP (**d**), and twin-based QTAB (**e**). **f**. Spatial correlation between every two heritability maps. Spatial autocorrelation is considered to permute the maps using the geodesic distance variogram and $P_{\text{variogram}}$ values are obtained based on 1000 permutations.

Genome-wide association studies (GWAS) of gradient similarities

In terms of finding a summary metric to capture individual variability in gradients and interpreting the gradients as a holistic description of brain organisation, we ran genome-wide association studies (GWAS) of the individual similarity scores instead of running node-wise GWAS. We calculated individual similarity (Pearson r) in the UKB with the high-quality HCP template to study individual variation in forming the whole sensory-association axis (**Fig. 2a**). A higher similarity score indicates a typical young-adult-like spatial pattern. The histogram of the individual distribution is shown in **Fig. 2b**. It showed a slightly right-skewed distribution (skewness = -1.19, -1.24, -0.74 in G1, G2, and G3). We also found that lower similarity scores were correlated with higher age for G1 (r = -0.151, $P < 0.001$), G2 (r = -0.142, $P < 0.001$), and G3 (r = -0.107, $P < 0.001$). To increase the robustness of our results, we also used the dissimilarity score (Euclidean distance, **Fig. 2c**). The histograms showed slightly left-skewed distributions (skewness = 0.46, 0.49, 0.75 in G1, G2, and G3). The heritability scores for Pearson r were G1: 0.17 (0.019), G2: 0.10 (0.018), and G3: 0.12 (0.019). The heritability scores for Euclidean distance were G1: 0.17 (0.019), G2: 0.11 (0.018), and G3: 0.09 (0.018). Standard errors are in brackets.

Following we conducted GWAS to uncover the related single-nucleotide polymorphisms (SNPs) using the GCTA [36] tool (https://yanglab.westlake.edu.cn/software/gcta/). We set up minor allele frequency (MAF) as 1% to find out the common variants, which include 9,116,883 remaining SNPs. We used the empirical significant threshold for P values (i.e., $P < 5 \times 10^{-8}$) and identified 3 loci containing 187 SNPs for G1 (Pearson r, **Fig. 2a**). Along the 3 loci (regional plot), the lead SNPs were rs1452628, rs3891783, and rs4880259. We observed 2 loci containing 155 SNPs for

G2 with lead SNPs: rs11187838 and rs10870314 (**Supplementary Fig. S1**). We observed 2 loci containing 42 SNPs for G3 with lead SNPs: rs1452628 and rs9645539 (**Supplementary Fig. S1**). Q-Q plots of the GWAS summary are shown in **Supplementary Fig. S2**. Detailed SNPs were reported in the **Source Data**.

For Euclidean distance, we identified 5 loci (3 repetitive loci in Pearson *r*) containing 408 SNPs for G1 (**Fig. 2b**). Along the two non-repetitive loci (regional plot), the lead SNPs were 6:96893385:C:CAG and rs59375526. We observed two loci containing 161 SNPs for G2 with lead SNPs: rs11187838 and rs796953214 (**Supplementary Fig. S3**). We observed 3 loci containing 51 SNPs for G3 with lead SNPs: rs3891783, rs9645539, and rs3781658 (**Supplementary Fig. S3**).

We performed various robustness checks. In addition to Pearson *r* and Euclidean distance, we performed GWAS for cosine similarity as an alternative measure of similarity. We found high consistency between Euclidean distance/Pearson *r* and cosine similarity (**Supplementary Fig. S4**). We also tested alternative gradient construction methods based on the top 50% of connections and full functional connectome and performed GWAS for Pearson *r*. These analyses showed similar eigenmaps and consistent significant SNPs centered on chromosome 10 (**Supplementary Fig. S5 and S6**). Nonetheless the top 10% has previously been shown to be more robust and offer better reproducibility and reliability [3,11,12,32].

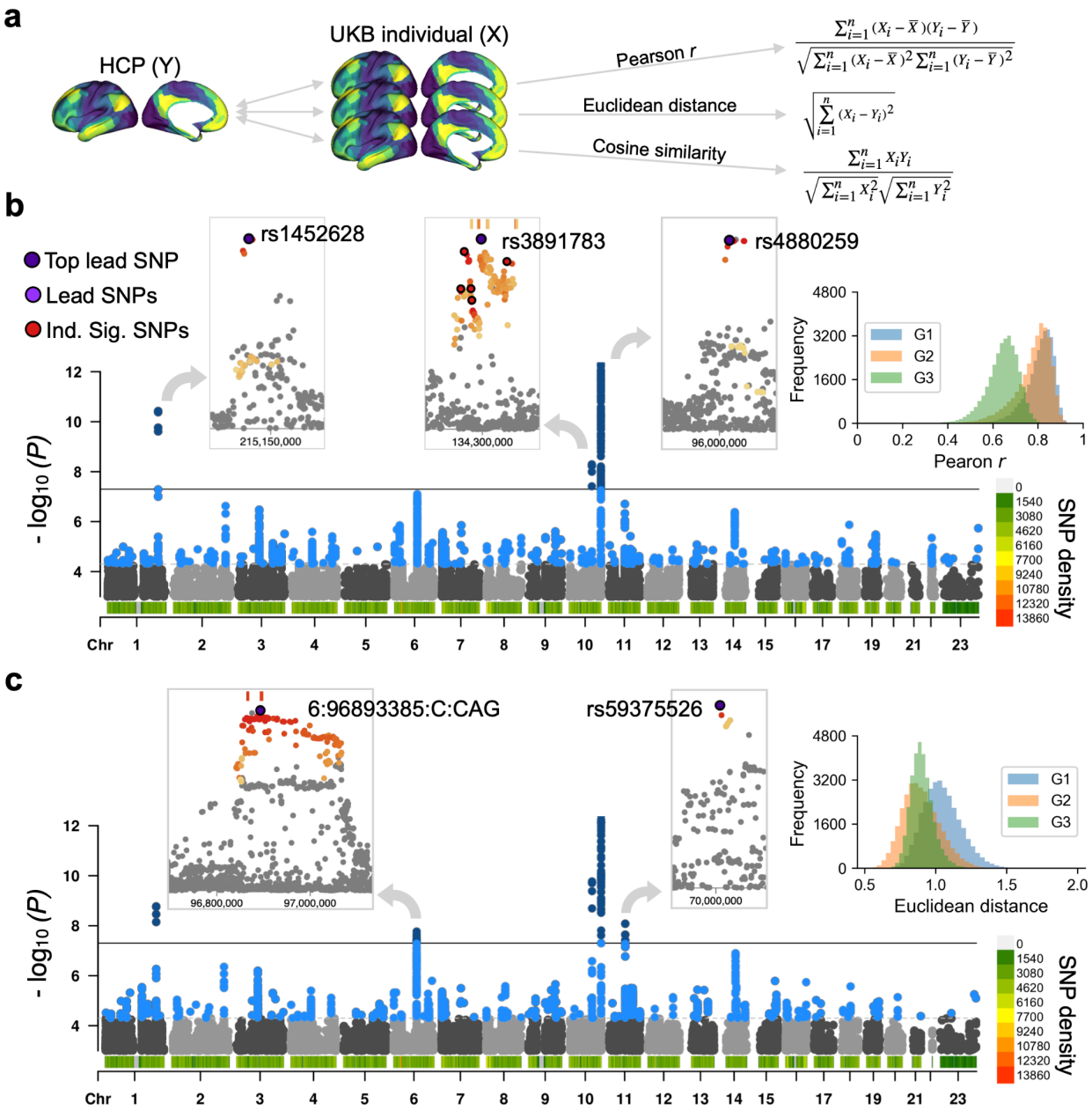

Fig. 2: Genome-wide association studies (GWAS) on the similarity score of Pearson *r* and Euclidean distance in G1. a. We associated the UKB individual gradient map with HCP group level gradient map to obtain (dis)similarity scores using Pearson *r*, Euclidean distance, and cosine similarity. Pearson *r* (**b**) uncovered three loci and Euclidean distance (**c**) uncovered two more loci in G1. Cosine similarity showed a similar Manhattan plot to Pearson *r*. So we didn't show it here but in **Supplementary Fig. S4**. Q-Q plots and GWAS summaries in G2 and G3 are shown in **Supplementary Figures** and **Source Data**.

Genetic correlates of functional gradients

To explore the wider spectrum of brain structural and disease correlates of functional gradients, we used linkage disequilibrium (LD) score [37,38] to estimate the SNPs-based genetic correlations between functional gradients and genetic profiles in 13 macro- and micro-structural brain phenotypes based on previous work [39] and the 11 neuropsychiatric GWAS summaries, detailed in **Methods**.

We identified stable genetic correlations for all three primary functional gradients and brain microstructural phenotypes derived from diffusion-weighted imaging (DWI) and reflecting microstructural integrity [39] based on Pearson *r* similarity metric (**Fig. 3a**). In particular, isotropic volume fraction (ISOVF, G1: r_g = -0.373, P_{FDR} < 0.001; G2: r_g = -0.307, P_{FDR} = 0.008; G3: r_g = -0.366, P_{FDR} = 0.003), mean diffusivity (G1: r_g = -0.294, P_{FDR} = 0.010; G2: r_g = -0.254, P_{FDR} = 0.049) and orientation dispersion index (G1: r_g = -0376, P_{FDR} = 0.004; G3: r_g = -0.508, P_{FDR} < 0.001) were all negatively correlated with the functional gradients. We found flipped genetic correlations for the Euclidean distance functional organization metric, where higher values indicate more similarity (**Fig. 3a**). Gradients were observed genetically inter-correlated amongst themselves (G1 and G2: r_g = 0.873; G2 and G3: r_g = 0.610; G3 and G1: r_g = 0.735, **Fig. 3b**). In addition, the genetic correlations were also intercorrelated across Euclidean distance (dissimilarity) and Pearson *r* (similarity) with r_g < -0.885 for G1 (**Fig. 3b**). It indicates different calculations for global features of gradients capture similar individual variance and global features are quite stable within subjects. Detailed scores are provided in **Source Data**.

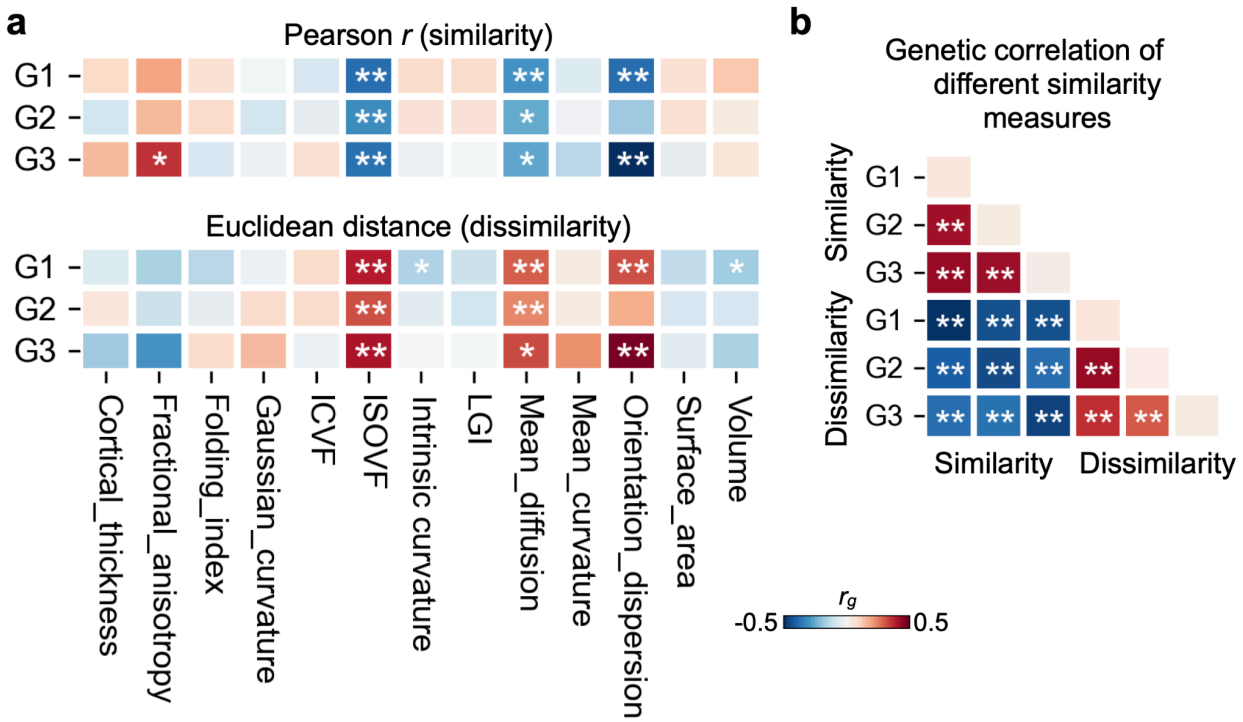

Fig. 3: SNPs-based genetic correlations (r_g) between functional gradients and 13 brain structural measures. a shows genetic correlations for Pearson *r* (similarity) and Euclidean distance (dissimilarity) features. **b** shows genetic correlations within gradients and between similarity measures. ICVF: intra-cellular volume fraction; ISOVF: isotropic volume fraction; LGI: local gyrification index; * and ** indicate $P < 0.05$ and $P_{FDR} < 0.05$.

Specific genes contribution: evidence from transcriptomics

Next we explored how the identified genetic loci relate to cortical gene expression, to uncover the molecular pathways through which genetic variation may influence functional gradient organization. First, we used the functional mapping and annotation (FUMA) [40] tool (https://fuma.ctglab.nl/) to map the nearest genes based on the GWAS results (maximum distance: 10kb) as well as MAGMA gene-based analysis [41]. We mapped 6 genes for the Pearson *r* including *FAM200A*, *NPC3L*, *PLCE1*, *GNA12, FHL5,* and *AMZ1* (**Supplementary Fig. S7**). Regarding Euclidean distance, in addition to the overlapping genes, we mapped 10 genes including *UFL1*, *FHL5*, *ANO1*, *FUT9*, *TBC1D12*, *HELLS*, *CYP2C9*, *CYP2C8*, *SLC35G1*, and *RP11-805J14.3* (**Supplementary Fig. S8**). Then, to understand how the transcriptomics of these genes spatially shape gradients, we extracted gene expression data from the Allen Human Brain

Atlas (AHBA) (**Fig. 4a**). AHBA contains approximately 500 brain locations per hemisphere across six donors with two of them having data from both hemispheres and four of them only having left hemisphere data [42]. When mapping the brain locations of probes onto the MMP atlas, there are 170 (of 180) regions available in the left hemisphere and 127 (of 180) regions available in the right hemisphere. The gene expression maps are offered by the ABAGEN [42–44] and cleaned using the ENIGMA toolbox [45]. We imputed the missing values using the variogram algorithm based on the central coordinates of the parcels and used imputed data for any downstream analyses. Among the 16 genes, we discovered 10 available in AHBA. The 10 gene expression maps were displayed in **Fig. 4b**.

To understand how the combination of the 10 genes shapes the gradients (R^2), we formed the ordinary least square (OLS) linear regression and bootstrapping analyses. We used (n) significant genes ($P < 0.05$) in OLS to compare the randomized bootstrapping (15,000 samples) combination using the same number (n) of genes shown as null distribution. We defined its specificity as the percentage of how many of the samples exceed the 'practical' R^2 (OLS with significant genes). R^2_{total} was 0.219, and $R^2_{practical}$ was 0.206 with six significant genes and specificity = 0.904 for G1 (**Fig. 4c**). Regarding G2, R^2_{total} was 0.189 and $R^2_{practical}$ was 0.177 with four significant genes with specificity = 0.978 (**Fig. 4d**). Regarding G3, R^2_{total} was 0.083 and $R^2_{practical}$ was 0.067 with three significant genes with specificity = 0.973 (**Fig. 4e**). It indicates that the GWAS-informed gene set reaches an excellent fitting at the transcriptome level.

As the above analyses were informed by GWAS hits, we also tested a purely data-driven approach to understand how a combination of transcriptomic maps could describe three gradients simultaneously based on spatial associations. We used a total of 15,634 genes as the input, then used elastic net (lasso ratio: 0.1-1) as the estimator to sparsify features, and finally detected which of them could be selected to predict three gradients (**Fig. 4f**). We detected 46 to 5 gene expression maps using lasso (L1) ratio from 0.1 to 1 (**Fig. 4g**). The weight of selected genes was displayed as a word cloud. Top 6 genes were *COL27A1*, *COL8A2*, *ADAD2*, *AMDHD1*, *CCDC158*, and *CD6*. The full list of the genes is provided in **Source Data**. Most of the parameters suggested an excellent specificity for G1, G2, and G3. We then used the same number as GWAS-informed of genes for each gradient (i.e., lasso ratio = 0.5, 10 genes, **Fig. 4h**) it

reached maximum specificity (8 genes for G1, 6 genes for G2, and 4 genes for G3), providing only limited discrimination.

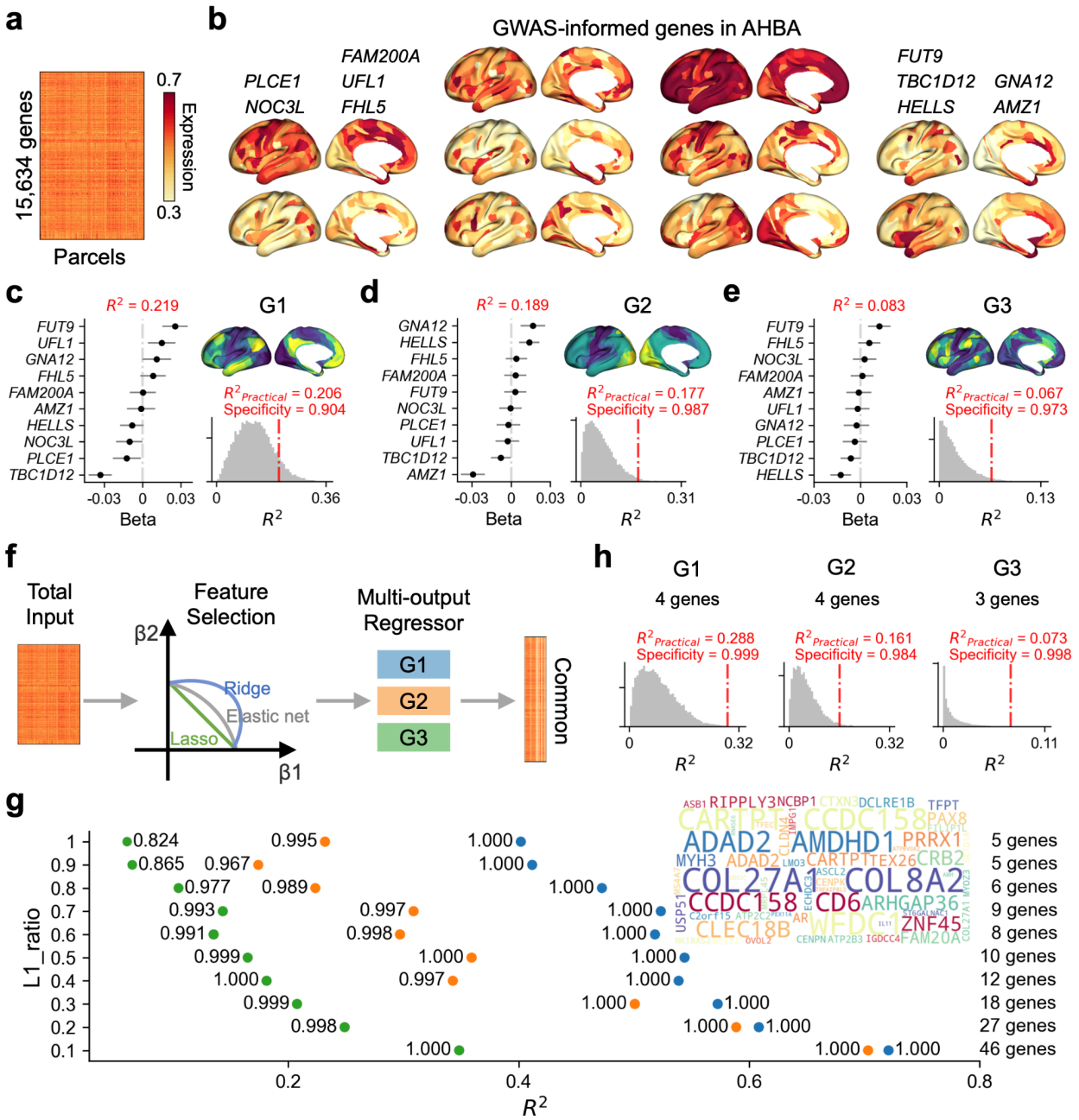

Fig. 4: Genetic validation of gradients using transcriptomic data. a shows gene expression extracted from Allen Human Brain Atlas (AHBA) with 15,634 genes by 360 parcels (297 available). **b** shows genes mapped from GWAS summary using the FUMA tool including positional mapping (maximum distance: 10kb) and MAGMA results. **c** displays the beta values of the six genes with a 95% confidence interval (CI) using an ordinary least squares (OLS) model to fit the G1 map as well as how many bootstrapping samples (15,000) below the practical model (four significant genes, $P < 0.05$) for G1. **d** OLS and bootstrapping (four significant genes, $P <$

0.05) for G2. **e**. OLS and bootstrapping (three significant genes, $P < 0.05$) for G3. **f** shows the data-driven approach using an elastic net to select seven common contributing genes for the three gradients. **g** tests different lasso (L1) ratios for detecting the common genes. The number next to each dot indicates the specificity of the selected gene set. The word cloud indicates the frequency of the selected genes (separate gradients in **Supplementary Fig. S9**). **h** shows bootstrapping for three gradients after the data-driven selection (L1_ratio = 0.5 for G1, G2, and G3).

Functional annotations for the genes

To understand how these genes are expressed in the brain across development, we used BrainSpan (https://www.brainspan.org/) to map the averaged gene expression level from 8 weeks post conception (PCW) to 40 years after birth (in total 29 time points). This data has been integrated into FUMA. We first predicted gradients-related gene expression using the model in **Fig. 4b**. For example, the OLS model returned the *t*-values for G1 was multiplied by normalized expression score, then the scores were summed up, finally, we looped it for the 29 development stages (**Fig. 5a**). We observed that G1-related genes were increasingly expressed after 13 years of age. G2- and G3-related genes were mostly expressed before birth. G2-related gene expression kept declining after birth whilst G3 remained stable in later life. The data points were smoothed by a Gaussian filter. We also performed this analysis for the AHBA-driven genes reported in the previous section. It showed that G1 expression remains stable but G2 increases expression of G2-related genes, and G3 expression initially decreases, but increases around pre-puberty (~age 9) (**Fig. 5a**). Specific genes (in grey tones) did not show much variance in expression over time.

Then, we performed an enrichment analysis to understand the biological annotations of related genes using ShinyGO 0.81 [46]. **Fig. 5b** showed 11 significant enriched terms for GWAS-informed genes. We used fold enrichment,a measure of the degree to which a particular term is overrepresented in the genes of interest. These biological processings for GWAS-informed genes were centered on metabolic pathways and extended to lipid-related biosynthesis, acid-related metabolism, and serotonergic synapse (**Fig. 5c**). When performing enrichment analysis for 10 AHBA-driven genes (L1_ratio = 0.5), we found there were two significant terms: histidine metabolism and protein digestion and absorption. There were no significant enriched terms using L1_ratio = 0.1, 0.2, 0.8, 0.9, and 1.

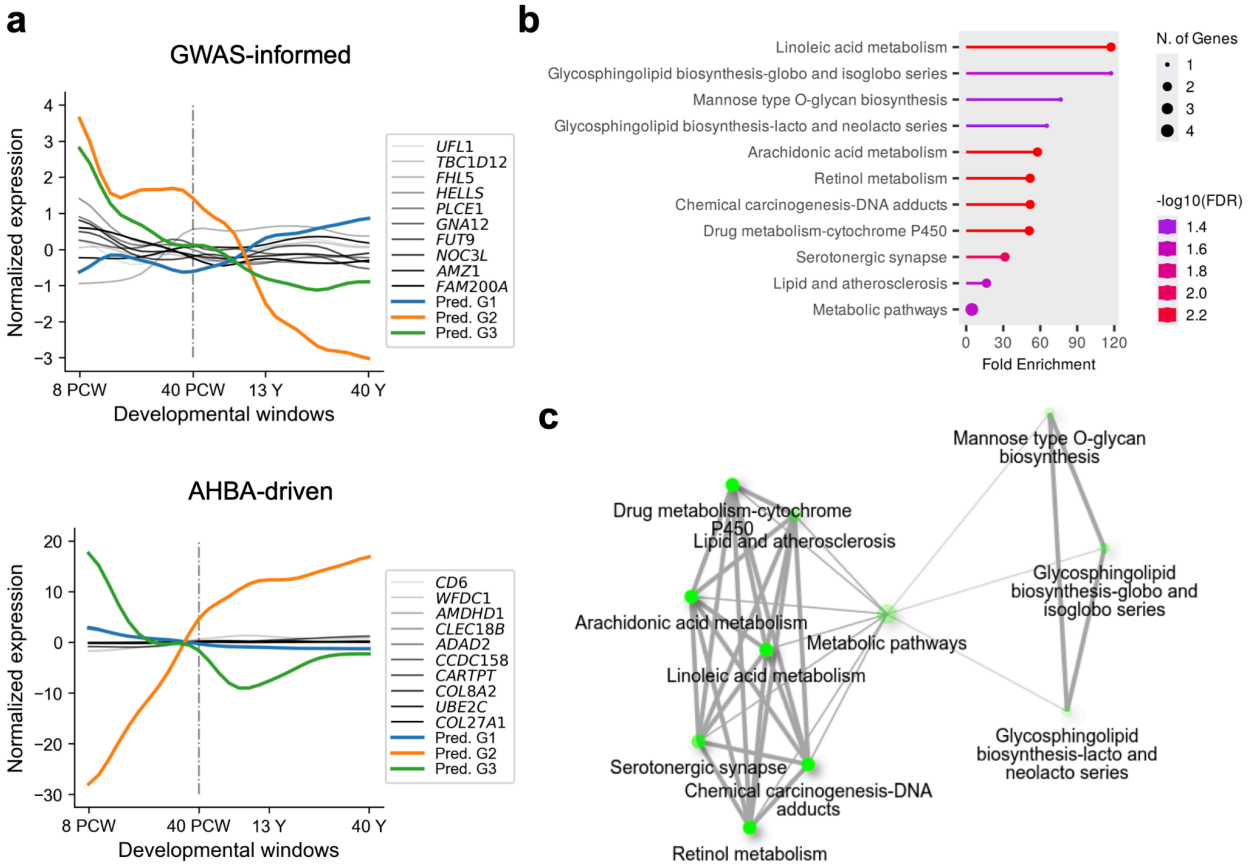

Fig. 5: Developmental and functional annotations for the mapped genes. a shows the cross-sectional trajectories of the GWAS-informed and AHBA-driven gene expressions. Post-conception weeks (PCW) and after birth years (Y) data from *post-mortem* brain. Gaussian normalized scores with averaging across the brain tissues are used here. **b** shows the enriched terms derived using ShinyGO version 0.81 [46] for the GWAS-informed genes. Fold enrichment is a measure of the degree to which a particular term is overrepresented in the genes of interest compared to a background set of genes. **c** displays enriched significant terms visualized as a network with an edge threshold setting of 0.3 for the GWAS-informed genes.

Association between regional gradients and neuropsychiatric PGSs

Following, we assessed the relationship between functional gradients and genetic risk for neuropsychiatric conditions, first using genetic correlations. We did not find any significant genetic correlation that survived multiple-testing correction (**Fig 6a**). Yet, alterations of regional gradients have been observed in multiple psychiatric diseases [16–19], possibly linked to indirect genetic effects. To further test this, we used multivariate analyses using polygenic risk scores (PGSs) for 11 disorders in relation to regional gradient loadings. Sex, age, and forty genetic

principal components, head motion, and data site were regressed out for individual regional gradient loadings.

Given the expected intercorrelation among the PGSs (**Fig 6b**), suggesting transdiagnostic signatures [47–49], we conducted a partial least squares (PLS) analysis with a machine learning pipeline to uncover the overall multivariate association between regional gradients (360 x 30,716) and PGSs (11 x 30,716 x 3 gradients). We split the subjects into training and test samples 100 times and performed PLS in the training sample and transformed the test sample to study the brain-PGS first dimension's correlation (**Fig. 6c**). We selected the model based on a balance between overfitting and underfitting (**Fig. 6d**). The first latent dimension suggested a positive correlation between PGS and gradient scores in the training sample (G1: $r = 0.067$, $P < 0.001$; G2: $r = 0.061$, $P < 0.001$; G1: $r = 0.076$, $P < 0.001$; **Fig. 6d**). When studying the brain loadings (**Fig. 6e**), we observed a pattern differentiating central regions from frontal/occipital regions for G1, a pattern highlighting occipital regions for G2, and a pattern highlighting central sulcus for G3. When summarizing into 12 functional networks [34], we found the averaged absolute loadings to follow the rank: auditory > language > secondary visual > orbito-affective > default mode > cingulo-opercular > frontoparietal > dorsal attention > posterior multimodal > somatomotor > ventral multimodal > primary visual networks. Regarding the PGS correlation with the brain patterns (**Fig. 6f**), the SCZ PGS showed a strong correlation with the latent brain pattern (r = 0.829, 0.871, and 0.797 for G1-3), followed by BP (r = 0.669, 0.826, and 0.512 for G1-3) and PTSD (r = 0.500, 0.592, and 0.758 for G1-3). ADHD showed reverse correlation patterns for G2 (r = -0.293). The regional correlation maps between gradients and PGS for each neuropsychiatric condition separately were shown in **Supplementary Fig. S10-12**. Here only few parcels survived the FDR correction (**Source Data**).

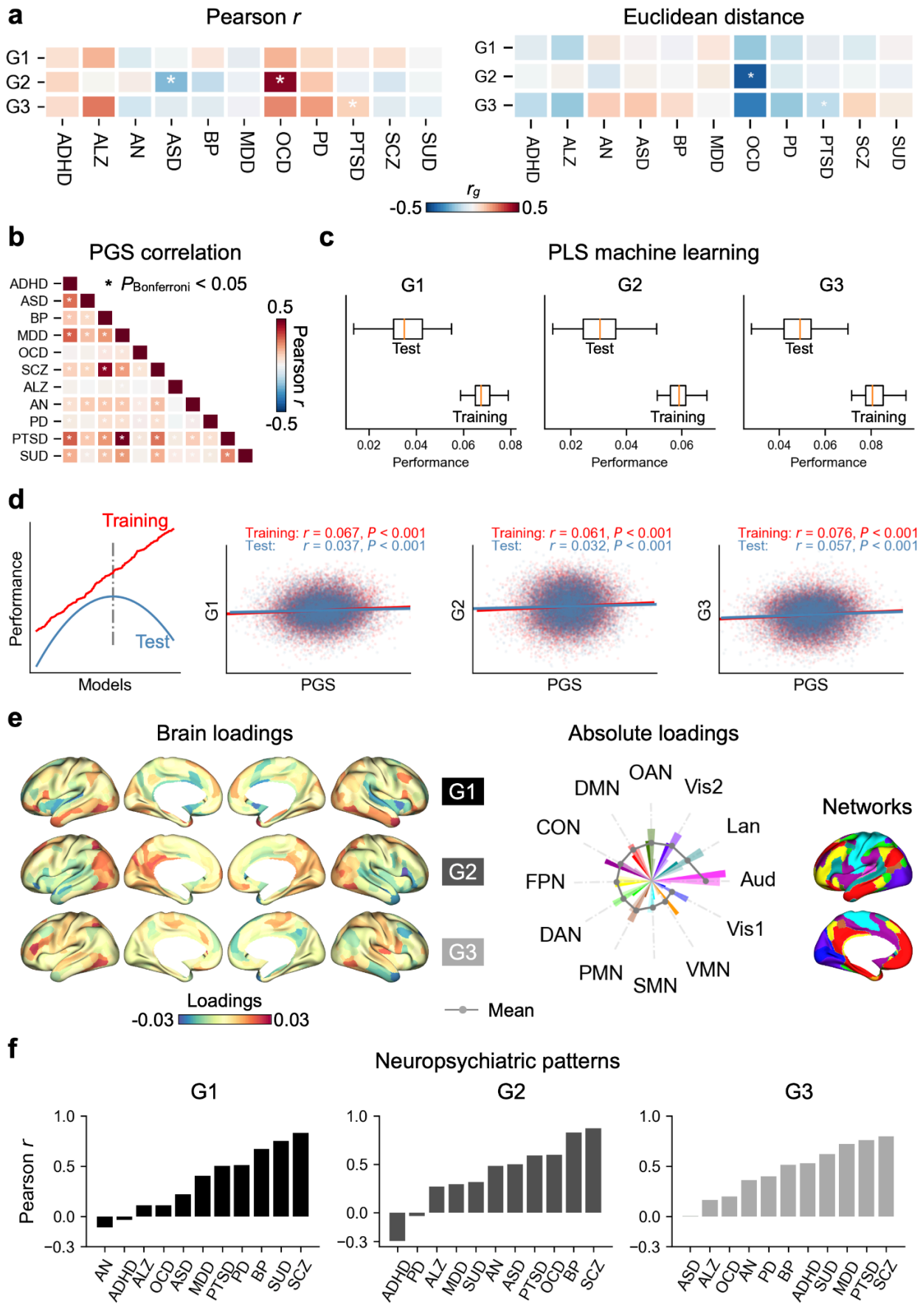
a
Pearson r
Euclidean distance
G1
G2
G3
ADHD
ALZ
AN
ASD
BP
MDD
OCD
PD
PTSD
SCZ
SUD
rg
-0.5
0.5
b
PGS correlation
* PBonferroni < 0.05
Pearson r
0.5
-0.5
c
PLS machine learning
G1
G2
G3
Test
Training
Performance
d
Training
Test
Performance
Models
Training: r = 0.067, P < 0.001
Test: r = 0.037, P < 0.001
Training: r = 0.061, P < 0.001
Test: r = 0.032, P < 0.001
Training: r = 0.076, P < 0.001
Test: r = 0.057, P < 0.001
PGS
e
Brain loadings
Absolute loadings
Networks
Loadings
-0.03
0.03
Mean
OAN
DMN
Vis2
CON
Lan
FPN
Aud
DAN
Vis1
PMN
SMN
VMN
f
Neuropsychiatric patterns
Pearson r
1.0
0.5
0.0
−0.3

Fig. 6: Association between regional gradients and polygenic scores for 11 neuropsychiatric conditions. a shows genetic correlations between 11 neuropsychiatric conditions and Pearson *r* and Euclidean distance features. * and ** indicate $P < 0.05$ and $P_{FDR} < 0.05$. **b** Correlation between polygenic risk scores (PGSs). The 11 psychiatric conditions include attention-deficit/hyperactivity disorder (ADHD), Alzheimer's disease (ALZ), anorexia nervosa (AN), autism spectrum disorder (ASD); bipolar disorder (BP), major depression disorder (MDD), obsessive-compulsive disorder (OCD), Parkinson's disease (PD), post-traumatic stress disorder (PTSD), schizophrenia (SCZ), substance use disorder (SUD). **c** Partial least squares (PLS) regression to associate multiple risks and regional gradients using machine learning. The sample was split into training and test subsamples for 100 times. Performance used Pearson *r*. **d** The performance of the 100 machine learning iterations revealed a most generalizable model with minimum distance between training and test and maximum training (i.e., min{(training - test)/train}). **e** PLS brain loadings for each gradient. We also summarized them into 12 networks [34] with the absolute loadings. The atlas-defined functional networks including primary visual (Vis1), secondary visual (Vis2), somatomotor (SMN), cingulo-opercular (CON), dorsal attention (DAN), language (LAN), frontoparietal (FPN), auditory network (AUD), default mode (DMN), posterior multimodal (PMN), ventral multimodal (VMN), orbito-affective (OAN). **f** Spatial correlation between brain loadings and regional *r*-map per neuropsychiatric condition and gradient in **Supplementary Fig. S10-12** and brain loading map in **d**.

Metabolic underpinnings using individual data

Together, our findings collectively provide a comprehensive view of the genetic and transcriptomic factors shaping cortical functional gradients and their potential implications for neuropsychiatric disorders. Finally, given the potential role of metabolism as central pathways in our enrichment analyses, we examined associations between 251 plasma metabolic biomarkers available in the UKB and functional gradients. Metabolic biomarkers were measured from randomly selected ethylenediaminetetraacetic acid (EDTA) plasma samples (aliquot 3) using a high-throughput nuclear magnetic resonance (NMR)-based metabolic biomarker profiling platform developed by Nightingale Health Ltd. We excluded outliers whose value exceeded the boundary of 3 standard deviations. Given the assessment lag (~10 years) between plasma and brain scans, we calculated the intraclass correlation coefficient (ICC) for each metabolite and used ICC > 0.7 to obtain the stable metabolites (**Fig. 7a**). We correlated gradient similarity (Person *r* and Euclidean distance) with the left 25 metabolites (**Fig. 7** and **Supplementary Fig. S13**). Detailed ICC and correlations at plasma T1 and T2 are shown in **Source Data**.

After regressing out covariates including age, sex, assessment lag, head motion, and data site, we observed that 17, 7, and 7 (P_{raw} < 0.05/25, Bonferroni correction) metabolites at plasma T1 were correlated with gradient similarity of G1, G2, and G3, respectively (**Fig. 7b**). We observed that 17, 1, and 13 metabolites at plasma T1 were correlated with gradient dissimilarity of G1, G2, and G3, respectively (**Fig. 7b**). Most of the metabolites were large (average diameter 12.1 nm) or very large (average diameter 14.3 nm) high-density lipoproteins (HDL), which were negatively and positively correlated with gradient similarity and dissimilarity. At plasma T2, we did not observe any significance due to the smaller sample size ($n_{\text{plasma T1}}$ = 17,027-18,995 and $n_{\text{plasma T2}}$ = 4,333-4,397). The effect size across 25 metabolites of two time points were correlated for similarity (G1: r = 0.882, P < 0.001; G2: r = 0.310, P = 0.131; G1: r = 0.985, P < 0.001) and dissimilarity (G1: r = 0.898, P < 0.001; G2: r = 0.124, P = 0.556; G1: r = 0.975, P < 0.001).

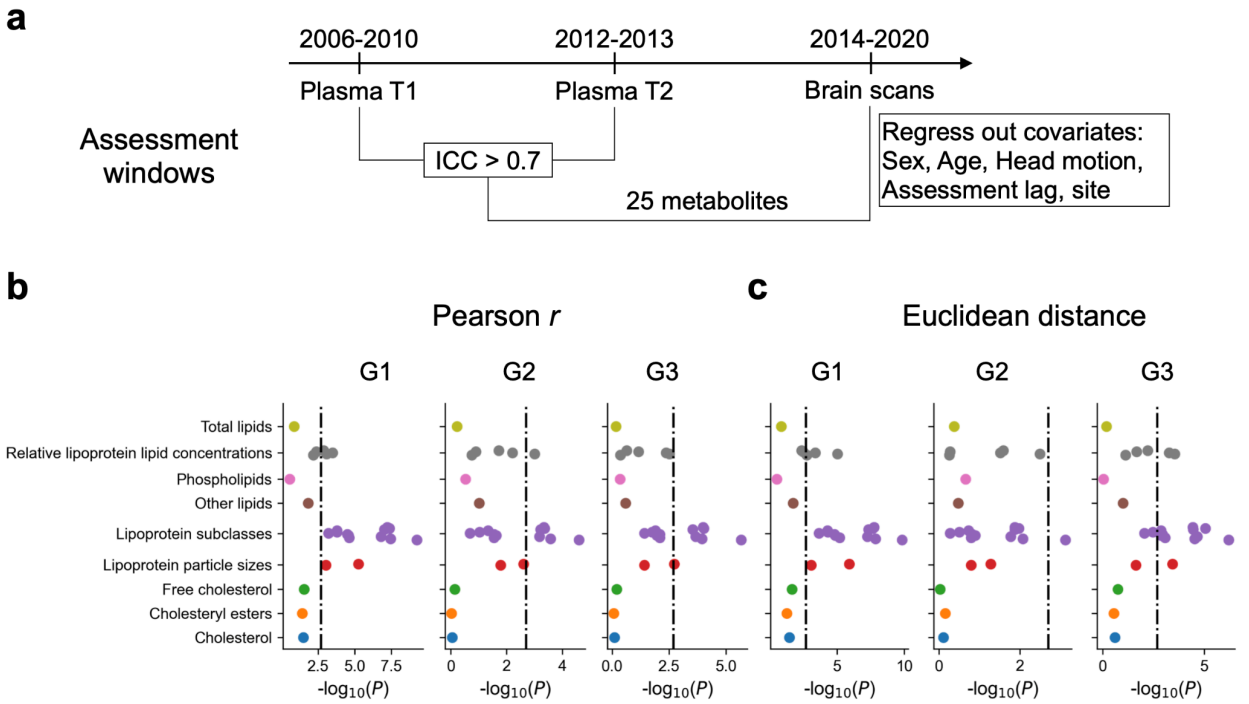

Fig. 7: Association between 24 metabolic biomarkers and gradient similarity in the UKB. After data cleaning (excluding outliers), 17,027 to 18,995 (Plasma T1) and 4,333 to 4,397 (Plasma T2) individuals who have both imaging and different biomarker data. **a** Intraclass correlation (ICC) analyses were used to filter the unstable metabolites due to assessment lag. We left 25 out of 251 (ICC > 0.7) biomarkers. *P*-values of Pearson correlation using plasma T1 and brain scan data were shown in the scatter plots with dashed lines (Bonferroni correction) for each gradient similarly (**b**) and dissimilarity (**c**). Model with regressing out covariates including age, sex, time gap between brain scanning and blood test, head motion, and data site. Detailed statistics of each biomarker are provided in **Source Data**.

Discussion

By integrating multimodal data from diverse cohorts, we demonstrated that genetic variation contributes to the organization of functional gradients across the lifespan. First, we examined the regional heritability of each gradient in three datasets of different age groups and found similar spatial maps across samples and gradients. This suggests that the proportion of individual variability in regional gradient loadings that can be explained by genetic variation is spatially consistent across the lifespan. We then used GWAS to investigate the common genetic variants associated with a global measure of gradient organization (i.e., the similarity of gradient to a functional organization template derived from the HCP sample), identifying 5 loci and 16 genes. There were strong genetic correlations between our data and GWAS from previous work based on diffusion-derived microstructure [39]; yet weak genetic correlations between our data and GWAS data from previous work studying neuropsychiatric conditions. This suggests that whilst there is a genetic overlap between function and microstructure, genetic variation supporting functional gradients only marginally relates to genes implicated in neuropsychiatric disorders. Yet, probing indirect genetic effects, by correlating the PGSs of 11 neuropsychiatric conditions and regional gradients, we uncovered the latent brain and neuropsychiatric patterns for three gradients separately, where SCZ, BP, and PTSD dominated the brain patterns. Finally, we observed that the identified genes linked to transcriptomic maps from a different dataset with expression related to each gradient varying across the lifespan and that overall the genes observed were largely involved in metabolic pathways. This observation was further corroborated using individual data indicating more normative gradient organization at the individual level was negatively correlated with molecular cardiovascular risk factors (i.e., large HDL). This supports the hypothesis that metabolic factors may be important to understanding brain function [50].

We observed a positive spatial correlation of regional heritability across UKB, HCP, and QTAB, even though the three datasets have different age ranges. This suggests that inter-individual variance in gradient loadings is continuously heritable across different age ranges. The degree of heritability was comparable for the two twin-based datasets and, as expected, lower for the SNP-based heritability. Longitudinal data are needed to understand how much the genetic

contributions, apart from the environmental contributions, change during development [51]. For example, the functional connectome in sensory areas matures earlier than in association areas during adolescence [52]. Theoretically, these developmental changes could be driven by similar gene sets, which may also be influenced by environmental factors.

The five genomic loci involved in shaping gradient axes are centered on chromosomes 1, 2, 10, and 11. Studying these loci, we found they contribute also to macrostructural features such as cortical thickness [53–57], surface area [54,57], and other features including morphological shape and sulcal depth [54,57]. Regarding brain functional features, rs38191783 and rs4880259 have been consistently observed associated with amplitude of low-frequency fluctuations in many regions [58] and independent components of amplitude [55]. As for non-brain features, they are mostly related to body weight and blood pressure [59–68]. The full GWAS summaries of the gradients were strongly related to DWI-derived microstructural measures but weakly related to macrostructural measures using GWAS summaries from a previous study [39]. This may indicate two things. Firstly, function is more closely reflective of microstructural integrity than macrostructure in the human brain; secondly, common genetic factors may push microstructural integrity (positive r_g for FA and negative r_g for mean diffusion, orientation dispersion, and ISOVF) and function gradients in similar directions. Of note, these DWI-derived measures are based on the mean score of the whole brain. Further regional patterns would benefit a detailed understanding of structure-function relationships, for example, to uncover the directionality of relationships (e.g. higher microstructural integrity may relate to both stronger between or within network integration, indirectly captured by gradient loadings). Phenotypic correlation studies show that microstructure and function correspond closely in sensory areas but diverge in association areas [21,69,70]. Possibly this divergence stems from different relationships between microstructure and short versus long-distance functional connectivity patterns or from environmentally induced plasticity impacting structure and function differently [21,71,72]. From a genetic perspective, Mendelian randomization (MR) could help us better understand their directionality in the future. For example, previous work using this approach has shown that surface area unidirectionally drives the folding of the cortex [39].

We found no overlapping genes between GWAS-informed and AHBA-driven analyses, yet both approaches achieved high specificity in modelling the spatial patterns of gradients. The lack of

concordance between the two imaging genetics approaches may be methodological. For example, there is a high possibility of type II error for the AHBA-driven method, i.e., the insignificant genes also contribute to the targeted brain maps, as we found that many GWAS-informed genes didn't show a spatial correlation with the gradients. Also, biological interpretations from AHBA association studies should be made with more caution. For example, while we found that metabolic pathways are informed by GWAS, almost no significant biological pathways are enriched by the genes observed using AHBA. Moreover, we found a large difference in the predicted G2 gene expression along development, i.e., reversed directions in GWAS-informed and AHBA-driven approaches. G2 is an early developmental axis [13,17,73] and we expected the related expression would decrease after adolescence, which fitted the GWAS-informed prediction. Yet this pattern completely reversed in AHBA-driven prediction. Further work, including rigorous data preprocessing and selection (Dear, 2024), may help bridge GWAS and expression-based observations.

Notably, the identified loci overlap with genes implicated in metabolic pathways, suggesting a biological bridge between genes, metabolism, and functional brain organization. This finding echoes emerging evidence on the role of energy metabolism in brain function [32,74] and highlights potential mechanisms through which genetic variation may impact cognitive and neuropsychiatric outcomes. Specifically, we found that large HDLs, cardiovascular risk factors, were negatively correlated with gradient similarity. It indicates that gradient similarity (to a young adult template) may reflect a health-related index. In previous work, we observed that gradients may form a connection between brain glucose metabolism and intrinsic functional organization [32], possibly also linked to notions of allostatic load and brain structure and function [75]. Importantly, synapse connection relies on not only glucose metabolism but also inhibition/excitation of neurotransmitters (e.g., GABA and dopamine) and microenvironment proteins (e.g., impacted by tau-protein).

Finally, clinical studies to date have focused mainly on case-control differences in G1, differentiating sensory from association regions. For example, higher gradient scores in default mode and lower gradient scores in the visual cortex for G1 have been found in young adults with early psychosis [20] and pediatric BP [76]. The reversed findings have been observed in ASD [16] and adults with BP [77]. In the current study, we did not find a direct genetic correlation, based on

GWAS markers, between neuropsychiatric disorders and functional gradients, yet when we correlated the regional gradient loadings with PGSs of 11 neuropsychiatric conditions we did observe multivariate associations between transdiagnostic genetic risk scores and gradient loadings. SCZ stood out among these PGSs, which is also the most heritable of disorders [78]. Our approach could identify links between both sensory and association areas, with strongest effects in auditory and language areas. The cortical patterns observed distinguish inferior frontal and middle temporal areas from sensory areas for G1, insular areas from visual cortices for G2, and default mode from frontoparietal networks for G3. Language-related areas might be relevant to transdiagnostic neuropsychiatry because of shared genetics between these regions and sociability [79], and their integrative role in cognition and emotion [80].

Our findings open several avenues for future research. Longitudinal studies are needed to explore how genetic and metabolic influences on functional gradients evolve across development and aging. Integrating multi-omics data, such as epigenomics and proteomics, could further elucidate the biological pathways linking genetics to brain function. Moreover, experimental studies investigating the causal role of metabolic pathways in shaping brain organization would provide more mechanistic insights. Understanding how metabolic health influences brain organization may also inform preventive strategies for cognitive decline and mental health disorders. Ultimately, our integrative framework lays the groundwork for personalized approaches in neuroscience, leveraging genetic and metabolic profiles to predict brain health and disease risk.

To close, while our multi-cohort approach enhances the generalizability of findings, several limitations warrant consideration. First, the reliance on predominantly European ancestry samples may limit the applicability of results to other populations. Future studies should prioritize diverse cohorts to assess the generalizability of genetic associations. Second, although we used robust statistical methods, GWAS inherently captures common variants, potentially overlooking rare genetic factors with significant effects. Additionally, the cross-sectional design limits our ability to infer causal relationships, particularly regarding the temporal dynamics of gradient development and metabolic changes.

Methods

Data used in the present study are all available to obtain through their data repositories. Ethics regarding each dataset have been approved by their review committee boards. UKB data used in this study under the application ID: 20904.

Datasets

UKB

UKB is a prospective cohort of approximately 500,000 individuals from the UK, who have SNP genotype data. Of these individuals, approximately 40,000 individuals have accomplished the MRI scanning when the current study commenced. Participants were excluded from the safety criteria such as metal implants, recent surgery, or conditions problematic for scanning such as hearing problems, breathing problems, or claustrophobia. The age range of the UKB sample was from 40 to more than 70 years. An overview of the data can be seen at the UKB website: https://www.ukbiobank.ac.uk.

HCP

HCP has 1,206 healthy young adults with high-quality four sessions of rs-fMRI scanning from the US [35]. It's a twin-based dataset with 298 monozygotic (MZ) and 188 dizygotic (DZ) twins as well as 720 singletons. We included individuals with a complete set of four fMRI scans that passed the HCP quality assessment [35,81]. We included 470 males with a mean age of 28.7 years (range: 22–37). Data can be openly obtained from the HCP DB website: https://db.humanconnectome.org.

QTAB

QTAB is a longitudinal dataset including 422 subjects aged from 9 to 14 years for the baseline and 304 subjects aged from 10 to 16 years for the follow-up [82]. Here, we used baseline data (48% female, mean age 11.3 ± 1.4 years) including 218 MZ, 194 DZ, 5 unpaired twins, and 5 without imaging data. Data can be openly obtained from OpenNeuro: https://openneuro.org/datasets/ds004146.

Genome-wide association

Before GWAS, genetic quality control and imputation of the UKB were done by the UKB team and described in detail elsewhere [83]. The following inclusion and exclusion criteria have been described in our previous publication [39]. Briefly, the inclusion criteria is self-identified white ethnicity and the exclusion criteria is outliners exceeding ±5 standard deviation from the means in the first two genetic principal components, which are also called predominantly European genetic ancestries. We further removed individuals whose genetic sex did not match their reported sex, or had excessive genetic heterozygosity, as provided by the UKB team. For the GWAS, we used all genotyped and imputed SNPs in the UKB that had a common variant with minor allele frequency (MAF) > 1% and were in Hardy–Weinberg equilibrium (HWE; $P < 1 \times 10^{-6}$) and, for imputed SNPs, had an imputation $r^2 > 0.4$. After quality control, we retained a maximum of 31,797 participants and approximately 130 million SNPs. We conducted three GWAS for the global features (described below in the functional gradients section) using fastGWA in the GCTA toolbox (version 1.94.1) [36]. For all GWAS, we included age, age2, sex, age × sex, age2 × sex, imaging center, first 40 genetic principal components, mean framewise displacement, maximum framewise displacement, and Euler Index as covariates. We identify associations using the common genetic variant significance level $P < 5 \times 10^{-8}$.

After obtaining GWAS results, we used the Functional Mapping and Annotation (FUMA) toolbox to map the genes. First, we found lead SNPs and defined genomic regions around each lead SNP with a window size of 10kb. Then we mapped SNPs to the nearest genes and extended the region around lead SNPs based on linkage disequilibrium (LD) patterns to capture SNPs in LD with lead SNPs.

Imaging preprocessing

UKB includes 34,830 subjects' imaging data. Details on data processing and acquisition can be found in the UKB Brain imaging documentation (https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf). Briefly, individual rs-fMRI data were motion corrected, intensity normalized, high-pass temporally filtered, and further denoised using the ICA-FIX pipeline, all implemented in FSL. Multimodal parcellation (MMP) [33] was warped to subject space based on the high-resolution T1-weighted anatomical image.

Individual warping parameters were applied to map the parcellation to the functional space following T1-rsfMRI alignment. Finally, FC was calculated by Fisher-z transforming the correlation matrix of MMP parcellated time series. After genetic and imaging quality control, 30,716 subjects from UKB were used in this study.

All HCP rs-fMRI data underwent HCP's minimal preprocessing [81] and were coregistered using a multimodal surface matching algorithm (MSMAll) (Robinson et al., 2014) to the HCP template 32k_LR surface space. The template consists of 32,492 total vertices per hemisphere (59,412 excluding the medial wall). We calculated the correlation matrix of MMP parcellated time series for two sessions (REST1_LR and REST1_RL). Then, we averaged the correlation matrix and Fisher-z transformed it to obtain the FC matrix. Finally, we analyzed 1,018 subjects.

For the QTAB, we followed the published pipeline [84]. Briefly, individual rs-fMRI data were motion and distortion corrected, high-pass temporally filtered, and then nuisance signal removed. The native surface was reconstructed from the T1w image using fastsurfer. Processed time series data then registered to the native cortical surface and further registered to fsLR_32k space with a 10mm Gaussian kernel smooth. We calculated the FC matrix of MMP parcellated time series for 422 subjects.

Functional gradients

Next, we employed the nonlinear dimensionality reduction technique [3] to generate the gradients of the FC matrix. We first made the FC matrix sparse with the top 10% of the connectivity remaining, then applied diffusion embedding on the affinity matrix, and finally extracted the first 10 eigenvectors with 360 eigenvalues for each eigenvector. All the steps were accomplished using the Python package BrainSpace [12]. Along these low-dimensional axes or gradients, cortical regions with similar functional connectivity profiles will be situated together and vice versa. The name of this approach, which belongs to the family of graph Laplacians, is derived from the equivalence of the Euclidean distance between points in the diffusion-embedded mapping [3,12,85]. It is controlled by a single parameter α, which controls the influence of the density of sampling points on the manifold (α=0, maximal influence; α=1, no influence). On the basis of the previous work [3], we followed recommendations and set α=0.5, a choice that retains the global relations

between data points in the embedded space and has been suggested to be relatively robust to noise in the covariance matrix.

Group-level gradients were constructed using the HCP averaged FC matrix and individual gradients were generated using Procrustes alignment from raw gradients to HCP group-level gradients (template), which makes individuals comparable. The global features are the similarity value (Pearson *r*) between the individual gradients after alignment and template gradients. The regional features are the regional eigenvalues after alignment for each individual. Here, we conducted the analyses for the first three gradients including sensory-association, somatomotor-visual, and default mode-multi demand.

Transcriptomics

We used BrainSpan and AHBA to investigate how the gradient gene expression in the brain develops over age and whether regional expression is correlated with gradient maps. BrainSpan contains transcriptomics of the 29 developing brain samples from 8 post-conceptual weeks (PCW) to 40 years old, data offered from FUMA [40,86] (https://fuma.ctglab.nl/). AHBA contains microarray expression data from approximately 20,000 genes [42]. A total of 3,702 brain tissue samples were collected from six neurotypical adults, including three Caucasian males, two African American males, and one Caucasian female. ENIGMA-toolbox [45] integrates the gene expression data from the abagen [44] and helps project the averaged gene expression data from Desikan–Killiany atlas to MMP, where 15,634 genes on 312 parcels are available to obtain the transcriptomic data. We imputed the missing values using the variogram algorithm based on the central coordinates of the parcels and used imputed data to do the transcriptomic analyses (full 360 parcels). The toolbox we used here is PyKirge (https://pykrige.readthedocs.io).

Heritability analyses

For the SNP-based heritability, we used GCTA–GREML (v1.94.1) [36,87] to calculate how much the phenotypic variance is explained by all the SNPs with the covariates mentioned in the GWAS. For the twin-based heritability, we used Sequential Oligogenic Linkage Analysis Routines (SOLAR, v9.0) [88]. Heritability (i.e. narrow-sense heritability h^2) represents the proportion of the phenotypic variance (σ^2_p) accounted for by the total additive genetic variance (σ^2_g), that is $h^2 = \sigma^2_g / \sigma^2_p$. Phenotypes exhibiting stronger covariances between genetically more similar individuals

(MZ twins) than between genetically less similar individuals (other pairs) have higher heritability. We added covariates to our twin-based heritability models including age, sex, age2, and age×sex. The spatial correlations between datasets across gradients were corrected using spatial autocorrelation [89], which generates surrogate maps using the variogram function on the geometric distance matrix. We generated 1000 surrogate maps as null models and tested how many null models the practical model exceeded to obtain the $P_{\text{variogram}}$ value.

Metabolic biomarkers

We used 251 plasma metabolic biomarkers available in the UKB and functional gradients. Metabolic biomarkers were measured from randomly selected ethylenediaminetetraacetic acid (EDTA) plasma samples (aliquot 3) using a high-throughput nuclear magnetic resonance (NMR)-based metabolic biomarker profiling platform developed by Nightingale Health Ltd. We excluded outliers whose value exceeded the boundary of 3 standard deviations (remaining 7,027-18,995 individuals for plasma T1 and 4,333-4,397 individuals for plasma T2) and the 251 metabolic biomarkers can be categorized into 18 classes including amino acids, apolipoproteins, cholesterol, cholesteryl esters, fatty acids, fluid balance, free cholesterol, glycolysis related metabolites, inflammation, ketone bodies, lipoprotein particle concentrations, lipoprotein particle sizes, lipoprotein subclasses, triglycerides, total lipids, relative lipoprotein lipid concentrations, phospholipids, and other lipids. These 251 metabolic biomarkers consist of 81 ratios and 170 direct measures. However, due to the assessment lag between plasma tests and brain scans, we only used stable metabolites in our analyses. Here, we filtered the metalites, whose ICC was below 0.7, and 25 metabolites were finally analyzed. Covaraites in the metabolic analyses included age (age at scanning, age^2 at scanning, and assessment lag between age at scanning and blood test), sex (biological sex, sex*age, and sex*age^2), scanning site, and head motion (Euler angles for left plus right hemispheres, mean framewise displacement, and maximum framewise displacement). We regressed these covariates for gradient similarity score and then conducted the correlation analyses. We performed these analyses for both plasma T1 and plasma T2 to make sure that the correlation directions could be stable. However, due to the smaller sample size of T2, we only showed the results of T1 in the figure. All results are provided in **Source Data**.

PGS for neuropsychiatry conditions

We calculated the PGS for 11 psychiatric conditions including ADHD [90], ALZ [91], AN [92], ASD [93], BP [94], MDD [95], OCD [96], PD, PTSD [53], SCZ [97], and SUD [98], using summary statistics in independent cohorts from the UKB, filtered for common variants also present in the UKB cohort in our GWAS. PGS weights were derived for each chromosome separately, using PRS by continuous shrinkage (PRS-cs) [99], based on the LD reference panel for predominantly European ancestry from the 1000 Genomes Project. Each individual was scored for each chromosome and the total PGS across 23 chromosomes was standardized to z-scores (zero mean, unit variance). We used published GWAS summary statistics for 13 imaging-derived phenotypes of global brain macro- and microstructure [39]. For these phenotypes, we estimated their genetic correlation with the functional gradients from the present study using LD score regression (LDSC) [37,38]. The same method was also applied to estimate the genetic correlation amongst the three functional gradients, and their heritability.

PLS machine learning

We initially conducted PLS analyses with a machine learning pipeline to uncover the overall multivariate association between regional gradients (360 x 30,716) and PGSs (11 x 30,716 x 3 gradients). We regressed out covariates for regional gradients here including age (age at scanning, age^2 at scanning, and time gap between age at scanning and blood test), sex (biological sex, sex*age, and sex*age^2), scanning site, and head motion (Euler angles for left plus right hemispheres, mean framewise displacement, and maximum framewise displacement), and first forty genetic components. Regarding the machine pipeline, we split the subjects into training and test samples 100 times and performed PLS in the training sample and transformed the test sample to look at the brain-PGS first dimension's correlation. Then, we selected the best model as the output, which maximally utilizes performance in the test sample and reduces the overfitting outcome, i.e., minimum distance between training and test and maximum training (min{(training - test)/train}. To capture the three gradients at the same time, we used the minimal averaged performance across three gradients as the final model. We focused on the first component of neuropsychiatric conditions and brain features for three gradients separately. After obtaining the latent brain loadings for each gradient, we calculated the Pearson *r* between brain loadings and

each neuropsychiatric PGS effect size map. Finally, we got how similar each neuropsychiatric PGS was to the latent brain loadings.

Data availability

The raw data we used in the current study are available through their website, which has been mentioned in the Datasets section. Source data is also provided.

Code availability

All scripts including genetic analyses, brain features computation, and statistical models can be found at a GitHub repository: https://github.com/wanb-psych/GWAS_FuncGradients.

Acknowledgments

BW is supported by the International Max Planck Research School on Neuroscience of Communication: Function, Structure, and Plasticity (IMPRS NeuroCom), the Federation of European Neuroscience Societies (FENS), and the IBRO Pan-Europe Regional Committee (IBRO-PERC) Exchange Fellowships Programme. SLV is funded by the Otto Hahn Award and Lise Meitner Excellence Program at the Max Planck Society and Helmholtz International BigBrain Analytics and Learning Laboratory (HIBALL), supported by the Helmholtz Association's Initiative and Networking Fund and the Healthy Brains, Healthy Lives initiative at McGill University.

Competing interests

The authors declare no conflict of interest.

Authors' contribution

Conceptualization: B.W., S.L.V. Methodology: B.W., W.V, R.A.I.B, S.L.V. Formal analysis: B.W., Y.H. Writing - Original Draft: B.W., S.L.V. Writing - Review & Editing: B.W., H.Y., W.V., A.J., S.B.E, R.A.I.B., S.L.V. Visualization: B.W. Data curation: B.W., R.A.I.B, S.L.V. Project administration: B.W. Funding acquisition: B.W., S.L.V. Supervision: S.L.V.

Reference

1. Bernhardt, B. C., Smallwood, J., Keilholz, S. & Margulies, D. S. Gradients in brain organization. *NeuroImage* **251**, 118987 (2022).
2. Huntenburg, J. M., Bazin, P.-L. & Margulies, D. S. Large-Scale Gradients in Human Cortical Organization. *Trends Cogn. Sci.* **22**, 21–31 (2018).
3. Margulies, D. S. *et al.* Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl. Acad. Sci.* **113**, 12574–12579 (2016).
4. Ito, T. & Murray, J. D. Multitask representations in the human cortex transform along a sensory-to-motor hierarchy. *Nat. Neurosci.* **26**, 306–315 (2023).
5. Markello, R. D. *et al.* neuromaps: structural and functional interpretation of brain maps. *Nat. Methods* **19**, 1472–1479 (2022).
6. Mesulam, M. M. From sensation to cognition. *Brain J. Neurol.* **121 (Pt 6)**, 1013–1052 (1998).
7. Sydnor, V. J. *et al.* Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for psychopathology. *Neuron* **109**, 2820–2846 (2021).
8. Zhang, X.-H. *et al.* The cell-type underpinnings of the human functional cortical connectome. *Nat. Neurosci.* 1–11 (2024) doi:10.1038/s41593-024-01812-2.
9. Dear, R. *et al.* Cortical gene expression architecture links healthy neurodevelopment to the imaging, transcriptomics and genetics of autism and schizophrenia. *Nat. Neurosci.* **27**, 1075–1086 (2024).
10. Hansen, J. Y. *et al.* Mapping gene transcription and neurocognition across human neocortex. *Nat. Hum. Behav.* **5**, 1240–1250 (2021).
11. Hong, S.-J. *et al.* Toward a connectivity gradient-based framework for reproducible biomarker discovery. *NeuroImage* **223**, 117322 (2020).
12. Vos de Wael, R. *et al.* BrainSpace: a toolbox for the analysis of macroscale gradients in neuroimaging and connectomics datasets. *Commun. Biol.* **3**, 1–10 (2020).
13. Dong, H.-M., Margulies, D. S., Zuo, X.-N. & Holmes, A. J. Shifting gradients of macroscale cortical organization mark the transition from childhood to adolescence. *Proc. Natl. Acad. Sci.* **118**, (2021).
14. Xia, Y. *et al.* Development of functional connectome gradients during childhood and adolescence. *Sci. Bull.* (2022) doi:10.1016/j.scib.2022.01.002.
15. Bethlehem, R. A. I. *et al.* Dispersion of functional gradients across the adult lifespan. *NeuroImage* **222**, 117299 (2020).
16. Hong, S.-J. *et al.* Atypical functional connectome hierarchy in autism. *Nat. Commun.* **10**, 1022 (2019).
17. Wan, B. *et al.* Diverging asymmetry of intrinsic functional organization in autism. *Mol. Psychiatry* 1–11 (2023) doi:10.1038/s41380-023-02220-x.
18. Xia, M. *et al.* Connectome gradient dysfunction in major depression and its association with gene expression profiles and treatment outcomes. *Mol. Psychiatry* **27**, 1384–1393 (2022).
19. Dong, D. *et al.* Compressed sensorimotor-to-transmodal hierarchical organization in schizophrenia. *Psychol. Med.* 1–14 (2021) doi:10.1017/S0033291721002129.
20. Holmes, A. *et al.* Disruptions of Hierarchical Cortical Organization in Early Psychosis and Schizophrenia. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **8**, 1240–1250 (2023).
21. Valk, S. L. *et al.* Genetic and phylogenetic uncoupling of structure and function in human transmodal cortex. *Nat. Commun.* **13**, 2341 (2022).
22. Wan, B. *et al.* Heritability and cross-species comparisons of human cortical functional organization

asymmetry. *eLife* **11**, e77215 (2022).
23. Kivimäki, M. & Steptoe, A. Effects of stress on the development and progression of cardiovascular disease. *Nat. Rev. Cardiol.* **15**, 215–229 (2018).
24. Schaare, H. L. *et al.* Associations between mental health, blood pressure and the development of hypertension. *Nat. Commun.* **14**, 1953 (2023).
25. Song, R. *et al.* Associations Between Cardiovascular Risk, Structural Brain Changes, and Cognitive Decline. *J. Am. Coll. Cardiol.* **75**, 2525–2534 (2020).
26. Lizano, P. *et al.* Multivariate relationships between peripheral inflammatory marker subtypes and cognitive and brain structural measures in psychosis. *Mol. Psychiatry* **26**, 3430–3443 (2021).
27. Schrepf, A. *et al.* A multi-modal MRI study of the central response to inflammation in rheumatoid arthritis. *Nat. Commun.* **9**, 2243 (2018).
28. Williams, J. A. *et al.* Inflammation and Brain Structure in Schizophrenia and Other Neuropsychiatric Disorders: A Mendelian Randomization Study. *JAMA Psychiatry* **79**, 498–507 (2022).
29. Hussain, G. *et al.* Lipids as biomarkers of brain disorders. *Crit. Rev. Food Sci. Nutr.* **60**, 351–374 (2020).
30. Palombit, A. *et al.* Variability of regional glucose metabolism and the topology of functional networks in the human brain. *NeuroImage* **257**, 119280 (2022).
31. Riedl, V. *et al.* Local Activity Determines Functional Connectivity in the Resting Human Brain: A Simultaneous FDG-PET/fMRI Study. *J. Neurosci.* **34**, 6260–6266 (2014).
32. Wan, B., Riedl, V., Castrillon, G., Kirschner, M. & Valk, S. L. Bridging Glucose Metabolism and Intrinsic Functional Organization of the Human Cortex. 2024.09.26.615152 Preprint at https://doi.org/10.1101/2024.09.26.615152 (2024).
33. Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).
34. Ji, J. L. *et al.* Mapping the human brain's cortical-subcortical functional network organization. *NeuroImage* **185**, 35–57 (2019).
35. Van Essen, D. C. *et al.* The WU-Minn Human Connectome Project: An overview. *NeuroImage* **80**, 62–79 (2013).
36. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
37. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
38. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
39. Warrier, V. *et al.* Genetic insights into human cortical organization and development through genome-wide analyses of 2,347 neuroimaging phenotypes. *Nat. Genet.* **55**, 1483–1493 (2023).
40. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
41. Leeuw, C. A. de, Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Comput. Biol.* **11**, e1004219 (2015).
42. Hawrylycz, M. J. *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).
43. ArnatkevičIūtė, A., Fulcher, B. D. & Fornito, A. A practical guide to linking brain-wide gene

expression and neuroimaging data. *Neuroimage* **189**, 353–367 (2019).
44. Markello, R. D. *et al.* Standardizing workflows in imaging transcriptomics with the abagen toolbox. *eLife* **10**, e72129 (2021).
45. Larivière, S. *et al.* The ENIGMA Toolbox: multiscale neural contextualization of multisite neuroimaging datasets. *Nat. Methods* **18**, 698–700 (2021).
46. Ge, S. X., Jung, D. & Yao, R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**, 2628–2629 (2020).
47. Cardno, A. G. & Owen, M. J. Genetic Relationships Between Schizophrenia, Bipolar Disorder, and Schizoaffective Disorder. *Schizophr. Bull.* **40**, 504–515 (2014).
48. Ghirardi, L. *et al.* The familial co-aggregation of ASD and ADHD: a register-based cohort study. *Mol. Psychiatry* **23**, 257–262 (2018).
49. Riglin, L. *et al.* ADHD and depression: investigating a causal explanation. *Psychol. Med.* **51**, 1890–1897 (2021).
50. Ng, B. *et al.* Integration across biophysical scales identifies molecular and cellular correlates of person-to-person variability in human brain connectivity. *Nat. Neurosci.* **27**, 2240–2252 (2024).
51. Schmitt, J. E. *et al.* The dynamic role of genetics on cortical patterning during childhood and adolescence. *Proc. Natl. Acad. Sci.* **111**, 6774–6779 (2014).
52. Dorfschmidt, L. *et al.* Sexually divergent development of depression-related brain networks during healthy human adolescence. *Sci. Adv.* **8**, eabm7825 (2022).
53. Nievergelt, C. M. *et al.* International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci. *Nat. Commun.* **10**, 4558 (2019).
54. Shadrin, A. A. *et al.* Vertex-wise multivariate genome-wide association study identifies 780 unique genetic loci associated with cortical morphology. *NeuroImage* **244**, 118603 (2021).
55. Smith, S. M. *et al.* An expanded set of genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nat. Neurosci.* **24**, 737–745 (2021).
56. van der Meer, D. *et al.* Understanding the genetic determinants of the brain with MOSTest. *Nat. Commun.* **11**, 3512 (2020).
57. van der Meer, D. *et al.* The genetic architecture of human cortical folding. *Sci. Adv.* **7**, eabj9446 (2021).
58. Zhao, B. *et al.* Common variants contribute to intrinsic human brain functional networks. *Nat. Genet.* **54**, 508–517 (2022).
59. German, C. A., Sinsheimer, J. S., Klimentidis, Y. C., Zhou, H. & Zhou, J. J. Ordered multinomial regression for genetic association analysis of ordinal phenotypes at Biobank scale. *Genet. Epidemiol.* **44**, 248–260 (2020).
60. Giri, A. *et al.* Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nat. Genet.* **51**, 51–62 (2019).
61. Guo, Y. *et al.* A genome-wide cross-phenotype meta-analysis of the association of blood pressure with migraine. *Nat. Commun.* **11**, 3368 (2020).
62. Helgeland, Ø. *et al.* Characterization of the genetic architecture of infant and early childhood body mass index. *Nat. Metab.* **4**, 344–358 (2022).
63. Hübel, C. *et al.* Genomics of body fat percentage may contribute to sex bias in anorexia nervosa. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet. Off. Publ. Int. Soc. Psychiatr. Genet.* **180**, 428–438 (2019).
64. Kichaev, G. *et al.* Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *Am. J.*

Hum. Genet. **104**, 65–75 (2019).
65. Plotnikov, D., Williams, C. & Guggenheim, J. A. Association between birth weight and refractive error in adulthood: a Mendelian randomisation study. *Br. J. Ophthalmol.* **104**, 214–219 (2020).
66. Plotnikov, D. *et al.* High Blood Pressure and Intraocular Pressure: A Mendelian Randomization Study. *Invest. Ophthalmol. Vis. Sci.* **63**, 29 (2022).
67. Sakaue, S. *et al.* A cross-population atlas of genetic associations for 220 human phenotypes. *Nat. Genet.* **53**, 1415–1424 (2021).
68. Warrington, N. M. *et al.* Maternal and fetal genetic effects on birth weight and their relevance to cardio-metabolic risk factors. *Nat. Genet.* **51**, 804–814 (2019).
69. Bignardi, G. *et al.* Distinct genetic underpinnings of inter-individual differences in the sensorimotor-association axis of cortical organisation. 2023.07.13.548817 Preprint at https://doi.org/10.1101/2023.07.13.548817 (2024).
70. Vázquez-Rodríguez, B. *et al.* Gradients of structure–function tethering across neocortex. *Proc. Natl. Acad. Sci.* **116**, 21219–21227 (2019).
71. Hettwer, M. D. *et al.* Longitudinal variation in resilient psychosocial functioning is associated with ongoing cortical myelination and functional reorganization during adolescence. *Nat. Commun.* **15**, 6283 (2024).
72. Paquola, C. *et al.* Shifts in myeloarchitecture characterise adolescent development of cortical gradients. *eLife* **8**, e50482 (2019).
73. Larivière, S. *et al.* Multiscale Structure–Function Gradients in the Neonatal Connectome. *Cereb. Cortex* **30**, 47–58 (2020).
74. Castrillon, G. *et al.* An energy costly architecture of neuromodulators for human brain evolution and cognition. *Sci. Adv.* **9**, eadi7632 (2023).
75. Chiappelli, J. *et al.* Allostatic Load and Reduced Cortical Thickness in Schizophrenia. *Psychoneuroendocrinology* **77**, 105–111 (2017).
76. Lei, W. *et al.* The disruption of functional connectome gradient revealing networks imbalance in pediatric bipolar disorder. *J. Psychiatr. Res.* **164**, 72–79 (2023).
77. Liu, X. *et al.* Functional Reorganization across Three Episodes in Bipolar Disorder. 2024.09.03.24312957 Preprint at https://doi.org/10.1101/2024.09.03.24312957 (2024).
78. Pettersson, E. *et al.* Genetic influences on eight psychiatric disorders based on family data of 4 408 646 full and half-siblings, and genetic data of 333 748 cases and controls. *Psychol. Med.* **49**, 1166–1173 (2019).
79. Fanelli, G. *et al.* Shared genetics linking sociability with the brain's default mode network. *medRxiv* 2024.05.24.24307883 (2024) doi:10.1101/2024.05.24.24307883.
80. Smallwood, J. *et al.* The default mode network in cognition: a topographical perspective. *Nat. Rev. Neurosci.* **22**, 503–513 (2021).
81. Glasser, M. F. *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80**, 105–124 (2013).
82. Strike, L. T. *et al.* The Queensland Twin Adolescent Brain Project, a longitudinal study of adolescent brain development. *Sci. Data* **10**, 195 (2023).
83. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
84. Cruces, R. R. *et al.* Micapipe: A pipeline for multimodal neuroimaging and connectome analysis. *NeuroImage* **263**, 119612 (2022).

85. Coifman, R. R. *et al.* Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proc. Natl. Acad. Sci.* **102**, 7426–7431 (2005).
86. Kang, H. J. *et al.* Spatiotemporal transcriptome of the human brain. *Nature* **478**, 483 (2011).
87. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
88. Almasy, L. & Blangero, J. Multipoint Quantitative-Trait Linkage Analysis in General Pedigrees. *Am. J. Hum. Genet.* **62**, 1198–1211 (1998).
89. Burt, J. B., Helmer, M., Shinn, M., Anticevic, A. & Murray, J. D. Generative modeling of brain maps with spatial autocorrelation. *NeuroImage* **220**, 117038 (2020).
90. Demontis, D. *et al.* Genome-wide analyses of ADHD identify 27 risk loci, refine the genetic architecture and implicate several cognitive domains. *Nat. Genet.* **55**, 198–208 (2023).
91. Wightman, D. P. *et al.* A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nat. Genet.* **53**, 1276–1282 (2021).
92. Watson, H. J. *et al.* Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nat. Genet.* **51**, 1207–1214 (2019).
93. Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **51**, 431–444 (2019).
94. Sullivan, P. bip2021. figshare https://doi.org/10.6084/m9.figshare.14102594.v2 (2021).
95. Als, T. D. *et al.* Depression pathophysiology, risk prediction of recurrence and comorbid psychiatric disorders using genome-wide analyses. *Nat. Med.* **29**, 1832–1844 (2023).
96. Arnold, P. D. *et al.* Revealing the complex genetic architecture of obsessive–compulsive disorder using meta-analysis. *Mol. Psychiatry* **23**, 1181–1188 (2018).
97. Trubetskoy, V. *et al.* Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).
98. Hatoum, A. S. *et al.* Multivariate genome-wide association meta-analysis of over 1 million subjects identifies loci underlying multiple substance use disorders. *Nat. Ment. Health* **1**, 210–223 (2023).
99. Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A. & Smoller, J. W. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat. Commun.* **10**, 1776 (2019).