ARTICLE OPEN

Check for updates

Genomic structural equation modeling reveals latent phenotypes in the human cortex with distinct genetic architecture

Rajendra A. Morey[1,2,3], Yuanchao Zheng[4,5], Henry Bayly[4,5], Delin Sun[1,2,3], Melanie E. Garrett[3,6], Marianna Gasperi[7,8,9], Adam X. Maihofer[8,9], C. Lexi Baird[1], Katrina L. Grasby[10], Ashley A. Huggins[1,2,3], Courtney C. Haswell[1,2], Paul M. Thompson[11], Sarah Medland[12], Daniel E. Gustavson[13], Matthew S. Panizzon[14], William S. Kremen[14], Caroline M. Nievergelt[7,8,9], Allison E. Ashley-Koch[3,6] and Mark W. Logue[4,5,15,16]✉

Genetic contributions to human cortical structure manifest pervasive pleiotropy. This pleiotropy may be harnessed to identify unique genetically-informed parcellations of the cortex that are neurobiologically distinct from functional, cytoarchitectural, or other cortical parcellation schemes. We investigated genetic pleiotropy by applying genomic structural equation modeling (SEM) to map the genetic architecture of cortical surface area (SA) and cortical thickness (CT) for 34 brain regions recently reported in the ENIGMA cortical GWAS. Genomic SEM uses the empirical genetic covariance estimated from GWAS summary statistics with LD score regression (LDSC) to discover factors underlying genetic covariance, which we are denoting *genetically informed brain networks* (GIBNs). Genomic SEM can fit a multivariate GWAS from summary statistics for each of the GIBNs, which can subsequently be used for LD score regression (LDSC). We found the best-fitting model of cortical SA identified 6 GIBNs and CT identified 4 GIBNs, although sensitivity analyses indicated that other structures were plausible. The multivariate GWASs of the GIBNs identified 74 genome-wide significant (GWS) loci ($p < 5 \times 10^{-8}$), including many previously implicated in neuroimaging phenotypes, behavioral traits, and psychiatric conditions. LDSC of GIBN GWASs found that SA-derived GIBNs had a positive genetic correlation with bipolar disorder (BPD), and cannabis use disorder, indicating genetic predisposition to a larger SA in the specific GIBN is associated with greater genetic risk of these disorders. A negative genetic correlation was observed between attention deficit hyperactivity disorder (ADHD) and major depressive disorder (MDD). CT GIBNs displayed a negative genetic correlation with alcohol dependence. Even though we observed model instability in our application of genomic SEM to high-dimensional data, jointly modeling the genetic architecture of complex traits and investigating multivariate genetic links across neuroimaging phenotypes offers new insights into the genetics of cortical structure and relationships to psychopathology.

Translational Psychiatry (2024)14:451 ; https://doi.org/10.1038/s41398-024-03152-y

INTRODUCTION

A number of different neurobiological markers have been employed in conjunction with various organizational schemes to map the human cortex. It is possible that individual differences in regional cortical surface area (SA) and cortical thickness (CT) may drive factors that affect each person and each region independently. However, the covariance structure of regional SA and CT reveals that individual differences are systematically coordinated within communities of brain regions, fluctuate in magnitude together within a population, may be instantiated as structural covariance networks (SCN) [1], and partially recapitulate established organizational schemes [2–5]. For instance, SCN organization is consistent with topological patterns of cortical maturation observed throughout developmental stages from childhood and adolescence into early adulthood [6], and the same patterns are then targeted by neurodegenerative diseases in late life [7, 8]. Second, brain regions with highly correlated CT or SA often represent networks that perform dedicated cognitive

[1]Brain Imaging and Analysis Center, Duke University, Durham, NC 27710, USA. [2]Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, NC, USA. [3]VISN 6 MIRECC, VA Health Care System, Croasdaile Drive, Durham, NC 27705, USA. [4]National Center for PTSD, VA Boston Healthcare System, Boston, MA 02130, USA. [5]Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118, USA. [6]Department of Medicine, Duke Molecular Physiology Institute, Carmichael Building, Duke University Medical Center, Durham, NC 27701, USA. [7]VA Center of Excellence for Stress and Mental Health, VA San Diego Healthcare System, San Diego, CA 92161, USA. [8]Research Service VA, San Diego Healthcare System, San Diego, CA 92161, USA. [9]Department of Psychiatry, University of California San Diego, La Jolla, CA 92093, USA. [10]Psychiatric Genetics, QIMR, Berghofer Medical Research Institute, 4006 Brisbane, QLD, Australia. [11]Imaging Genetics Center, Stevens Neuroimaging & Informatics Institute Keck School of Medicine University of Southern California, Los Angeles, CA 90033, USA. [12]Queensland Institute for Medical Research, Berghofer Medical Research Institute, 4006 Brisbane, QLD, Australia. [13]Institute for Behavioral Genetics, University of Colorado Boulder, Boulder, CO 80303, USA. [14]Stein Institute for Research on Aging, University of California San Diego, La Jolla, CA 92093, USA. [15]Department of Psychiatry, Boston University School of Medicine, Boston, MA 02118, USA. [16]Biomedical Genetics, Boston University School of Medicine, Boston, MA 02118-2526, USA. ✉email: loguem@bu.edu

Received: 10 August 2023 Revised: 30 September 2024 Accepted: 3 October 2024
Published online: 24 October 2024

processes [1, 9, 10]. Third, regions within SCNs tend to be directly connected by white matter tracts. Indeed, about 40% of SCN connections show convergent white matter fiber connections, although other relationships captured by SCNs are distinct from fiber connectivity [5].

The correlation structure between regions represented by a SCN is influenced by both the environment and genetics. The genetic factors underlying structural correlations closely resemble functional and developmental patterns [4, 5, 11]. We will refer to these patterns of genetic correlations between brain regions as *genetically informed brain networks* (GIBNs). Genetic correlations of CT and SA regions have been examined in twin studies [12, 13]. These genetic influences were recapitulated in over 400 twin pairs, to show that the cortex is organized genetically into communities of structural and functional regions, is hierarchical, modular, and bilaterally symmetric [11]. Their genetically informed parcellation identified 12 spatially contiguous regions that qualify as GIBNs.

While twin studies have laid important groundwork regarding genetic correlations of the brain, they have several limitations. First, twin studies do not provide specific genetic variants associated with genetically correlated regions [11]. Second, twin studies rely on the *equal environment* assumption, which may be invalid for some traits. Third, quantifying the genetic correlation between CT/SA and assembling a well-powered cohort of low prevalence traits such as schizophrenia (0.5% prevalence) [14] or bipolar disorder (1% prevalence) [15] is extremely difficult due to the rarity of pairs with twins affected by one or both traits. Recently, genetic correlations between brain regions derived from GWAS results have been applied to estimate the contribution of common genetic variation to CT/SA heritability [16]. This method confers several advantages over twin studies as they do not have the same assumptions, allow effect-size estimation for individual variants, and have the ability to test genetic correlations with other traits in different populations. These SA and CT GWAS results reveal pleiotropy and genetic correlation across many neuroimaging phenotypes [17, 18].

Genomic structural equation modeling (gSEM) is a multivariate statistical method that can leverage the genetic architecture of multiple genetically correlated phenotypes to derive relatively few latent phenotypes, which describe the observed genetic correlation and elucidate loadings of multiple phenotypes onto the latent phenotype [19]. Therefore, gSEM applied to GWAS offers a genetically informed clustering of the cortex that may be neurobiologically distinct from functional and cytoarchitectural parcellations [6, 20]. Multiple regions that have significant loadings on a particular factor define the brain regions that can be described as GIBN. Importantly, gSEM can be used to estimate the strength of association between genetic variants and each latent factor in a multivariate GWAS of each GIBN using GWAS summary statistics for the individual traits. Thus, gSEM provides a description of the underlying genetic architecture of the traits being examined and effect size estimates for specific SNPs and their association with the latent factors.

In the present study, we sought to elucidate the genetic architecture of 34 regional SA and CT phenotypes reported in the ENIGMA-3 GWAS of over 50,000 primarily healthy individuals. We hypothesized that gSEM might identify cortical SA networks consistent with the clusters described by Chen et al. [11], along with other viable solutions. The genetic correlations reported in Grasby et al. [18], were stronger within major anatomical lobes than across lobes. Thus, while we predicted gross lobar structure may be reflected by GIBNs, we further predicted that GIBNs may represent regions corresponding to functional networks, canonical resting-state networks (RSN), fiber tract networks, and other neurobiological systems [6, 11]. We hypothesized that most genetic variants discovered by the ENIGMA-3 cortical GWAS would influence GIBNs. We also sought to discover novel genetic markers and links between known genetic variants and GIBNs.

Our motivation for the present analysis was that there is robust evidence of disrupted cortical structure and function for most psychiatric disorders [21, 22]. We also know psychiatric disorders are polygenic and there is significant genetic correlation between disorders [23–27]. We further hypothesized genetic correlations between GIBNs and major neuropsychiatric disorders that are stronger than correlations between global measures of SA or CT.

METHODS

Data

We used the results of the ENIGMA-3 cortical GWAS as reported in Grasby et al. [18]. that identified genetic loci associated with variation in cortical SA and CT measures in 51,665 individuals, primarily (~94%) of European descent, from 60 international cohorts. All subjects provided informed consent to participate in study procedures approved by the local ethics board or IRB. The present study was deemed exempt by the Duke University IRB reviewed as Protocol ID: Pro00079963 and Protocol Title: Trauma and Genomics Modulate Brain Structure across Common Psychiatric Disorders. In ENIGMA-3, phenotype measures were extracted from structural MRI scans for 34 regions defined by the Desikan-Killiany atlas using gyral anatomy, which establishes coarse partitions of the cortex [28]. This study analyzed global measures of total cortical SA and average CT, as well as 34 regional measures of SA and CT averaged across left and right hemisphere structures to yield 70 distinct phenotypes. Multiple testing correction in the ENIGMA-3 GWAS was based on 70 independent phenotypes with a GWS threshold of $P \leq 8.3 \times 10^{-10}$. We accessed the GWAS summary results for the 34-regional bilateral analyses. The primary GWAS analyses presented in Grasby et al. adjusted for global SA and mean CT. However, we utilized alternate results without global adjustments, as the global-adjusted results produce multiple negative genetic correlations between regions (see Supplementary Figs. S1 and S2), which might be artifactual and lead to uninterpretable factor loadings. Regional SA and CT metrics were analyzed separately due to computational limitations and because negative genetic correlations between SA and CT could complicate model interpretation [16, 18, 29].

Ethics approval and consent to participate

All subjects provided informed consent to participate in study procedures approved by the local ethics board or IRB. The present study was deemed exempt by the Duke University IRB reviewed as Protocol ID: Pro00079963 and Protocol Title: Trauma and Genomics Modulate Brain Structure across Common Psychiatric Disorders.

Analysis

Our analyses were performed using the GenomicSEM R package [19]. The gSEM was performed twice, once for 34 SA regions and once for 34 CT regions. The gSEM fitting process includes an exploratory factor analysis (EFA) stage and a confirmatory factor analysis (CFA) stage. In gSEM, these steps are usually not performed on the basis of independent cohorts. Many gSEM studies include the same GWAS data in both EFA and CFA [30–33]. To avoid overfitting, other studies split the chromosomes into two distinct sets, usually odd-numbered autosomal chromosomes and even autosomal chromosomes, and use one for EFA and the other set for CFA [34–37] and sensitivity analysis [38]. In this study, we chose the second strategy and analyzed odd chromosomes in the EFA and even chromosomes in the CFA. Whereas SEM often fits multiple models corresponding to a priori hypotheses built on theoretical models, we took a hypothesis-free (data driven) approach. In the EFA, we fit models allowing for 1 to 10 factors, for each of SA and CT. Scree plots were examined to ensure that 10 factors would be sufficient (see Figs. S3 and S4). In the EFA step, positive factor loading estimates greater than a pre-specified threshold from the EFA were carried forward to the CFA to be re-estimated, and the remaining loading parameters were set to zero [39]. As there was no consensus on factor loading cutoffs [19, 40], we tested two thresholds: 0.3 and 0.5. Cross-loadings were allowed if they exceeded the threshold. Factors that loaded on only a single region were removed as single regions do not constitute factors. Therefore, some models with a large number of factors ended up as redundant and were not carried forward to CFA as the investigation of single regions was already carried out by Grasby et al. [18]. We additionally ran a sensitivity analysis in which EFA was run using the even chromosomes and CFA was performed using the odd chromosomes to determine if order affected the final model.

The Akaike Information Criteria (AIC) was used as our primary measure of model fit. For our purposes, a model which minimized the AIC was deemed optimal. Standardized root-mean square residual (SRMR), model χ^2, and Comparative Fit Index (CFI) were also calculated. A lower SRMR indicates a better model fit, with SRMR < 0.1 is indicative of acceptable fit, and SRMR < 0.05 is indicative of excellent fit [19]. A higher CFI indicates better model fit. A CFI > 0.9 is indicative of a good fit, and a CFI > 0.95 is indicative of excellent fit [19]. As opposed to regression modeling, where significant statistics represent the strength of association between the predictors in the response, genomic SEM, a significant χ^2 statistic represents a lack of fit between the observed genetic covariance matrix and the covariance matrix implied by the model [19]. However, with large sample sizes, the χ^2 statistics can be significant regardless of the model, which is not informative. We found all χ^2 statistics were highly significant ($p \sim 0$), and therefore not reported.

The top-performing factor models in the CFA were further optimized by successive removal of non-significant factor loadings, which is considered standard practice [41]. We additionally fit a model as part of the CFA step to account for the observed correlation between the factors. Specifically, we fit a bifactor model where a "total" CT or SA factor was added, which loaded on all regions, a second bifactor+ model where a total CT or SA factor loaded on all regions and the GWAS results from the corresponding average CT or total SA GWAS from Grasby et al. [18], and a multi-level model where all EFA factors loaded onto a 2nd order factor. The bifactor and bifactor+ models failed to converge and the multilevel models failed to improve model fit in all cases; hence these results are not reported.

GIBN overlap with alternate networks and parcellations

To explore the possible relevance of GIBNs to other parcellations of the cortex, we used Dice's Coefficient to measure percent volume overlap. We used permutation testing to determine the significance for each Dice's coefficient by estimating the probability that the magnitude of overlap occurred by chance. We used 1000 iterations of populating a given network with randomly selected brain regions, calculating its Dice's coefficient relative to the parcellation of interest, and then comparing the GIBNs true Dice's coefficient to the null distribution of 1000 Dice's coefficients. The relative position of Dice's coefficient for a particular GIBN-to-parcellation comparison within the probability distribution provided the significance. False Discovery Rate (FDR) was used to correct for multiple testing with GIBNs, receptors, networks, and clusters.

First, we conducted a quantitative analysis of the overlap between GIBN's and 7 canonical RSNs reported by Yeo and colleagues [42]. We quantitatively analyzed the overlap between GIBNs and networks based on 20 neuroreceptor density maps of Hansen et al. [43]. We calculated Dice's coefficient between each the 4 CT and 6 SA GIBNs and both high and low neuroreceptor densities defined by the top 20%, and bottom 20% receptor densities for serotonin-1a (5-HT1a), serotonin-1b (5HT-1b), serotonin-2a (5-HT2a), serotonin-4 (5-HT4), serotonin-6 (5-HT6), serotonin transporter (5-HTT), alpha-4 beta-2 nicotinic (α4β2), cannabinoid type-1 (CB1), dopamine D1 (D1), dopamine D2 (D2), dopamine transporter (DAT), fluorodopa (fDOPA), gamma-aminobutyric acid A (GABAa), histamine type-3 (H3), muscarinic acetylcholine (M1), metabotropic glutamate receptor-5 (mGluR5), opioid (MOR), norepinephrine (NorEpi), N-methyl-D-aspartic acid (NMDA), vesicular acetylcholine transporter (VachT).

Multivariate GWAS analysis

Using the GIBNs from our best-fitting model, we used gSEM to generate a multivariate GWAS of each GIBN. The GenomicSEM package *sumstats* program was used to perform the GWAS, with options set for a linear model (continuous outcome) and default parameters for the 'info' and 'maf' filters (info ≤ 0.6, MAF ≤ 0.01). MAF was determined based on 1000 G Phase3 EUR reference panel. The GWS associations ($p < 5 \times 10^{-8}$) for each GIBN were compared to the significant SNPs reported by Grasby et al. with and without the global correction. The FUnctional Mapping and Annotations (FUMA) package [44] was used to annotate results from each GIBN GWAS, including annotating SNPs to specific genes, identifying independent loci, and identifying potential functional variants. FUMA was run using LD in the 1000 G Phase3 EUR reference panel and the default FUMA parameters.

While CT and SA were examined separately, both for computational limitations and conceptual reasons, we used the multivariate GWAS results to estimate the genetic correlation between CT GIBNs and SA GIBNs, hypothesizing that they would be consistent with the negative genetic correlation between average CT and total SA [16]. Additionally, to examine the degree to which the CT and SA GIBNs genetically resembled the overall CT and SA measures, we estimated the genetic correlation between each GIBN and the average CT and total SA (uncorrected for ICV) as reported in the Grasby et al.

Polygenicity analysis

We examined the significant SNPs from the GIBN GWAS, as well as SNPs in LD using FUMA to test for functional associations with established behavioral traits and major neuropsychiatric disorders. First, we examined whether observed variants from the GWAS recapitulated GWS SNPs from previous GWASs of neuroimaging traits including cortical GWASs and other structural neuroimaging parameters [17, 45–50]. We also looked for SNPs that were significant in GWASs of 12 neuropsychiatric disorders from the Psychiatric Genomics Consortium (PGC): ADHD [51], alcohol dependence [52], anorexia nervosa [53], autism spectrum disorder [54], bipolar [55], cannabis use [56], MDD [57], obsessive-compulsive disorder (OCD) [58] posttraumatic stress disorder (PTSD) [59], schizophrenia [60], Tourette's syndrome [61], and anxiety [62]. Finally, FUMA was used to functionally annotate loci that overlapped with previously published GWAS results.

Genetic correlation with psychopathology

We used cross-trait LDSC to identify links between psychiatric disorders and CT-derived GIBNs as well as psychiatric disorders and SA-derived GIBNs [63]. We estimated the genetic correlation between CT- and SA-derived GIBNs and neuropsychiatric disorders using their GWAS summary statistics [63]. To limit our need for a multiple testing correction, we limited our analyses to the 12 neuropsychiatric disorders noted above. A false discovery rate corrected p value (P_{FDR}) was used to correct for the number of GIBNs [10] and disorders [12].

RESULTS

Model fit

The SA-derived 6-GIBN solution resulted in the best overall model fit to the genetic covariances generated from the GWAS summary statistics (AIC = 22,712,604, CFI = 0.920, SRMR = 0.062). See Supplementary Table S1 for fit statistics for each evaluated model. The 6 SA-derived GIBNs (SA1-SA6) loaded on 24 of the 34 brain regions [18]. The standardized estimates for the 6 SA-derived GIBN models (standardized factor loadings) are presented in Supplementary Table S2 and Fig. 1A. The GIBNs generally encompass contiguous brain regions and many correspond to known neuroanatomical features. SA1 contains loadings for inferior temporal, isthmus cingulate, postcentral, precuneus, superior parietal, supramarginal, and temporal pole. SA2 contains loadings for caudal anterior cingulate, caudal middle frontal, medial orbitofrontal, paracentral, and rostral anterior cingulate. SA3 contains loadings for banks superior temporal sulcus (STS), inferior parietal, and middle temporal. SA4 contains loadings for pars opercularis, pars orbitalis, and pars triangularis, SA5 contains loadings for cuneus, lateral occipital, lingual, and pericalcarine, and SA6 corresponds to the auditory cortex. The 6-factor model indicated substantial correlation between GIBNs ($r_g = 0.61$ to 0.91) as reported in Supplementary Table S3.

The CT-derived 4-GIBN solution resulted in the best model fit (AIC = 17761928, CFI = 0.932, SRMR = 0.077; Supplementary Table S4). Significant non-zero loadings for CT-derived GIBNS loaded on 25 of the 34 brain regions from Grasby et al. See Supplementary Table S5 for the estimated loadings that are visualized in Fig. 1B. As observed with SA models, the CT-derived GIBNs generally encompassed contiguous cortical regions. CT1 contains loadings for banks STS, caudal middle frontal, inferior parietal, paracentral, pars opercularis, post-central, pre-central, precuneus, rostral middle frontal, superior frontal, superior parietal, and supramarginal cortices. CT2 contains loadings for the caudal anterior cingulate, frontal pole, insula, lateral orbitofrontal, medial orbitofrontal, pars orbitalis, rostral anterior cingulate, and rostral middle frontal. CT3 contains loadings for banks STS, superior temporal, and temporal pole. CT4 contains loadings for cuneus, lateral

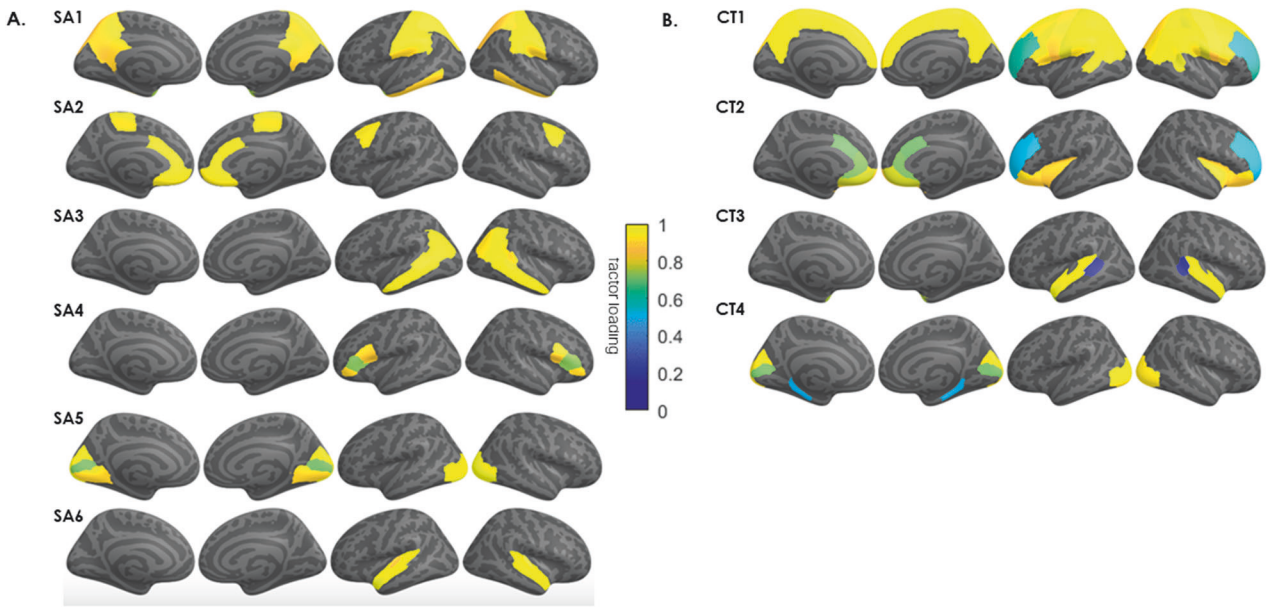

Fig. 1 Genetically Informed Brain Networks (GIBNs). Genomic structural equation modeling (gSEM) jointly modeled the genetic architecture of (**A**) cortical surface area (SA), and (**B**) cortical thickness, for 34 brain regions based on GWAS results of Grasby et al. [18]. The model generated 6 *genetically informed brain networks* (GIBNs) from SA phenotype measures. The color overlay on cortical regions represents the magnitude of the factor loadings indicated in the color gradient (yellow = high; blue = low). Subsequent GWAS identified several genome-wide significant hits ($p < 5 \times 10^{-8}$) associated with each GIBN.

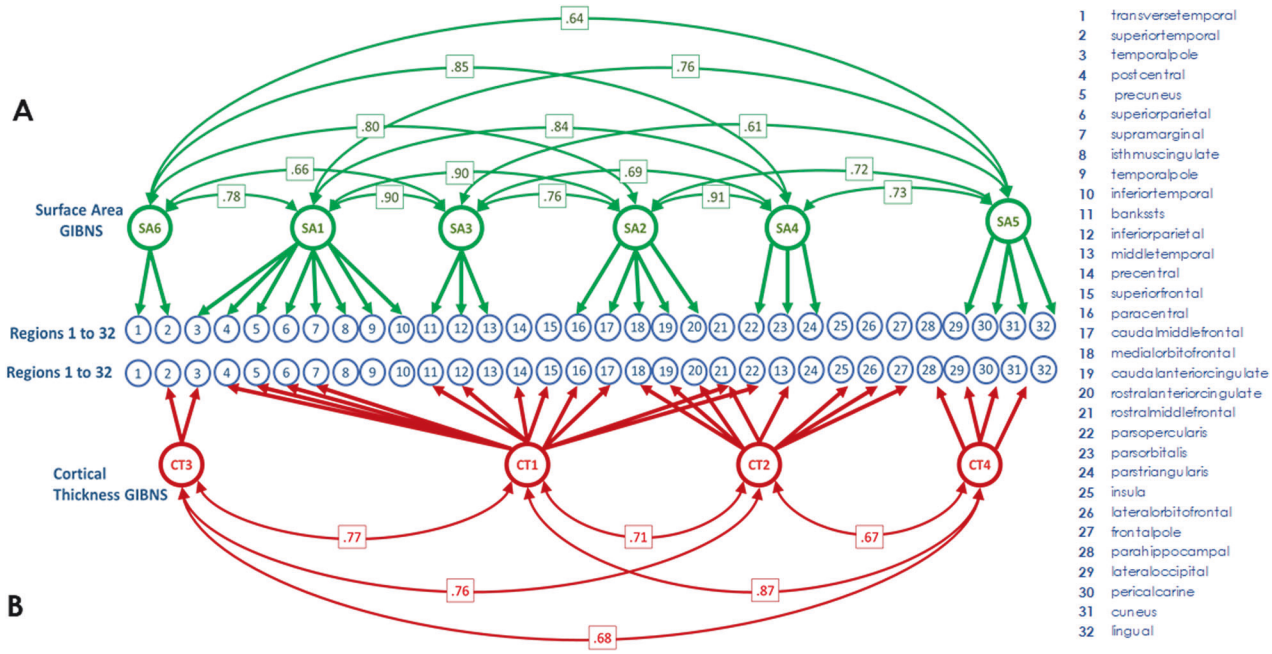

Fig. 2 Graph of genomic structural equation modeling (gSEM) results. The blue circles, numbered from 1 to 32, represent the cortical surface area (SA) and cortical thickness (CT) of regions defined by the Desikan–Killiany atlas in the figure legend. Two separate gSEMs were carried out on **A** Latent SA variables, indicated by green circles, representing the genetic contributions from regional SA, which are specified by thick green lines and arrows. Thin green lines connect genetically related SA variables with their genetic correlation strength (r_g) indicated in green boxes. **B** Latent CT variables, indicated by red circles, represent the genetic contributions from regional CT, which are specified by thick red lines and arrows. Thin red lines connect genetically related CT variables with their genetic correlation strength (r_g) indicated in red boxes. SA surface area, CT cortical thickness, GIBN genetically informed brain network.

occipital, parahippocampal, and pericalcarine cortices. The CT-derived GIBNs were moderately to highly correlated ($r_g = 0.67$ to 0.87; Supplementary Table S6).

Factor diagrams for SA- and CT-derived GIBNs are presented in Fig. 2. Consistent with prior work, the SA-derived GIBNs were largely distinct from CT-derived GIBNs, although some regional overlap exists. For example, SA5 and CT4 are both 4-region GIBNs, with 3 overlapping regions.

We then performed a sensitivity analysis by switching the order of the chromosomes, with the even chromosomes used in the EFA and the odd chromosomes used in the CFA. See Supplementary Tables S7–S12 for fit statistics and model estimates. The best

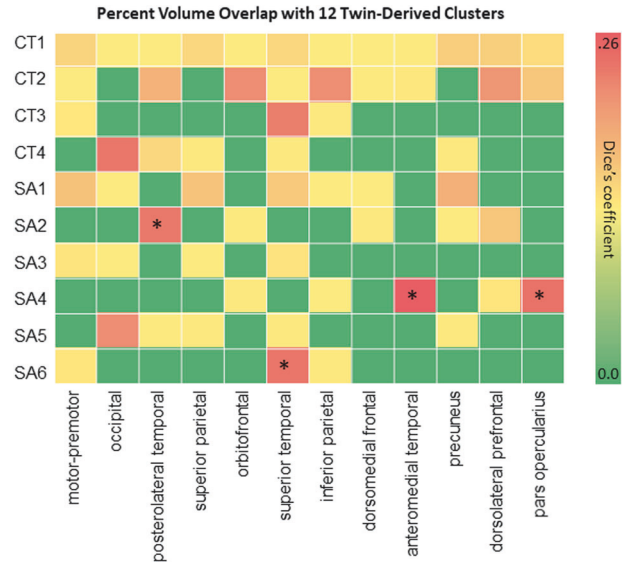

Fig. 3 Genetic parcellation of twin brains. We examined the percent volume overlap as measured by Dice's coefficients for the overlap between 6 SA GIBNs and the 12 clusters reported by Chen et al. [11] derived from twin brain data. *FDR-corrected *p* values for percent volume overlap with 12 clusters are indicated by an asterisk. SA surface area, CT cortical thickness, GIBN genetically informed brain network.

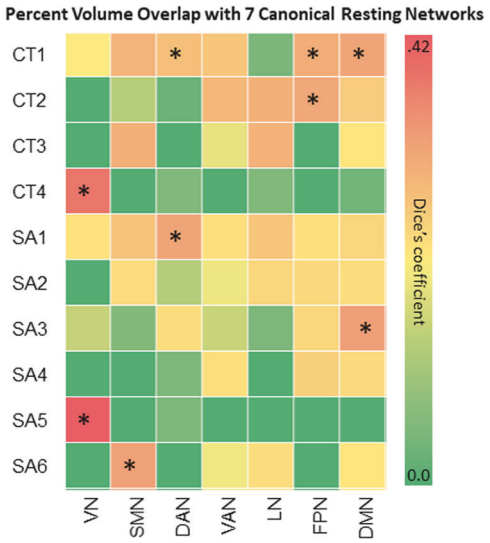

Fig. 4 Canonical resting-state networks. Percent volume overlap (Dice's coefficients) between GIBNs (4 CT and 6 SA) and canonical resting-state networks for the 7-network parcellation scheme by Yeo and colleagues (2011). *FDR-corrected *p* values for percent volume overlap between GIBNs and 7 canonical resting-state networks. VN visual network, SMN somatomotor network, DAN dorsal attention network, VAN ventral attention network, LN limbic network, FPN frontoparietal network, DMN default mode network, SA surface area, CT cortical thickness, GIBN genetically informed brain network.

fitting model from this sensitivity analyses had seven SA GIBNs and four CT GIBNs, with very similar, but slightly worse CFI fit statistics than the model described above (SA CFI = 0.89 and SRMR = 0.064, vs CFI of 0.92 and SRMR of 0.62 for the original SA GIBNs; CT CFI = 0.86 and SRMR = 0.077 vs CFI of 0.93, SRMR = 0.077 for the original CT GIBNs; AIC are not comparable). Interestingly, even though the fit was similar, the factor structure differed substantially from our original runs, although there were points of correspondence. For example, SA factor 2 from the 7-factor sensitivity analysis model corresponds to SA5 of the original model. The CT1 and CT2 GIBNs from the original runs were subsets of larger regions identified as CT1 and CT2 in the sensitivity analysis. We decided to continue the investigation of the SA and CT GIBNs from our original models based on CFI and SRMR, but we note that the sensitivity analyses indicate other solutions may be viable alternatives to this factor structure.

Overlap of SA GIBNs and twin-derived genetic SA parcellations

We computed Dice's coefficients (DCs) and corresponding *p* values between the SA GIBNs and the 12 SA correlation networks reported by Chen et al. [11]. Our results (Fig. 3) showed a high overlap between SA2 and the posterolateral temporal (DC = 0.205, $p_{FDR} = 0.02$), SA4 and pars opercularis (DC = 0.219, $p_{FDR} = 0.02$), SA4 and anteromedial temporal network (DC = 0.256, $p_{FDR} = 0.02$), SA4 and pars opercularis network (DC = 0.219, $p_{FDR} = 0.02$), SA4 and anteromedial temporal network (DC = 0.259, $p_{FDR} = 0.02$), and SA6 and superior temporal network (DC = 0.211, $p_{FDR} = 0.02$).

Overlap of GIBNs and RSNs

The Dice's coefficients and corresponding *p* values of GIBNs and the Yeo et al. [42] 7 RSNs (Fig. 4) showed that CT4 and SA5 had relatively high overlap with the visual network (CT4:DC = 0.353, $p_{FDR} = 0.008$; SA5:DC = 0.424, $p_{FDR} = 0.008$), while CT1 and SA3 had high overlap with the DMN (CT1:DC = 0.232, $p_{FDR} = 0.008$, SA3:DC = 0.249, $p_{FDR} = .008$, CT1 and SA1 with DAN (CT1:DC = 0.159, $p_{FDR} = 0.008$; SA1:DC = 0.232, $p = 0.008$), CT1 and CT2 with FPN (CT1:DC = 0.210, $p_{FDR} = 0.008$; CT2:DC = 0.225, $p_{FDR} = 0.008$), and SA6 with SMN (DC = 0.247, $p_{FDR} = 0.008$.

GIBN overlap with high/low neuroreceptor density regions

We examined the overlap between CT and SA GIBNs and regions of highest (top 20%; Fig. 5A) and lowest (bottom 20%, Fig. 5B) neuroreceptor densities. We found that CT1 overlapped a region of high neuroreceptor densities for many types of neuroreceptors and a region of low fDOPA receptor density. CT2 and SA2 overlapped regions of high 5-HT1a, 5-HT4, and 5HTT receptor density. SA5 overlapped the high 5HTT receptor-density region. (Fig. 5A).

GWAS of GIBNs

To identify specific genetic variants that may be influencing the GIBNs, we performed a multivariate GWAS on each SA- and CT-derived GIBN. Manhattan plots for SA- and CT-derived GIBN GWASs, their associated quantile-quantile (QQ) plots, and genomic inflation factors (λ) are provided in Supplementary Figs. S5–14. We observed moderate *p* value inflation (λ values between 1.06 and 1.16). However, the single-trait LD Score regression intercepts for SA- and CT-derived GIBNs were all less than 1.02, indicating that the apparent inflation was likely due to polygenicity. A total of 5,843 GWS ($p < 5 \times 10^{-8}$) variants were associated with the GIBNs. FUMA [44] mapped these variants to 74 independent regions, including 64 loci associated with the 6 SA-derived GIBNs and 10 loci associated with the 4 CT-derived GIBNs. A phenogram [64] of the genetic associations is presented in Fig. 6. A list of all GWS loci is provided in Table S13. Except for two novel SNPs, all others were previously identified in Grasby et al. [18]. in either the analyses adjusted for global SA/CT or the unadjusted analyses. The first novel SNP, rs3006933, near the genes *SDCCAG8* and *AKT3* on chromosome 1, was associated with SA1 ($p = 4.08 \times 10^{-9}$). The other novel SNP, rs1004763, on chromosome 22 in the vicinity of the gene *SLC16A8*, was associated with CT2 ($p = 3.41 \times 10^{-08}$). Notably, many of the 75 GWS loci associated with GIBNs were not associated with global measures, but only with individual CT/SA

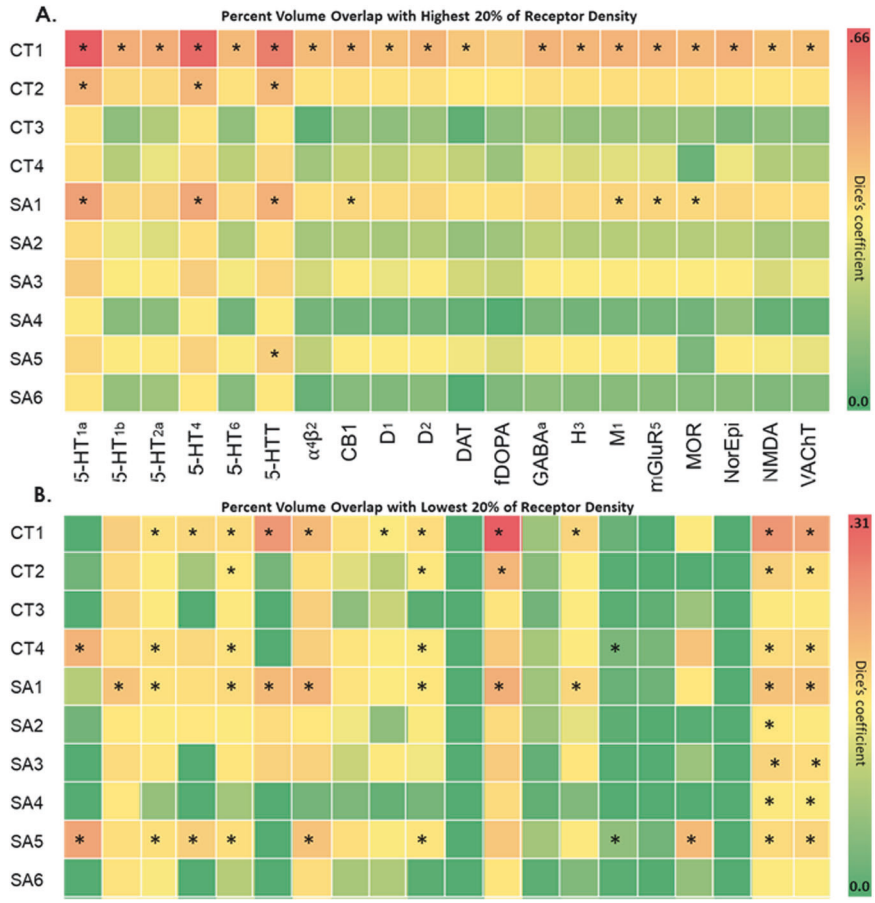

Fig. 5 Receptor density. Dice's coefficients and corresponding *p* values between each of the 4 CT and 6 SA GIBNs and neuroreceptor density. The 20 receptors included serotonin 1a (5-HT1a), serotonin 1b (5-HT1b), serotonin 2a (5-HT2a), serotonin 4 (5-HT4), serotonin 6 (5-HT6), serotonin transporter (5-HTT), alpha-4 beta-2 nicotinic (α4β2), cannabinoid type 1 (CB1), dopamine D1 (D1), dopamine D2 (D2), dopamine transporter (DAT), fluorodopa (fDOPA), gamma aminobutyric acid A (GABAa), histamine type 3 (H3), muscarinic acetylcholine (M1), metabotropic glutamate receptor 5 (mGluR5), opioid (MOR), norepinephrine (NorEpi), N-methyl-D-aspartic acid (NMDA), vesicular acetylcholine transporter (VAchT). In Hansen et al. [43] and the 10 GIBNs maps were used to calculate Dice's coefficient and corresponding FDR-corrected *p* values for **A** the highest 20% receptor density, and **B** the lowest 20% receptor density (* FDR-corrected *p* value < 0.05). The results, consisting of 200 Dice's coefficients (20 receptors × 10 GIBNs), are displayed in heatmaps. SA surface area, CT cortical thickness, GIBN genetically informed brain network.

regional measures, and half (37 out of 75) were more significantly associated with the GIBNs than the corresponding global measures. Using FUMA we found no significant enrichment of a particular tissue type in either CT- or SA-derived GIBNs and no enriched expression of developmental genes or regulators.

Genetic correlation between CT and SA

Although CT and SA regions were analyzed separately, we examined the genetic correlation between CT and SA using LDSC analysis of the GIBN GWAS results. The mean genetic correlation between SA GIBNs and CT GIBNs is −0.22 (−0.43 to −0.08; Table S14), whereas the mean genetic correlation between the 6 SA GIBNs is 0.77 (0.61 to 0.91; Table S3) and the mean genetic correlation between the 4 CT GIBNs is 0.76 (0.71 to 0.87; Table S6). The dramatically lower correlation between CT and SA compared to within SA and compared to within CT GIBNs supports separate gSEM analyses of CT and SA phenotypes.

LDSC analysis of genetic correlation

We examined the genetic correlation between CT and SA GIBNs and psychiatric disorders. The LDSC analysis of SA GIBNs is reported in Table S15 and Figure S15. ADHD exhibited a significant negative genetic correlation with all SA-derived GIBNs except SA4 ($r_g = -0.13$ to -0.20, $p = 3.29 \times 10^{-6}$ to 0.0038, $p_{FDR} = 0.00040$ to 0.039). Significant positive genetic correlations were observed between bipolar disorder and SA1, SA2, SA4, and SA5 (r_g=0.10 to 0.14, $p = 3.00 \times 10^{-4}$ to 0.0047, $p_{FDR} = 0.012$ to 0.043). Interestingly, we observed significant genetic correlations between MDD and SA-derived GIBNs, but in the opposite direction as bipolar disorder. We found a significant negative correlation between MDD and SA6, which was not associated with bipolar disorder ($r_g = -0.10$, $p = 0.0011$, $p_{FDR} = 0.17$). Negative correlations were non-significant after multiple-testing correction between MDD SA1-SA3, and SA5 ($r_g = -0.057$ to -0.080, $p_{unc} = 0.0090$ to 0.046), while SA4 was not genetically correlated with MDD ($p = 0.12$). SA4 was significantly correlated with cannabis use disorder (r_g=0.15, $p = 4.00 \times 10^{-4}$, $p_{FDR} = 0.012$), while SA2 correlation with cannabis use was non-significant (r_g=0.11, $p_{unc} = 0.011$).

Fewer genetic correlations were significant between CT-derived GIBN regions and psychiatric disorders (Table S16 and Figure S16). CT3 and CT4 were negatively correlated with alcohol use disorder, exhibiting the strongest correlations with any traits that we examined (CT3 $r_g = -0.35$, $p = 3 \times 10^{-4}$, $p_{FDR} = 0.012$; CT4 $r_g = -0.31$, $p = 7 \times 10^{-4}$, $p_{FDR} = 0.014$). We found a negative nominally significant correlation between alcohol use disorder and CT1 ($r_g = -0.18$, $p_{unc} = 0.035$, $p_{FDR} = 0.22$). CT3 had a positive nominally significant correlation with OCD (r_g=0.22, $p_{unc} = 0.0091$, $p_{FDR} = 0.078$).

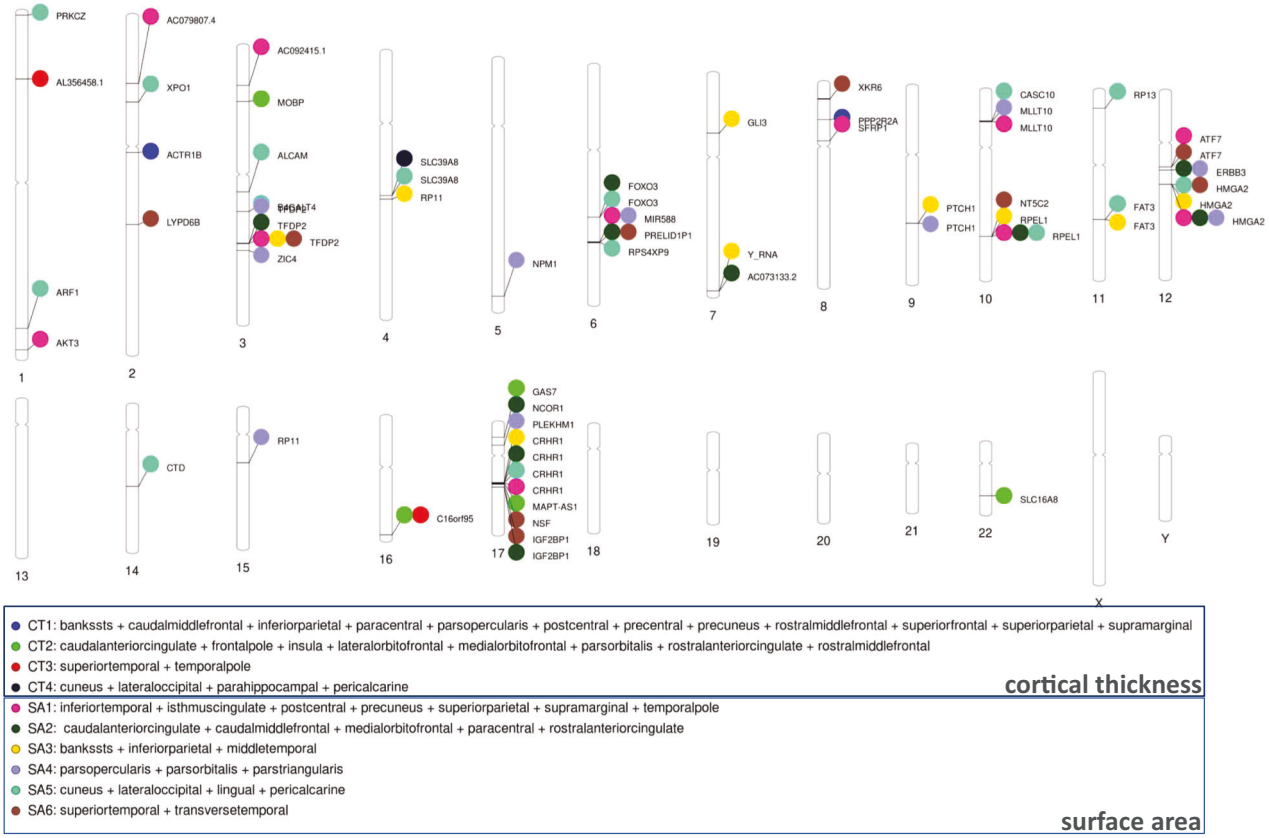

Fig. 6 SNPs for GIBNS derived from surface area and cortical thickness. Phenogram of GWS SNPS associated with six genetically informed brain networks (GIBNs) derived from surface area (SA) and four GIBNs derived from the cortical thickness (CT).

The global measures for thickness and SA are genetically correlated to many of the same psychopathology traits as the GIBNs, but the genetic correlations with global measures are almost always less significant than the genetic correlations with the most strongly associated GIBNs. The details of these results are provided in Tables S15 and S16.

DISCUSSION

The goal of the present study was to leverage the pleiotropic architecture of the human cortex to investigate genetic factors underlying CT and SA, and to identify further links between the genetics of CT, SA, and psychopathology. We applied gSEM to jointly model the genetic architecture of 34 brain regions using results from the ENIGMA-3 GWAS [18]. The process was undertaken with gSEM to generate several possible solutions, from which the best-model fit was selected. This solution organized brain regions to optimally assign genetic pleiotropy to 6 SA- and 4 CT-derived latent factors, which we have termed *genetically informed brain networks* (GIBNs).

The GIBNs we generated may be compared to similar structures generated from twin studies. Using 400 twin pairs, Chen et al. generated twelve genetically-informed clusters from vertex-based SA measures [11]. Chen et al. [11]. reported heritability estimates and genetic correlations between genetically informed parcels that are more consistent with classical anatomically-defined sulcal and gyral boundaries, Brodmann definitions, and cytoarchitectural patterns than our GIBNs. The slight differences in model fit between potential alternating models from our main analyses and sensitivity analyses indicate that this is one of several viable solutions with similar fit statistics. Even so, it is instructive to examine ways this solution corresponds to prior parcellations of the cortex based on genetic correlation as determined in twin studies as well as neurobiological and functional parcellations.

The best-fitting SA-derived GIBNs overlap with several canonical RSNs, such as visual network and SA5 (Dice's coefficient=0.424), which is composed of cuneus, lateral occipital, lingual, and pericalcarine cortices [42]. Twin-based non-linear multidimensional heritability estimates are among the highest for the visual network (left $h^2_m = 0.53$; right $h^2_m = 0.45$) and auditory network (left $h^2_m = 0.44$; right $h^2_m = 0.60$) [65]. SA6, which includes superior and transverse temporal cortices, overlaps the auditory cortex from twin-derived genetic parcels (Dice's coefficient = 0.211; Fig. 3). The functional specializations of the human auditory cortex [42], which include parts of the lateral prefrontal cortex, Broca's area, and subcentral regions, are needed for human vocalization and language [66, 67]. The dorsal attention network (DAN), which directs voluntary allocation of attention, has substantial overlap with SA1 (Dice's coefficient = 0.232, Fig. 4) that is comprised of superior parietal, supramarginal, postcentral, precuneus, isthmus cingulate, and inferior temporal regions. A noteworthy omission from SA1, an important feature of the DAN, are the frontal eye fields (FEF) [68]. Since FEF is not a FreeSurfer parcellation output, it may be poorly represented in the ENIGMA cortical GWAS. The DAN has relatively high twin heritability estimates (left $h^2 = 0.45$; right $h^2 = 0.40$) [65]. SA4 partially overlaps (Dice's coefficient=0.259) the frontoparietal network (FPN), which includes pars opercularis, pars orbitalis, and pars triangularis, but lacks the critical temporoparietal structures [69] (Fig. 4). One advantage of using gSEM and identifying GIBNS as a strategy is that it allows us to align the genetic correlation of clusters of regions with neurobiological features such as connectivity and gene expression, and implicates specific GIBN-

associated genetic variants and the degree to which they align with other traits such as psychopathology.

The overlap of GIBN GWS loci with prior GWAS of neuroimaging phenotypes or psychiatric disorders firmly points to the relevance of GIBN-related (multi-regional SA and CT) variants to brain structure and cognition. First, we note that novel variant rs3006933 has been previously associated with subcortical volumes [70]. Novel variants rs3006933 and rs1004763 [17, 48] have been associated with neuroimaging phenotypes of corpus callosum white matter microstructure [71]. A comparison of our GIBN GWAS with published psychiatric disorder GWAS results found that multiple SNPs linked to SA-derived GIBNs were also implicated in a GWAS of schizophrenia [60]. Specifically, we identified a cluster of 4 loci in the *CRHR1* gene strongly associated with SA-derived GIBNs (rs62057153 associated with SA1) in our GWAS ($p = 5.22 \times 10^{-17}$ to 8.45×10^{-21}). We also observed an association between CT1 and rs11692435 ($p = 1.17 \times 10^{-12}$), a schizophrenia-related locus, within the *ACTR1B* gene. Finally, CT- and SA-derived GIBNs were associated with schizophrenia risk variants in the *SLC39A8* gene; namely rs13107325 was associated with CT5 and rs13135092 was associated with SA5. No other traits had GWS variants associated with any of the GIBNs.

Many GIBN-associated SNPs have been associated with other cognitive, behavioral, neuroanatomical, neurofunctional, and neuropsychiatric phenotypes. In addition to rs3006933 noted above [16], SA6-linked locus rs9909861 [72] and SA5-linked SNP rs7570830 [16] have been associated with subcortical volumes. Multiple loci associated with SA-derived GIBNs that encompass temporal, parietal, and temporo-parietal association cortices, including the SA1-linked locus rs10109434 [73], the SA3-linked SNP rs2299148 [74], and the SA6-linked locus rs9909861 [74–78] have been implicated in academic attainment and cognitive ability. Regions in SA6, namely superior temporal gyrus, and SA3, namely supramarginal gyrus were the most strongly linked to academic attainment in the UKB sample [79]. The same regions were reported independently in the Queensland Twin Imaging and HCP samples [80]. The SA5-linked locus rs6701689 has been reported for risk tolerance [81]. However, a role for SA5 in risk tolerance is unsupported. Risk tolerance is linked to cerebellar, midbrain, and prefrontal cortical anatomy, as well as glutamatergic and GABAergic neurotransmission [81, 82]. The CT4-associated locus rs13107325 has been associated with many traits including schizophrenia [83–90], bipolar disorder [87, 88], Parkinson's disease [89, 90], sedentary behavior [70, 91] and risk taking [81], as well as cognition, intelligence, and educational attainment [74–78, 92]. CT4 includes the parahippocampal and fusiform gyri, which have firmly established links to schizophrenia [93] and sedentary behavior [94].

Behavioral traits and neuropsychiatric disorders showed distinct genetic correlations with SA-derived GIBNs that differ markedly from correlations with CT-derived GIBNs. CT3, located in the middle and superior temporal cortices, and CT4, located in the visual perceptual cortex, were strongly negatively correlated with alcohol use disorder. This divergent relationship between CT-derived and SA-derived networks is consistent with the ENIGMA-3 cortical GWAS where a similar pattern of positive and negative correlations between total brain SA and behavioral traits/disorders was found, but average CT correlations with behavioral traits/disorders were non-significant [18]. Specifically, the ENIGMA-3 GWAS found that total SA was significantly positively correlated with cognitive function, educational attainment, Parkinson's disease, and anorexia nervosa, but significantly negatively correlated with MDD, ADHD, depressive symptoms, neuroticism, and insomnia. In addition, the SA-derived GIBNs showed distinct genetic relationships to several psychiatric disorders. Several SA-derived GIBNs (SA1, SA2, SA4, SA5) were positively correlated with bipolar disorder, whereas SA-derived GIBNs (SA1, SA2, SA3, SA5, SA6) were negatively correlated with MDD, buttressing prior evidence that MDD and Bipolar are distinct conditions with diverging genetics [27]. While the relationship between these SA-derived GIBNs and MDD converge with the findings from the ENIGMA total SA results, the relationship with bipolar disorder was novel. Thus, GIBNs may provide additional power to detect genetic relationships when their strength across cortical regions is heterogenous.

Interestingly, although several GIBN-associated SNPs were associated with schizophrenia, no GIBNs were significantly genetically correlated with schizophrenia (r_g=0.029 to 0.034; p values > 0.30). While this may be counterintuitive, genetic correlation between phenotypes predicts an overlap in SNPs, but the reverse may not be true. A genetic correlation could be zero when many variants affect both traits, but the direction of effects are uncorrelated across variants [95].

There is ample evidence that genetic variants that influence SA are distinct from genetic variants that influence CT [18]. The results of Panizzon et al. [29] and Grasby et al. [18] compared to van der Meer et al. [16] are focused on different, albeit related measures. van der Meer et al. [16] primarily focused on the overlap of individual genetic variants associated with CT and SA and only secondarily on their genetic correlation. By contrast, Panizzon et al. [29] focused on genetic correlation from twin data, which means genetic marker associations were not available. Grasby et al. [18] was focused primarily on GWAS results and secondarily on reporting genetic correlations between CT and SA. Indeed, the results of van der Meer [16] are completely consistent with the results of Grasby et al. [18], with the former reporting a genetic correlation between SA and CT of – 0.26, and the latter of – 0.32. However, van der Meer et al. [16] reports that the 4016 out of 7941 causal variants are shared between CT and SA. Therefore, the relatively high genetic overlap, but low genetic correlation, is due to a mixture of opposing and agreeing effects from variants. Genetic variation affecting gene regulation in progenitor cell types, present in fetal development, affects adult cortical SA [96]. An increase in proliferative divisions of neural progenitor cells leads to an expanded pool of progenitors, resulting in increased neuronal production and larger cortical SA, which is more prevalent in gyrencephalic species (e.g. humans, primates) [97]. By contrast, loci near genes implicated in cell differentiation, migration, adhesion, and myelination are associated with CT. Our findings suggest this distinction holds for SA-derived compared to CT-derived GIBNs. We hypothesize that the unique genetic correlations of SA-derived GIBNs and CT-derived GIBNs with behavioral traits/disorders may be explained by the distinct developmental functions of their associated genes [98].

Limitations

A number of limitations deserve consideration in interpreting the present findings. First, we note that even with the large sample size of the Grasby et al. [17, 18] GWAS, there were many models that fit approximately as well as our final model in terms of the number of GIBNs and the clustering of the regions. Our sensitivity analyses, in which the order of the chromosomes in the EFA and CFA were switched provided alternate solutions. Many of these solutions had similar CFI and SRMR scores, indicating acceptable fit according to the CFI and SRMR criteria, but none had a consistently excellent fit. This may be an artifact of the number of regions examined and the dimensionality of the model. Therefore, our set of GIBNs should be considered a working model rather than the only possible genetic parcellation of the cortex. This model is useful, in that it allows us to generate a set of genetic associations with a smaller number of tests than considering each of the regions individually.

There are also other limitations to consider. We only examined genetic correlation with a handful of traits to reduce the multiple testing burden, but note that there are other traits worthy of interest, but beyond the scope of this manuscript. This also

includes traits like insomnia and restless sleep, given the reported genetic relationship between psychiatric disorders and diurnal sleep patterns [32]. Other traits such as Parkinson's and Alzheimer's disease would also be of interest [99, 100]. There are other analyses that can be run to further interpret the multivariate GWAS results, including Mendelian Randomization to support a causal relationship between cortical structure and psychiatric disorders and neurological traits. However, given the instability of the model estimates in the sensitivity analyses, and large number of models with similar fit indices, we think it is premature to exhaustively investigate the characteristics of these GIBNs. It is likely the factor structure we have presented will be supplanted in the future as higher spatial resolution genetic associations become available. The present gSEM was based on the GWAS results of Grasby et al. [17, 18], which averaged left and right hemisphere phenotypic measures. Additionally, Grasby et al. examined the 34 cortical regions as defined by the Desikan-Killiany atlas. A high-resolution GWAS of the cortex would allow more flexibility and redefining parcellation boundaries informed by genetic pleiotropy and likely yield GIBNs which correspond more closely to anatomical and functional boundaries. However, simultaneously analyzing GWAS results for each of 100 s or 1000 s of vertices would present a computational challenge without major advances in gSEM methodology. Therefore, the GIBNs we presented here are a proof-of-concept that genetic correlation can be used to enhance the interpretation of high-dimensional GWAS results and provide novel insights into the relationship between neuroimaging phenotypes and psychiatric disorders.

Conclusion

We harnessed the pervasive pleiotropy of the human cortex to realize a genetically-informed parcellation that is neurobiologically distinct from functional, cytoarchitectural, and other established cortical parcellations, yet harbors meaningful topographic similarities to other network schemas. Strong genetic correlation between GIBNs and several major neuropsychiatric conditions, coupled with clear confirmation that nearly all GIBN-associated SNPs play a role in cognitive, behavioral, neuroanatomical, and neurofunctional phenotypes, begins to expose the deeply interconnected architecture of the human cortex. Applying gSEM to model the joint genetic architecture of complex traits and investigate multivariate genetic links across phenotypes offers a new vantage point for mapping genetically informed cortical networks, although with limitations that must be carefully considered when results are interpreted.

DATA AVAILABILITY

GWAS summary statistics used in this paper are available on the ENIGMA consortium website (http://enigma.ini.usc.edu/research/download-enigma-gwas-results). The Genomic SEM package used to analyze the data is publicly available at https://github.com/GenomicSEM/GenomicSEM. The *ldsc* package is publicly available at https://github.com/bulik/ldsc. The results of the multivariate GWASs of the CT- and SA-derived GIBNs are available at https://pgc-ptsd.com/about/workgroups/imaging-workgroup/. All methods were performed in accordance with relevant guidelines and regulations.

REFERENCES

1. Zielinski BA, Gennatas ED, Zhou J, Seeley WW. Network-level structural covariance in the developing brain. Proc Natl Acad Sci. 2010;107:18191–6.
2. Romero-Garcia R, Whitaker KJ, Váša F, Seidlitz J, Shinn M, Fonagy P, et al. Structural covariance networks are coupled to expression of genes enriched in supragranular layers of the human cortex. Neuroimage. 2018;171:256–67.
3. Feng J, Chen C, Cai Y, Ye Z, Feng K, Liu J, et al. Partitioning heritability analyses unveil the genetic architecture of human brain multidimensional functional connectivity patterns. Hum Brain Mapp. 2020;41:3305–17.
4. Hawrylycz M, Miller JA, Menon V, Feng D, Dolbeare T, Guillozet-Bongaarts AL, et al. Canonical genetic signatures of the adult human brain. Nat Neurosci. 2015;18:1832–44.
5. Gong G, He Y, Chen ZJ, Evans AC. Convergence and divergence of thickness correlations with diffusion connections across the human cerebral cortex. Neuroimage. 2012;59:1239–48.
6. Alexander-Bloch A, Giedd JN, Bullmore E. Imaging structural co-variance between human brain regions. Nat Rev Neurosci. 2013;14:322–36.
7. Seeley WW, Crawford RK, Zhou J, Miller BL, Greicius MD. Neurodegenerative diseases target large-scale human brain networks. Neuron. 2009;62:42–52.
8. He Y, Chen Z, Evans A. Structural insights into aberrant topological patterns of large-scale cortical networks in alzheimer's disease. J Neurosci. 2008;28:4756–66.
9. Segall JM, Allen EA, Jung RE, Erhardt EB, Arja SK, Kiehl KA, et al. Correspondence between structure and function in the human brain at rest. Front Neuroinform. 2012;6:10.
10. Zhang Z, Liao W, Zuo X-N, Wang Z, Yuan C, Jiao Q, et al. Resting-state brain organization revealed by functional covariance networks. PLoS One. 2011;6:e28817.
11. Chen C-H, Gutierrez E, Thompson W, Panizzon MS, Jernigan TL, Eyler LT, et al. Hierarchical genetic organization of human cortical surface area. Science. 2012;335:1634–6.
12. Lenroot RK, Schmitt JE, Ordaz SJ, Wallace GL, Neale MC, Lerch JP, et al. Differences in genetic and environmental influences on the human cerebral cortex associated with development during childhood and adolescence. Hum Brain Mapp. 2009;30:163–74.
13. Schmitt J, Lenroot R, Wallace G, Ordaz S, Taylor K, Kabani N, et al. Identification of genetically mediated cortical networks: a multivariate study of pediatric twins and siblings. Cereb Cortex. 2008;18:1737–47.
14. Tichenor M, Sridhar D. Metric partnerships: global burden of disease estimates within the world bank, the world health organisation and the institute for health metrics and evaluation. Wellcome Open Res. 2019;4:35.
15. Rowland TA, Marwaha S. Epidemiology and risk factors for bipolar disorder. Ther Adv Psychopharmacol. 2018;8:251–69.
16. van der Meer D, Frei O, Kaufmann T, Chen C-H, Thompson WK, O'Connell KS, et al. Quantifying the polygenic architecture of the human cerebral cortex: extensive genetic overlap between cortical thickness and surface area. Cereb Cortex. 2020;30:5597–603.
17. Smith SM, Douaud G, Chen W, Hanayik T, Alfaro-Almagro F, Sharp K, et al. An expanded set of genome-wide association studies of brain imaging phenotypes in uk biobank. Nat Neurosci. 2021;24:737–45.
18. Grasby KL, Jahanshad N, Painter JN, Colodro-Conde L, Bralten J, Hibar DP, et al. The genetic architecture of the human cerebral cortex. Science. 2020;367:eaay6690.
19. Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, et al. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. Nat Hum Behav. 2019;3:513–25.
20. Kovas Y, Plomin R. Generalist genes: Implications for the cognitive sciences. Trends Cogn Sci. 2006;10:198–203.
21. McTeague LM, Rosenberg BM, Lopez JW, Carreon DM, Huemer J, Jiang Y, et al. Identification of common neural circuit disruptions in emotional processing across psychiatric disorders. Am J Psychiatry. 2020;177:411–21.
22. Fornito A, Bullmore ET, Zalesky A. Opportunities and challenges for psychiatry in the connectomic era. Biol Psychiatry: Cogn Neurosci Neuroimaging. 2017;2:9–19.
23. Gandal MJ, Haney JR, Parikshak NN, Leppa V, Ramaswami G, Hartl C, et al. Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. Science. 2018;359:693–7.
24. Grotzinger AD. Shared genetic architecture across psychiatric disorders. Psychol Med. 2021;51:2210–6.
25. Ward J, Tunbridge EM, Sandor C, Lyall LM, Ferguson A, Strawbridge RJ, et al. The genomic basis of mood instability: Identification of 46 loci in 363,705 uk biobank participants, genetic correlation with psychiatric disorders, and association with gene expression and function. Mol Psychiatry. 2020;25:3091–9.
26. Consortium C-DGotPG (2013): identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet 2013;381:1371–9.
27. Lee PH, Feng Y-CA, Smoller JW. Pleiotropy and cross-disorder genetics among psychiatric disorders. Biol Psychiatry. 2021;89:20–31.
28. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. Neuroimage. 2006;31:968–80.
29. Panizzon MS, Fennema-Notestine C, Eyler LT, Jernigan TL, Prom-Wormley E, Neale M, et al. Distinct genetic influences on cortical surface area and cortical thickness. Cereb Cortex. 2009;19:2728–35.
30. Abdellaoui A, Verweij KJ. Dissecting polygenic signals from genome-wide association studies on human behaviour. Nat Hum Behav. 2021;5:686–94.
31. Waldman ID, Poore HE, Luningham JM, Yang J. Testing structural models of psychopathology at the genomic level. World Psychiatry. 2020;19:350–9.

32. Morrison CL, Winiger EA, Rieselbach MM, Vetter C, Wright JrKP, LeBourgeois MK, et al. Sleep health at the genomic level: six distinct factors and their relationships with psychopathology. Biol Psychiatry Glob Open Sci. 2023;3:530–40.
33. Karlsson Linnér R, Mallard TT, Barr PB, Sanchez-Roige S, Madole JW, Driver MN, et al. Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. Nat Neurosci. 2021;24:1367–76.
34. Grotzinger AD, Mallard TT, Akingbuwa WA, Ip HF, Adams MJ, Lewis CM, et al. Genetic architecture of 11 major psychiatric disorders at biobehavioral, functional genomic and molecular genetic levels of analysis. Nat Genet. 2022;54:548–59.
35. Baselmans B, Van de Weijer M, Abdellaoui A, Vink J, Hottenga J, Willemsen G, et al. A genetic investigation of the well-being spectrum. Behav Genet. 2019;49:286–97.
36. Clifford RE, Maihofer AX, Chatzinakos C, Coleman JR, Daskalakis NP, Gasperi M, et al. Genetic architecture distinguishes tinnitus from hearing loss. Nat Commun. 2024;15:614.
37. Foote IF, Jacobs BM, Mathlin G, Watson CJ, Bothongo PL, Waters S, et al. The shared genetic architecture of modifiable risk for alzheimer's disease: a genomic structural equation modelling study. Neurobiol Aging. 2022;117:222–35.
38. Breunig S, Lawrence JM, Foote IF, Gebhardt HJ, Willcutt EG, Grotzinger AD. Examining differences in the genetic and functional architecture of attention-deficit/hyperactivity disorder diagnosed in childhood and adulthood. Biol Psychiatry Glob Open Sci. 2024;4:100307.
39. Howard MC. A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve? Int J Hum-Comput Int. 2016;32:51–62.
40. de Vries A, Tiemens B, Cillessen L, Hutschemaekers G. Construction and validation of a self-direction measure for mental health care. J Clin Psychol. 2021;77:1371–83.
41. Streiner DL. Building a better model: an introduction to structural equation modelling. Can J Psychiatry Rev. 2006;51:317–24.
42. Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol. 2011;106:1125–65.
43. Hansen JY, Markello RD, Tuominen L, Nørgaard M, Kuzmin E, Palomero-Gallagher N, et al. Correspondence between gene expression and neurotransmitter receptor and transporter density in the human brain. Neuroimage. 2022;264:119671.
44. Watanabe K, Taskesen E, Van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with fuma. Nat Commun. 2017;8:1–11.
45. Shadrin AA, Kaufmann T, van der Meer D, Palmer CE, Makowski C, Loughnan R, et al. Vertex-wise multivariate genome-wide association study identifies 780 unique genetic loci associated with cortical morphology. NeuroImage. 2021;244:118603.
46. Hofer E, Roshchupkin GV, Adams HH, Knol MJ, Lin H, Li S, et al. Genetic correlations and genome-wide associations of cortical structure in general population samples of 22,824 adults. Nat Commun. 2020;11:1–16.
47. Adams HH, Hibar DP, Chouraki V, Stein JL, Nyquist PA, Rentería ME, et al. Novel genetic loci underlying human intracranial volume identified through genome-wide association. Nat Neurosci. 2016;19:1569–82.
48. Elliott LT, Sharp K, Alfaro-Almagro F, Shi S, Miller KL, Douaud G, et al. Genome-wide association studies of brain imaging phenotypes in UK biobank. Nature. 2018;562:210–6.
49. Arfan Ikram M, Fornage M, Smith AV, Seshadri S, Schmidt R, Debette S, et al. Common variants at 6q22 and 17q21 are associated with intracranial volume. Nat Genet. 2012;44:539–44.
50. Hibar DP, Stein JL, Renteria ME, Arias-Vasquez A, Desrivieres S, Jahanshad N, et al. Common genetic variants influence human subcortical brain structures. Nature. 2015;520:224–9.
51. Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nat Genet. 2019;51:63.
52. Walters RK, Polimanti R, Johnson EC, McClintick JN, Adams MJ, Adkins AE, et al. Transancestral gwas of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. Nat Neurosci. 2018;21:1656–69.
53. Watson HJ, Yilmaz Z, Thornton LM, Hübel C, Coleman JR, Gaspar HA, et al. Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. Nat Genet. 2019;51:1207–14.
54. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, et al. Identification of common genetic risk variants for autism spectrum disorder. Nat Genet. 2019;51:431–44.
55. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. Nat Genet. 2019;51:793–803.
56. Johnson EC, Demontis D, Thorgeirsson TE, Walters RK, Polimanti R, Hatoum AS, et al. A large-scale genome-wide association study meta-analysis of cannabis use disorder. Lancet Psychiatry. 2020;7:1032–45.
57. Howard DM, Adams MJ, Clarke T-K, Hafferty JD, Gibson J, Shirali M, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. Nat Neurosci. 2019;22:343–52.
58. Arnold PD, Askland KD, Barlassina C, Bellodi L, Bienvenu O, Black D, et al. Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. Mol Psychiatry. 2018;23:1181.
59. Maihofer AX, Choi KW, Coleman JR, Daskalakis NP, Denckla CA, Ketema E, et al. Enhancing discovery of genetic variants for posttraumatic stress disorder through integration of quantitative phenotypes and trauma exposure information. Biol Psychiatry. 2021;91:626–636.
60. Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature. 2022;604:502–8.
61. Yu D, Sul JH, Tsetsos F, Nawaz MS, Huang AY, Zelaya I, et al. Interrogating the genetic determinants of tourette's syndrome and other tic disorders through genome-wide association studies. Am J Psychiatry. 2019;176:217–27.
62. Otowa T, Hek K, Lee M, Byrne EM, Mirza SS, Nivard MG, et al. Meta-analysis of genome-wide association studies of anxiety disorders. Mol Psychiatry. 2016;21:1391–9.
63. Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Patterson N, et al. Ld score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015;47:291–5.
64. Wolfe D, Dudek S, Ritchie MD, Pendergrass SA. Visualizing genomic information across chromosomes with phenogram. BioData Min. 2013;6:1–12.
65. Anderson KM, Ge T, Kong R, Patrick LM, Spreng RN, Sabuncu MR, et al. Heritability of individualized cortical network topography. Proc Natl Acad Sci. 2021;118:e2016271118.
66. Forseth KJ, Hickok G, Rollo P, Tandon N. Language prediction mechanisms in human auditory cortex. Nat Commun. 2020;11:1–14.
67. Celesia GG. Organization of auditory cortical areas in man. Brain. 1976;99:403–14.
68. Mengotti P, Käsbauer A-S, Fink GR, Vossel S. Lateralization, functional specialization, and dysfunction of attentional networks. Cortex. 2020;132:206–22.
69. Greene CM, Soto D. Functional connectivity between ventral and dorsal frontoparietal networks underlies stimulus-driven and working memory-driven sources of visual distraction. NeuroImage. 2014;84:290–8.
70. van der Meer D, Frei O, Kaufmann T, Shadrin AA, Devor A, Smeland OB, et al. Understanding the genetic determinants of the brain with mostest. Nat Commun. 2020;11:1–9.
71. Zhao B, Li T, Yang Y, Wang X, Luo T, Shan Y, et al. Common genetic variation influencing human white matter microstructure. Science. 2021;372:eabf3736.
72. Zhao B, Luo T, Li T, Li Y, Zhang J, Shan Y, et al. Genome-wide association analysis of 19,629 individuals identifies variants influencing regional brain volumes and refines their genetic co-architecture with cognitive and mental health traits. Nat Genet. 2019;51:1637–44.
73. Donati G, Dumontheil I, Pain O, Asbury K, Meaburn EL. Evidence for specificity of polygenic contributions to attainment in english, maths and science during adolescence. Sci Rep. 2021;11:1–11.
74. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet. 2018;50:1112–21.
75. Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, De Leeuw CA, et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nat Genet. 2018;50:912–9.
76. Hill WD, Marioni RE, Maghzian O, Ritchie SJ, Hagenaars SP, McIntosh A, et al. A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. Mol Psychiatry. 2019;24:169–81.
77. Demange PA, Malanchini M, Mallard TT, Biroli P, Cox SR, Grotzinger AD, et al. Investigating the genetic architecture of noncognitive skills using gwas-by-subtraction. Nat Genet. 2021;53:35–44.
78. Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD, et al. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nat Commun. 2018;9:1–16.
79. Ge T, Chen C-Y, Doyle AE, Vettermann R, Tuominen LJ, Holt DJ, et al. The shared genetic basis of educational attainment and cerebral cortical morphology. Cereb Cortex. 2019;29:3471–81.
80. Mitchell BL, Cuéllar-Partida G, Grasby KL, Campos AI, Strike LT, Hwang L-D, et al. Educational attainment polygenic scores are associated with cortical total surface area and regions important for language and memory. Neuroimage. 2020;212:116691.

81. Linnér K, Biroli P, Kong E, Meddens SFW, Wedow R, Fontana MA, et al. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. Nat Genet. 2019;51:245–57.
82. Karlsson Linnér R, Biroli P, Kong E, Meddens SFW, Wedow R, Fontana MA, et al. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. Nat Genet. 2019;51:245–57.
83. Goes FS, McGrath J, Avramopoulos D, Wolyniec P, Pirooznia M, Ruczinski I, et al. Genome-wide association study of schizophrenia in Ashkenazi Jews. Am J Med Genet Part B: Neuropsychiatr Genet. 2015;168:649–59.
84. Lam M, Chen C-Y, Li Z, Martin AR, Bryois J, Ma X, et al. Comparative genetic architectures of schizophrenia in east asian and european populations. Nat Genet. 2019;51:1670–8.
85. Pardiñas AF, Holmans P, Pocklington AJ, Escott-Price V, Ripke S, Carrera N, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat Genet. 2018;50:381–9.
86. Sullivan PF, Agrawal A, Bulik CM, Andreassen OA, Børglum AD, Breen G, et al. Psychiatric genomics: an update and an agenda. Am J Psychiatry. 2018;175:15–27.
87. Wu Y, Cao H, Baranova A, Huang H, Li S, Cai L, et al. Multi-trait analysis for genome-wide association study of five psychiatric disorders. Transl psychiatry. 2020;10:1–11.
88. Yao X, Glessner JT, Li J, Qi X, Hou X, Zhu C, et al. Integrative analysis of genome-wide association studies identifies novel loci associated with neuropsychiatric disorders. Transl psychiatry. 2021;11:1–12.
89. Smeland OB, Shadrin A, Bahrami S, Broce I, Tesli M, Frei O, et al. Genome-wide association analysis of parkinson's disease and schizophrenia reveals shared genetic architecture and identifies novel risk loci. Biol psychiatry. 2021;89:227–35.
90. Pickrell JK, Berisa T, Liu JZ, Ségurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. Nat Genet. 2016;48:709–17.
91. van de Vegte YJ, Said MA, Rienstra M, van der Harst P, Verweij N. Genome-wide association studies and mendelian randomization analyses for leisure sedentary behaviours. Nat Commun. 2020;11:1–10.
92. de la Fuente J, Davies G, Grotzinger AD, Tucker-Drob EM, Deary IJ. A general dimension of genetic sharing across diverse cognitive traits inferred from molecular data. Nat Hum Behav. 2021;5:49–58.
93. Takayanagi Y, Sasabayashi D, Takahashi T, Furuichi A, Kido M, Nishikawa Y, et al. Reduced cortical thickness in schizophrenia and schizotypal disorder. Schizophr Bull. 2020;46:387–94.
94. Siddarth P, Burggren AC, Eyre HA, Small GW, Merrill DA. Sedentary behavior associated with reduced medial temporal lobe thickness in middle-aged and older adults. PLoS One. 2018;13:e0195549.
95. Maier RM, Visscher PM, Robinson MR, Wray NR. Embracing polygenicity: a review of methods and tools for psychiatric genetics research. Psychol Med. 2018;48:1055–67.
96. Munji RN, Choe Y, Li G, Siegenthaler JA, Pleasure SJ. Wnt signaling regulates neuronal differentiation of cortical intermediate progenitors. J Neurosci. 2011;31:1676–87.
97. Rakic P. Evolution of the neocortex: a perspective from developmental biology. Nat Rev Neurosci. 2009;10:724–35.
98. Polimanti R, Ratanatharathorn A, Maihofer AX, Choi KW, Stein MB, Morey RA, et al. Association of economic status and educational attainment with post-traumatic stress disorder: a mendelian randomization study. JAMA Netw open. 2019;2:e193447.
99. García-Marín LM, Reyes-Pérez P, Diaz-Torres S, Medina-Rivera A, Martin NG, Mitchell BL, et al. Shared molecular genetic factors influence subcortical brain morphometry and parkinson's disease risk. npj Parkinson's Dis. 2023;9:73.
100. Williams ME, Elman JA, Bell TR, Dale AM, Eyler LT, Fennema-Notestine C, et al. Higher cortical thickness/volume in alzheimer's-related regions: protective factor or risk factor? Neurobiol Aging. 2023;129:185–94.

AUTHOR CONTRIBUTIONS

RAM contributed to obtaining funding, writing the manuscript, and conceptual design. YZ, HB, DS contributed to data analysis. MEG and CCH contributed to data analysis and revising the manuscript. AXM contributed to data analysis and revising the manuscript. MG contributed to statistical modeling and data analysis. AXM contributed to data analysis and conceptual design. CLB contributed to data curation and revising the manuscript. KLG, PMT, SM, KLG, DEG, MSP, WSK, and CMN contributed to the conceptual design and revising of the manuscript. AAH contributed to revising the manuscript. AEAK contributed to obtaining funding, conceptual design, and revising the manuscript. MWL contributed to obtaining funding, data analysis, writing the manuscript, conceptual design, and scientific leadership.

FUNDING

National Institute for Mental Health Grant No. R01-MH111671, R01-MH129832, and VISN6 MIRECC (to RAM); VA Merit Grant Nos. 1I01RX000389-01 (to RAM) and 1I01CX000748-01A1 (to RAM); National Institute of Neurological Disorders and Stroke Grant Nos. R01-NS086885 and K23 MH073091-01 (to RAM); National Health and Medical Research Council APP1173025 (to KLG). VA Career Development Award #1IK2CX002107 - US Department of Veterans Affairs CSR&D. ENIGMA was supported partly by NIH U54 EB020403 from the Big Data to Knowledge (BD2K) program, R56AG058854, R01MH116147, and P41 EB015922 (to PMT); NIMH R01MH106595 (to CMN). We thank Cohen Veterans Bioscience for ongoing support and building a collaborative scientific environment. We thank all members of the respective site laboratories within the ENIGMA who contributed to general study organization, recruitment, data collection, and management, as well as subsequent analyses. Most importantly, we thank all of our study participants for their efforts to take part in this study. The funding agencies had no part in the analysis of data or approval of the final publication. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the US government.

COMPETING INTERESTS

Dr. Thompson received partial research support from Biogen, Inc. (Boston, USA) for research unrelated to the topic of this manuscript. The remaining authors declare no competing interests. The material presented is original research that has not been previously published and has not been submitted for publication elsewhere.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41398-024-03152-y.

Correspondence and requests for materials should be addressed to Mark W. Logue.

Reprints and permission information is available at http://www.nature.com/reprints

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