Genomic Structural Equation Modeling Reveals Latent Phenotypes in the Human Cortex with Distinct Genetic Architecture

Rajendra Morey（✉ rajendra.morey@duke.edu ）
Duke University

Yuanchao Zheng
https://orcid.org/0000-0001-7571-3054

Delin Sun
Duke University Medical Center   https://orcid.org/0000-0003-3283-423X

Melanie Garrett

Marianna Gasperi

Adam Maihofer
UCSD

C. Lexi Baird

Katrina Grasby
QIMR Berghofer Medical Research Institute   https://orcid.org/0000-0001-8539-0228

Ashley Huggins
Duke University   https://orcid.org/0000-0001-7936-2116

Courtney Haswell

Paul Thompson
Imaging Genetics Center, Mark and Mary Stevens Neuroimaging & Informatics Institute, Keck School of Medicine of the University of Southern California, Marina del Rey, California, USA

Sarah Medland
QIMR Berghofer Medical Research Institute   https://orcid.org/0000-0003-1382-380X

Daniel Gustavson
Vanderbilt University Medical Center   https://orcid.org/0000-0002-1470-4928

Matthew Panizzon

William Kremen

Caroline Nievergelt

Allison Ashley-Koch
Duke University   https://orcid.org/0000-0001-5409-9155

Logue Logue

Article

Abstract

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 the 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. The multivariate GWASs of these 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 with attention deficit hyperactivity disorder (ADHD), major depressive disorder (MDD), and insomnia, indicating genetic predisposition to a larger SA in the specific GIBN is associated with lower genetic risk of these disorders. CT GIBNs displayed a negative genetic correlation with alcohol dependence. Jointly modeling the genetic architecture of complex traits and investigating multivariate genetic links across phenotypes offers a new vantage point for mapping the cortex into genetically informed networks.

1. 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 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] and therefore offer an incomplete characterization of cortical pleiotropy. 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 where 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 a 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 12 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 would reflect the complex relationships captured by 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 hypothesize that genetics may influence psychiatric illness by initially acting to disrupt brain structures, or impart vulnerabilities that lead to disrupted brain structures in the presence (or absence) of specific environmental exposures [28–30]. We further hypothesized genetic correlations between GIBNs and major neuropsychiatric disorders that are stronger than correlations between global measures of SA or CT.

2. METHODS

2.1 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. In that study, 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[31]. 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, 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[16,18,32].

2.2 *Analysis*

Our analyses were performed using the genomic-SEM 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. To avoid overfitting, we 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 **Figures S1** and **S2**). 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 [33]. As there was no consensus on factor loading cutoffs [19,34], 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]

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. As opposed to regression modeling, where significant statistics represent the strength of association between the predictors in the response, in SEM modeling, a significant χ^2 statistic represents a lack of fit. 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 ~ 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[35]. We additionally fit a bifactor 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 and a multi-level model where all EFA factors loaded onto a 2nd order factor. The bifactor models failed to converge and the multilevel models failed to improve model fit in all cases; hence these results are not reported.

2.3 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 1,000 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 1,000 Dice's coefficients. The relative position of the 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 [36]. We quantitatively analyzed the overlap between GIBNs and networks based on 20 neuroreceptor density maps of Hansen et al [37]. 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).

2.4 Multivariate GWAS Analysis

Using the GIBNs from our best fitting model, we used gSEM to generate a multivariate GWAS of each GIBN. The GWS associations ($p < 5x10^{-8}$) 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[38] 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 1000G 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.

2.5 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[17,39–44]. We also looked for SNPs that were significant in GWASs of 12 neuropsychiatric disorders from the Psychiatric Genomics Consortium (PGC): ADHD[45], alcohol dependence[46], anorexia nervosa[47], autism spectrum disorder[48], bipolar[49], cannabis use[50], MDD[51], obsessive compulsive disorder (OCD)[52] posttraumatic stress disorder (PTSD)[53], schizophrenia[54], Tourette's syndrome[55], and anxiety[56]. Finally, FUMA was used to functionally annotate loci that overlapped with previously published GWAS results.

2.6 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[57]. We estimated the genetic correlation between CT- and SA-derived GIBNs and neuropsychiatric disorders using their GWAS summary statistics[57]. 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).

2.7 Data and Code 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/.

3. RESULTS

3.1 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 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 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.

3.2 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).

3.3 Overlap of GIBNs and RSNs

The Dice's coefficients and corresponding p-values of GIBNs and the Yeo et al. (2011) 7 RSNs (Fig. 4) showed that CT4 and SA5 had relatively high overlap with the visual network (CT4:DC = 0.353,p_{FDR}=.008; SA5:DC = 0.424,p_{FDR}=.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.

3.4 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).

3.6 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 **Figures S3** to **S12**. 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 [38] 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[58] of the genetic associations is presented in Fig. 6. A list of all GWS loci is provided in **Table S7.** 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 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.

3.7 *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 S10**), 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.

3.8 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 S8**. 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 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 S9**). 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).

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 S8** and **S9**.

4. 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 applying 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. However, in one respect, the present results are more informative than Chen et al.[11], as we provide SNPs associated with each of the GIBNs.

The overlap of GIBN GWS loci with prior GWAS of neuroimaging phenotypes or psychiatric disorders firmly points to the relevance of GIBN-related variants to brain structure and cognition. First, we note that novel variant rs3006933 has been previously associated with subcortical volumes[59]. Novel variants rs3006933 and rs1004763[17,42] have been associated with neuroimaging phenotypes of corpus callosum white matter microstructure[60]. 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 [54]. 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 = 5.22×10^{-17} to 8.45×10^{-21}). We also observed an association between CT1 and rs11692435 (p = 1.17×10^{-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[61] 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[62], the SA3-linked SNP rs2299148[63], and the SA6-linked locus rs9909861[63–67] 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 [68]. The same regions were reported independently in the Queensland Twin Imaging and HCP samples [69]. The SA5-linked locus rs6701689 has been reported for risk tolerance[70]. 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[70,71]. The CT4-associated locus rs13107325 has been associated with many traits including schizophrenia[72–79], bipolar disorder[76,77], Parkinson's disease[78,79], sedentary behavior[59,80] and risk taking[70], as well as cognition, intelligence, and educational attainment[63–67,81]. CT4 includes the parahippocampal and fusiform gyri, which have firmly established links to schizophrenia[82] and sedentary behavior[83].

Resting-state functional connectivity evinces robust patterns of synchronous activity that intrinsically organize into canonical networks [84]. We found that 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[36]. 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)[85]. SA6, which includes superior and transverse temporal cortices, overlaps the auditory cortex from twin-derived genetic parcels (Dice's coefficient = 0.211; Fig. 5). The functional specializations of the human auditory cortex[36], which include parts of the lateral prefrontal cortex, Broca's area, and subcentral regions, are needed for human vocalization and language[86,87]. The dorsal attention network (DAN), which directs voluntary allocation of attention, has substantial overlap with SA1 (Dice's coefficient = 0.232, Fig. 5) 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)[88]. Since FEF is not a FreeSurfer parcellation output, it may be poorly represented in ENIGMA cortical GWAS. The DAN has relatively high twin heritability estimates (left h^2 = 0.45; right h^2 = 0.40)[85]. 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[89] (Fig. 6).

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[90].

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 [32] 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 [32] 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 variant. Genetic variation affecting gene regulation in progenitor cell types, present in fetal development, affects adult cortical SA[91]. 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)[92]. 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[93].

4.1 Limitations

A number of limitations deserve consideration in interpreting the present findings. 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 [17,18] 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. However, simultaneously analyzing GWAS results for each of 100s or 1000s of vertices would present a computational challenge. Therefore, the GIBNs 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.

4.2 Conclusion

We harnessed the pervasive pleiotropy of the human cortex to realize a unique 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.

Declarations

SUPPORT

CONFLICTS OF INTEREST

References

1. Zielinski, B.A., Gennatas, E.D., Zhou, J., and Seeley, W.W. (2010). Network-level structural covariance in the developing brain. Proceedings of the National Academy of Sciences *107*, 18191-18196.

2. Romero-Garcia, R., Whitaker, K.J., Váša, F., Seidlitz, J., Shinn, M., Fonagy, P., Dolan, R.J., Jones, P.B., Goodyer, I.M., and Bullmore, E.T. (2018). Structural covariance networks are coupled to expression of genes enriched in supragranular layers of the human cortex. Neuroimage *171*, 256-267.

3. Feng, J., Chen, C., Cai, Y., Ye, Z., Feng, K., Liu, J., Zhang, L., Yang, Q., Li, A., and Sheng, J. (2020). Partitioning heritability analyses unveil the genetic architecture of human brain multidimensional functional connectivity patterns. Human brain mapping *41*, 3305-3317.

4. Hawrylycz, M., Miller, J.A., Menon, V., Feng, D., Dolbeare, T., Guillozet-Bongaarts, A.L., Jegga, A.G., Aronow, B.J., Lee, C.-K., and Bernard, A. (2015). Canonical genetic signatures of the adult human brain. Nature neuroscience *18*, 1832-1844.

5. Gong, G., He, Y., Chen, Z.J., and Evans, A.C. (2012). Convergence and divergence of thickness correlations with diffusion connections across the human cerebral cortex. Neuroimage *59*, 1239-1248.

6. Alexander-Bloch, A., Giedd, J.N., and Bullmore, E. (2013). Imaging structural co-variance between human brain regions. Nature Reviews Neuroscience *14*, 322-336.

7. Seeley, W.W., Crawford, R.K., Zhou, J., Miller, B.L., and Greicius, M.D. (2009). Neurodegenerative diseases target large-scale human brain networks. Neuron *62*, 42-52.

8. He, Y., Chen, Z., and Evans, A. (2008). Structural insights into aberrant topological patterns of large-scale cortical networks in Alzheimer's disease. Journal of Neuroscience *28*, 4756-4766.

9. Segall, J.M., Allen, E.A., Jung, R.E., Erhardt, E.B., Arja, S.K., Kiehl, K.A., and Calhoun, V.D. (2012). Correspondence between structure and function in the human brain at rest. Frontiers in neuroinformatics *6*, 10.

10. Zhang, Z., Liao, W., Zuo, X.-N., Wang, Z., Yuan, C., Jiao, Q., Chen, H., Biswal, B.B., Lu, G., and Liu, Y. (2011). Resting-state brain organization revealed by functional covariance networks. PLoS One *6*, e28817.

11. Chen, C.-H., Gutierrez, E., Thompson, W., Panizzon, M.S., Jernigan, T.L., Eyler, L.T., Fennema-Notestine, C., Jak, A.J., Neale, M.C., and Franz, C.E. (2012). Hierarchical genetic organization of human cortical surface area. Science *335*, 1634-1636.

12. Lenroot, R.K., Schmitt, J.E., Ordaz, S.J., Wallace, G.L., Neale, M.C., Lerch, J.P., Kendler, K.S., Evans, A.C., and Giedd, J.N. (2009). Differences in genetic and environmental influences on the human cerebral cortex associated with development during childhood and adolescence. Human brain mapping *30*, 163-174.

13. Schmitt, J., Lenroot, R., Wallace, G., Ordaz, S., Taylor, K., Kabani, N., Greenstein, D., Lerch, J., Kendler, K., and Neale, M. (2008). Identification of genetically mediated cortical networks: a multivariate study of pediatric twins and siblings. Cerebral cortex *18*, 1737-1747.

14. Tichenor, M., and Sridhar, D. (2019). 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 Research *4*.

15. Rowland, T.A., and Marwaha, S. (2018). Epidemiology and risk factors for bipolar disorder. Therapeutic advances in psychopharmacology *8*, 251-269.

16. van der Meer, D., Frei, O., Kaufmann, T., Chen, C.-H., Thompson, W.K., O'Connell, K.S., Monereo Sánchez, J., Linden, D.E., Westlye, L.T., and Dale, A.M. (2020). Quantifying the polygenic architecture of the human cerebral cortex: extensive genetic overlap between cortical thickness and surface area. Cerebral Cortex *30*, 5597-5603.

17. Smith, S.M., Douaud, G., Chen, W., Hanayik, T., Alfaro-Almagro, F., Sharp, K., and Elliott, L.T. (2021). An expanded set of genome-wide association studies of brain imaging phenotypes in UK Biobank. Nature neuroscience *24*, 737-745.

18. Grasby, K.L., Jahanshad, N., Painter, J.N., Colodro-Conde, L., Bralten, J., Hibar, D.P., Lind, P.A., Pizzagalli, F., Ching, C.R., and McMahon, M.A.B. (2020). The genetic architecture of the human cerebral cortex. Science *367*.

19. Grotzinger, A.D., Rhemtulla, M., de Vlaming, R., Ritchie, S.J., Mallard, T.T., Hill, W.D., Ip, H.F., Marioni, R.E., McIntosh, A.M., and Deary, I.J. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. Nature human behaviour *3*, 513-525.

20. Kovas, Y., and Plomin, R. (2006). Generalist genes: implications for the cognitive sciences. Trends in cognitive sciences *10*, 198-203.

21. McTeague, L.M., Rosenberg, B.M., Lopez, J.W., Carreon, D.M., Huemer, J., Jiang, Y., Chick, C.F., Eickhoff, S.B., and Etkin, A. (2020). Identification of common neural circuit disruptions in emotional processing across psychiatric disorders. American Journal of Psychiatry *177*, 411-421.

22. Fornito, A., Bullmore, E.T., and Zalesky, A. (2017). Opportunities and challenges for psychiatry in the connectomic era. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging *2*, 9-19.

23. Gandal, M.J., Haney, J.R., Parikshak, N.N., Leppa, V., Ramaswami, G., Hartl, C., Schork, A.J., Appadurai, V., Buil, A., and Werge, T.M. (2018). Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. Science *359*, 693-697.

24. Grotzinger, A.D. (2021). Shared genetic architecture across psychiatric disorders. Psychological Medicine *51*, 2210-2216.

25. Ward, J., Tunbridge, E.M., Sandor, C., Lyall, L.M., Ferguson, A., Strawbridge, R.J., Lyall, D.M., Cullen, B., Graham, N., and Johnston, K.J. (2020). 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. Molecular psychiatry *25*, 3091-3099.

26. Consortium, C.-D.G.o.t.P.G. (2013). Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. The Lancet *381*, 1371-1379.

27. Lee, P.H., Feng, Y.-C.A., and Smoller, J.W. (2021). Pleiotropy and cross-disorder genetics among psychiatric disorders. Biological psychiatry *89*, 20-31.

28. Egervari, G., Kozlenkov, A., Dracheva, S., and Hurd, Y.L. (2019). Molecular windows into the human brain for psychiatric disorders. Molecular psychiatry *24*, 653-673.

29. Heck, A., Fastenrath, M., Ackermann, S., Auschra, B., Bickel, H., Coynel, D., Gschwind, L., Jessen, F., Kaduszkiewicz, H., and Maier, W. (2014). Converging genetic and functional brain imaging evidence links neuronal excitability to working memory, psychiatric disease, and brain activity. Neuron *81*, 1203-1213.

30. Sullivan, P.F., and Geschwind, D.H. (2019). Defining the genetic, genomic, cellular, and diagnostic architectures of psychiatric disorders. Cell *177*, 162-183.

31. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., and Hyman, B.T. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage *31*, 968-980.

32. Panizzon, M.S., Fennema-Notestine, C., Eyler, L.T., Jernigan, T.L., Prom-Wormley, E., Neale, M., Jacobson, K., Lyons, M.J., Grant, M.D., Franz, C.E., et al. (2009). Distinct genetic influences on cortical surface area and cortical thickness. Cereb Cortex *19*, 2728-2735. 10.1093/cercor/bhp026.

33. Howard, M.C. (2016). 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 *32*, 51-62. 10.1080/10447318.2015.1087664.

34. de Vries, A., Tiemens, B., Cillessen, L., and Hutschemaekers, G. (2021). Construction and validation of a self-direction measure for mental health care. Journal of Clinical Psychology *77*, 1371-1383.

35. Streiner, D.L. (2006). Building a better model: An introduction to structural equation modelling. Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie *51*, 317-324. Doi 10.1177/070674370605100507.

36. Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zoller, L., Polimeni, J.R., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of Neurophysiology *106*, 1125-1165. DOI 10.1152/jn.00338.2011.

37. Hansen, J.Y., Markello, R.D., Tuominen, L., Nørgaard, M., Kuzmin, E., Palomero-Gallagher, N., Dagher, A., and Misic, B. (2022). Correspondence between gene expression and neurotransmitter receptor and transporter density in the human brain. Neuroimage *264*, 119671.

38. Watanabe, K., Taskesen, E., Van Bochoven, A., and Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. Nature communications *8*, 1-11.

39. Shadrin, A.A., Kaufmann, T., van der Meer, D., Palmer, C.E., Makowski, C., Loughnan, R., Jernigan, T.L., Seibert, T.M., Hagler, D.J., and Smeland, O.B. (2021). Vertex-wise multivariate genome-wide association study identifies 780 unique genetic loci associated with cortical morphology. NeuroImage *244*, 118603.

40. Hofer, E., Roshchupkin, G.V., Adams, H.H., Knol, M.J., Lin, H., Li, S., Zare, H., Ahmad, S., Armstrong, N.J., and Satizabal, C.L. (2020). Genetic correlations and genome-wide associations of cortical structure in general population samples of 22,824 adults. Nature communications *11*, 1-16.

41. Adams, H.H., Hibar, D.P., Chouraki, V., Stein, J.L., Nyquist, P.A., Rentería, M.E., Trompet, S., Arias-Vasquez, A., Seshadri, S., and Desrivières, S. (2016). Novel genetic loci underlying human intracranial volume identified through genome-wide association. Nature neuroscience *19*, 1569-1582.

42. Elliott, L.T., Sharp, K., Alfaro-Almagro, F., Shi, S., Miller, K.L., Douaud, G., Marchini, J., and Smith, S.M. (2018). Genome-wide association studies of brain imaging phenotypes in UK Biobank. Nature *562*, 210-216.

43. Arfan Ikram, M., Fornage, M., Smith, A.V., Seshadri, S., Schmidt, R., Debette, S., Vrooman, H.A., Sigurdsson, S., Ropele, S., and Taal, H.R. (2012). Common variants at 6q22 and 17q21 are associated with intracranial volume. Nature genetics *44*, 539-544.

44. Hibar, D.P., Stein, J.L., Renteria, M.E., Arias-Vasquez, A., Desrivieres, S., Jahanshad, N., Toro, R., Wittfeld, K., Abramovic, L., Andersson, M., et al. (2015). Common genetic variants influence human subcortical brain structures. Nature. 10.1038/nature14101.

45. Demontis, D., Walters, R.K., Martin, J., Mattheisen, M., Als, T.D., Agerbo, E., Baldursson, G., Belliveau, R., Bybjerg-Grauholm, J., and Bækvad-Hansen, M. (2019). Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nature genetics *51*, 63.

46. Walters, R.K., Polimanti, R., Johnson, E.C., McClintick, J.N., Adams, M.J., Adkins, A.E., Aliev, F., Bacanu, S.-A., Batzler, A., and Bertelsen, S. (2018). Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. Nature neuroscience *21*, 1656-1669.

47. Watson, H.J., Yilmaz, Z., Thornton, L.M., Hübel, C., Coleman, J.R., Gaspar, H.A., Bryois, J., Hinney, A., Leppä, V.M., and Mattheisen, M. (2019). Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. Nature genetics *51*, 1207-1214.

48. Grove, J., Ripke, S., Als, T.D., Mattheisen, M., Walters, R.K., Won, H., Pallesen, J., Agerbo, E., Andreassen, O.A., and Anney, R. (2019). Identification of common genetic risk variants for autism spectrum disorder. Nature genetics *51*, 431-444.

49. Stahl, E.A., Breen, G., Forstner, A.J., McQuillin, A., Ripke, S., Trubetskoy, V., Mattheisen, M., Wang, Y., Coleman, J.R., and Gaspar, H.A. (2019). Genome-wide association study identifies 30 loci associated with bipolar disorder. Nature genetics *51*, 793-803.

50. Johnson, E.C., Demontis, D., Thorgeirsson, T.E., Walters, R.K., Polimanti, R., Hatoum, A.S., Sanchez-Roige, S., Paul, S.E., Wendt, F.R., and Clarke, T.-K. (2020). A large-scale genome-wide association

study meta-analysis of cannabis use disorder. The Lancet Psychiatry *7*, 1032-1045.

51. Howard, D.M., Adams, M.J., Clarke, T.-K., Hafferty, J.D., Gibson, J., Shirali, M., Coleman, J.R., Hagenaars, S.P., Ward, J., and Wigmore, E.M. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. Nature neuroscience *22*, 343-352.

52. Arnold, P.D., Askland, K.D., Barlassina, C., Bellodi, L., Bienvenu, O., Black, D., Bloch, M., Brentani, H., Burton, C.L., and Camarena, B. (2018). Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. Molecular psychiatry *23*, 1181-1181.

53. Maihofer, A.X., Choi, K.W., Coleman, J.R., Daskalakis, N.P., Denckla, C.A., Ketema, E., Morey, R.A., Polimanti, R., Ratanatharathorn, A., and Torres, K. (2021). Enhancing Discovery of Genetic Variants for Posttraumatic Stress Disorder Through Integration of Quantitative Phenotypes and Trauma Exposure Information. Biological psychiatry.

54. Trubetskoy, V., Pardiñas, A.F., Qi, T., Panagiotaropoulou, G., Awasthi, S., Bigdeli, T.B., Bryois, J., Chen, C.-Y., Dennison, C.A., and Hall, L.S. (2022). Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature *604*, 502-508.

55. Yu, D., Sul, J.H., Tsetsos, F., Nawaz, M.S., Huang, A.Y., Zelaya, I., Illmann, C., Osiecki, L., Darrow, S.M., and Hirschtritt, M.E. (2019). Interrogating the genetic determinants of Tourette's syndrome and other tic disorders through genome-wide association studies. American Journal of Psychiatry *176*, 217-227.

56. Otowa, T., Hek, K., Lee, M., Byrne, E.M., Mirza, S.S., Nivard, M.G., Bigdeli, T., Aggen, S.H., Adkins, D., and Wolen, A. (2016). Meta-analysis of genome-wide association studies of anxiety disorders. Molecular psychiatry *21*, 1391-1399.

57. Bulik-Sullivan, B.K., Loh, P.-R., Finucane, H.K., Ripke, S., Yang, J., Patterson, N., Daly, M.J., Price, A.L., Neale, B.M., and Consortium, S.W.G.o.t.P.G. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nature genetics *47*, 291-295.

58. Wolfe, D., Dudek, S., Ritchie, M.D., and Pendergrass, S.A. (2013). Visualizing genomic information across chromosomes with PhenoGram. BioData mining *6*, 1-12.

59. van der Meer, D., Frei, O., Kaufmann, T., Shadrin, A.A., Devor, A., Smeland, O.B., Thompson, W.K., Fan, C.C., Holland, D., and Westlye, L.T. (2020). Understanding the genetic determinants of the brain with MOSTest. Nature communications *11*, 1-9.

60. Zhao, B., Li, T., Yang, Y., Wang, X., Luo, T., Shan, Y., Zhu, Z., Xiong, D., Hauberg, M.E., and Bendl, J. (2021). Common genetic variation influencing human white matter microstructure. Science *372*, eabf3736.

61. Zhao, B., Luo, T., Li, T., Li, Y., Zhang, J., Shan, Y., Wang, X., Yang, L., Zhou, F., and Zhu, Z. (2019). 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. Nature genetics *51*, 1637-1644.

62. Donati, G., Dumontheil, I., Pain, O., Asbury, K., and Meaburn, E.L. (2021). Evidence for specificity of polygenic contributions to attainment in English, maths and science during adolescence. Scientific reports *11*, 1-11.

63. Lee, J.J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T.A., Bowers, P., Sidorenko, J., and Karlsson Linnér, R. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nature genetics *50*, 1112-1121.

64. Savage, J.E., Jansen, P.R., Stringer, S., Watanabe, K., Bryois, J., De Leeuw, C.A., Nagel, M., Awasthi, S., Barr, P.B., and Coleman, J.R. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nature genetics *50*, 912-919.

65. Hill, W.D., Marioni, R.E., Maghzian, O., Ritchie, S.J., Hagenaars, S.P., McIntosh, A., Gale, C.R., Davies, G., and Deary, I.J. (2019). A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. Molecular psychiatry *24*, 169-181.

66. Demange, P.A., Malanchini, M., Mallard, T.T., Biroli, P., Cox, S.R., Grotzinger, A.D., Tucker-Drob, E.M., Abdellaoui, A., Arseneault, L., and Van Bergen, E. (2021). Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. Nature genetics *53*, 35-44.

67. Davies, G., Lam, M., Harris, S.E., Trampush, J.W., Luciano, M., Hill, W.D., Hagenaars, S.P., Ritchie, S.J., Marioni, R.E., and Fawns-Ritchie, C. (2018). Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nature communications *9*, 1-16.

68. Ge, T., Chen, C.-Y., Doyle, A.E., Vettermann, R., Tuominen, L.J., Holt, D.J., Sabuncu, M.R., and Smoller, J.W. (2019). The shared genetic basis of educational attainment and cerebral cortical morphology. Cerebral Cortex *29*, 3471-3481.

69. Mitchell, B.L., Cuéllar-Partida, G., Grasby, K.L., Campos, A.I., Strike, L.T., Hwang, L.-D., Okbay, A., Thompson, P.M., Medland, S.E., and Martin, N.G. (2020). Educational attainment polygenic scores are associated with cortical total surface area and regions important for language and memory. Neuroimage *212*, 116691.

70. Linnér, K., Biroli, P., Kong, E., Meddens, S.F.W., Wedow, R., Fontana, M.A., Lebreton, M., Tino, S.P., Abdellaoui, A., and Hammerschlag, A.R. (2019). Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. Nature genetics *51*, 245-257.

71. Karlsson Linnér, R., Biroli, P., Kong, E., Meddens, S.F.W., Wedow, R., Fontana, M.A., Lebreton, M., Tino, S.P., Abdellaoui, A., and Hammerschlag, A.R. (2019). Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. Nature genetics *51*, 245-257.

72. Goes, F.S., McGrath, J., Avramopoulos, D., Wolyniec, P., Pirooznia, M., Ruczinski, I., Nestadt, G., Kenny, E.E., Vacic, V., and Peters, I. (2015). Genome-wide association study of schizophrenia in Ashkenazi Jews. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics *168*, 649-659.

73. Lam, M., Chen, C.-Y., Li, Z., Martin, A.R., Bryois, J., Ma, X., Gaspar, H., Ikeda, M., Benyamin, B., and Brown, B.C. (2019). Comparative genetic architectures of schizophrenia in East Asian and European populations. Nature genetics *51*, 1670-1678.

74. Pardiñas, A.F., Holmans, P., Pocklington, A.J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S.E., Bishop, S., Cameron, D., and Hamshere, M.L. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nature genetics *50*, 381-389.

75. Sullivan, P.F., Agrawal, A., Bulik, C.M., Andreassen, O.A., Børglum, A.D., Breen, G., Cichon, S., Edenberg, H.J., Faraone, S.V., and Gelernter, J. (2018). Psychiatric genomics: an update and an agenda. American Journal of Psychiatry *175*, 15-27.

76. Wu, Y., Cao, H., Baranova, A., Huang, H., Li, S., Cai, L., Rao, S., Dai, M., Xie, M., and Dou, Y. (2020). Multi-trait analysis for genome-wide association study of five psychiatric disorders. Translational psychiatry *10*, 1-11.

77. Yao, X., Glessner, J.T., Li, J., Qi, X., Hou, X., Zhu, C., Li, X., March, M.E., Yang, L., and Mentch, F.D. (2021). Integrative analysis of genome-wide association studies identifies novel loci associated with neuropsychiatric disorders. Translational psychiatry *11*, 1-12.

78. Smeland, O.B., Shadrin, A., Bahrami, S., Broce, I., Tesli, M., Frei, O., Wirgenes, K.V., O'Connell, K.S., Krull, F., and Bettella, F. (2021). Genome-wide association analysis of Parkinson's disease and schizophrenia reveals shared genetic architecture and identifies novel risk loci. Biological psychiatry *89*, 227-235.

79. Pickrell, J.K., Berisa, T., Liu, J.Z., Ségurel, L., Tung, J.Y., and Hinds, D.A. (2016). Detection and interpretation of shared genetic influences on 42 human traits. Nature genetics *48*, 709-717.

80. van de Vegte, Y.J., Said, M.A., Rienstra, M., van der Harst, P., and Verweij, N. (2020). Genome-wide association studies and Mendelian randomization analyses for leisure sedentary behaviours. Nature communications *11*, 1-10.

81. de la Fuente, J., Davies, G., Grotzinger, A.D., Tucker-Drob, E.M., and Deary, I.J. (2021). A general dimension of genetic sharing across diverse cognitive traits inferred from molecular data. Nature Human Behaviour *5*, 49-58.

82. Takayanagi, Y., Sasabayashi, D., Takahashi, T., Furuichi, A., Kido, M., Nishikawa, Y., Nakamura, M., Noguchi, K., and Suzuki, M. (2020). Reduced cortical thickness in schizophrenia and schizotypal disorder. Schizophrenia bulletin *46*, 387-394.

83. Siddarth, P., Burggren, A.C., Eyre, H.A., Small, G.W., and Merrill, D.A. (2018). Sedentary behavior associated with reduced medial temporal lobe thickness in middle-aged and older adults. PloS one *13*, e0195549.

84. Noble, S., Scheinost, D., and Constable, R.T. (2019). A decade of test-retest reliability of functional connectivity: A systematic review and meta-analysis. Neuroimage *203*, 116157.

85. Anderson, K.M., Ge, T., Kong, R., Patrick, L.M., Spreng, R.N., Sabuncu, M.R., Yeo, B.T., and Holmes, A.J. (2021). Heritability of individualized cortical network topography. Proceedings of the National Academy of Sciences *118*.

86. Forseth, K.J., Hickok, G., Rollo, P., and Tandon, N. (2020). Language prediction mechanisms in human auditory cortex. Nature communications *11*, 1-14.

87. Celesia, G.G. (1976). Organization of auditory cortical areas in man. Brain *99*, 403-414.

88. Mengotti, P., Käsbauer, A.-S., Fink, G.R., and Vossel, S. (2020). Lateralization, functional specialization, and dysfunction of attentional networks. Cortex *132*, 206-222.

89. Greene, C.M., and Soto, D. (2014). Functional connectivity between ventral and dorsal frontoparietal networks underlies stimulus-driven and working memory-driven sources of visual distraction. NeuroImage *84*, 290-298.

90. Maier, R.M., Visscher, P.M., Robinson, M.R., and Wray, N.R. (2018). Embracing polygenicity: a review of methods and tools for psychiatric genetics research. Psychological Medicine *48*, 1055-1067. 10.1017/S0033291717002318.

91. Munji, R.N., Choe, Y., Li, G., Siegenthaler, J.A., and Pleasure, S.J. (2011). Wnt signaling regulates neuronal differentiation of cortical intermediate progenitors. Journal of Neuroscience *31*, 1676-1687.

92. Rakic, P. (2009). Evolution of the neocortex: a perspective from developmental biology. Nature Reviews Neuroscience *10*, 724-735.

93. Polimanti, R., Ratanatharathorn, A., Maihofer, A.X., Choi, K.W., Stein, M.B., Morey, R.A., Logue, M.W., Nievergelt, C.M., Stein, D.J., and Koenen, K.C. (2019). Association of economic status and educational attainment with posttraumatic stress disorder: a Mendelian randomization study. JAMA network open *2*, e193447-e193447.
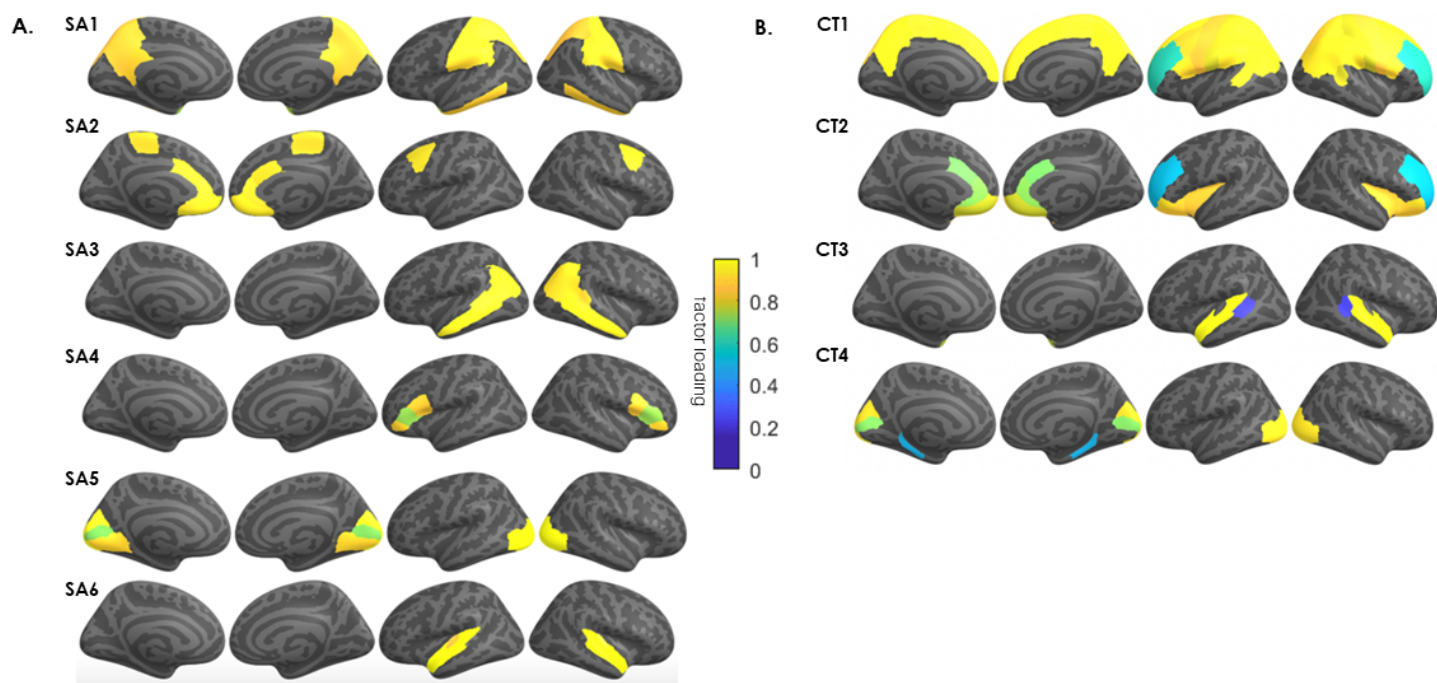
Figures

Figure 1

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 (2020). 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 ($p < 5\times10^{-8}$) associated with each GIBN.
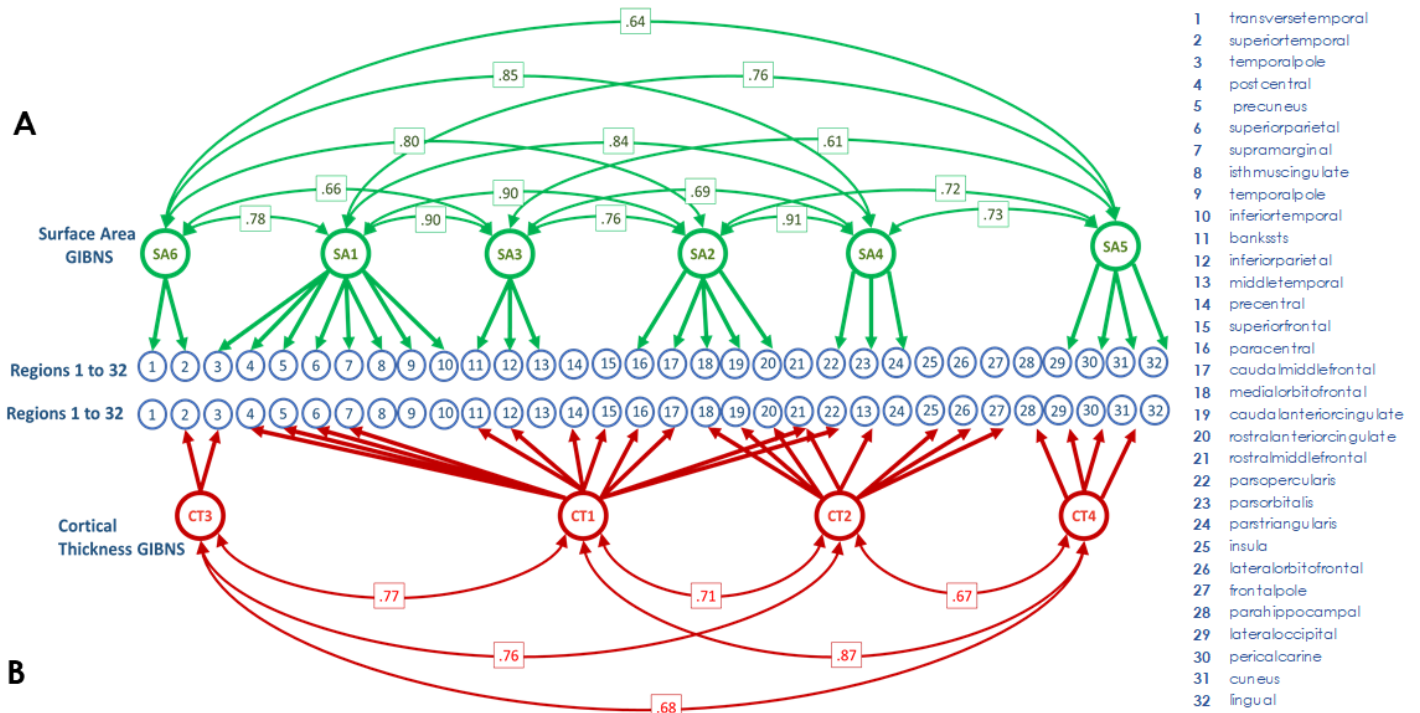
Figure 2

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, represent 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. Abbreviations: SA=surface area, CT=corticial thickness, GIBN=genetically informed brain network.
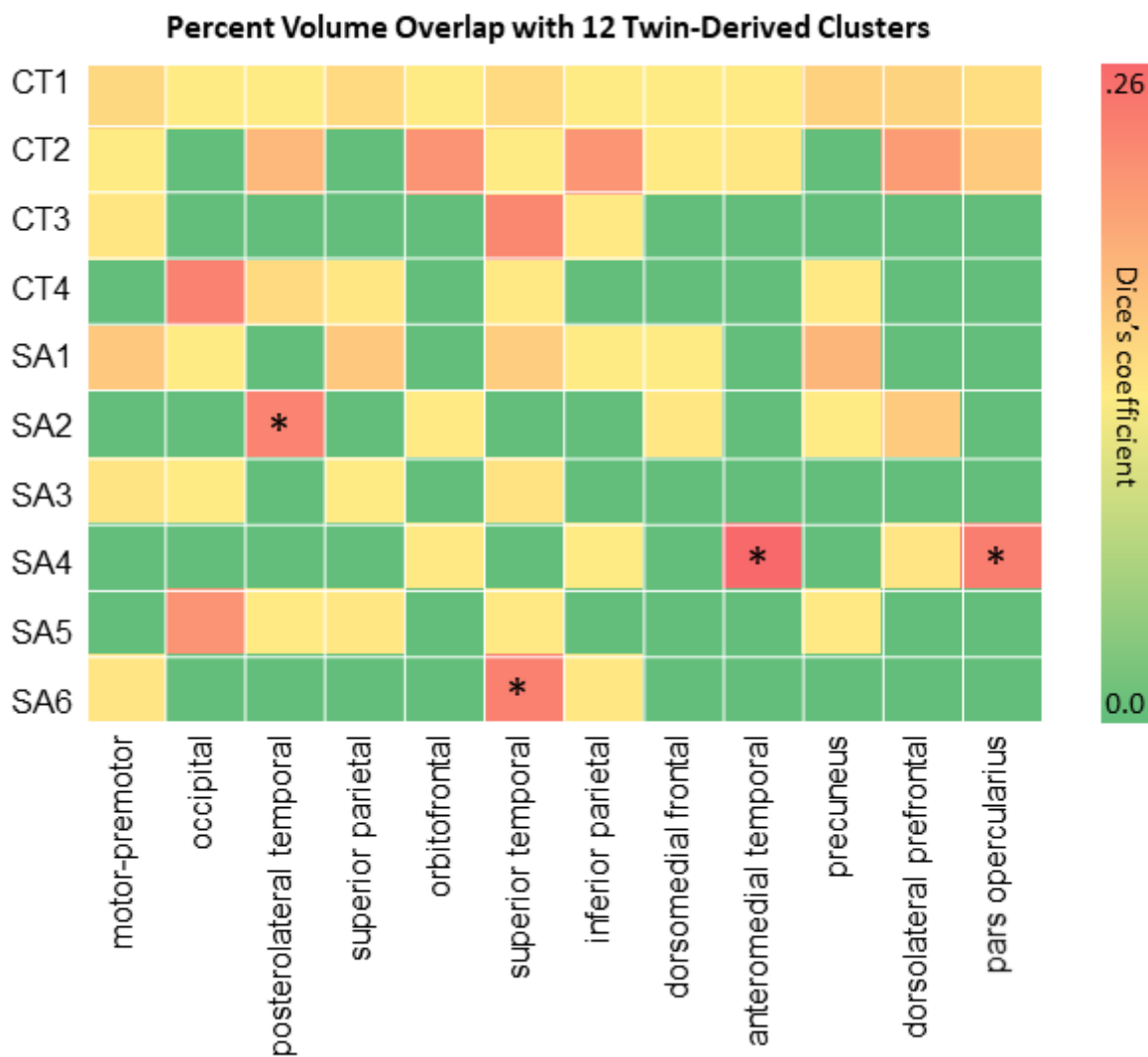
Percent Volume Overlap with 12 Twin-Derived Clusters

Figure 3

Genetic Parcellation of Twin Brains (A) 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. (2012) derived from twin brain data. *FDR-corrected p-values for percent volume overlap with 12 clusters are indicated by an asterisk. <u>Abbreviations</u>: SA=surface area, CT=cortical thickness, GIBN=genetically informed brain network.

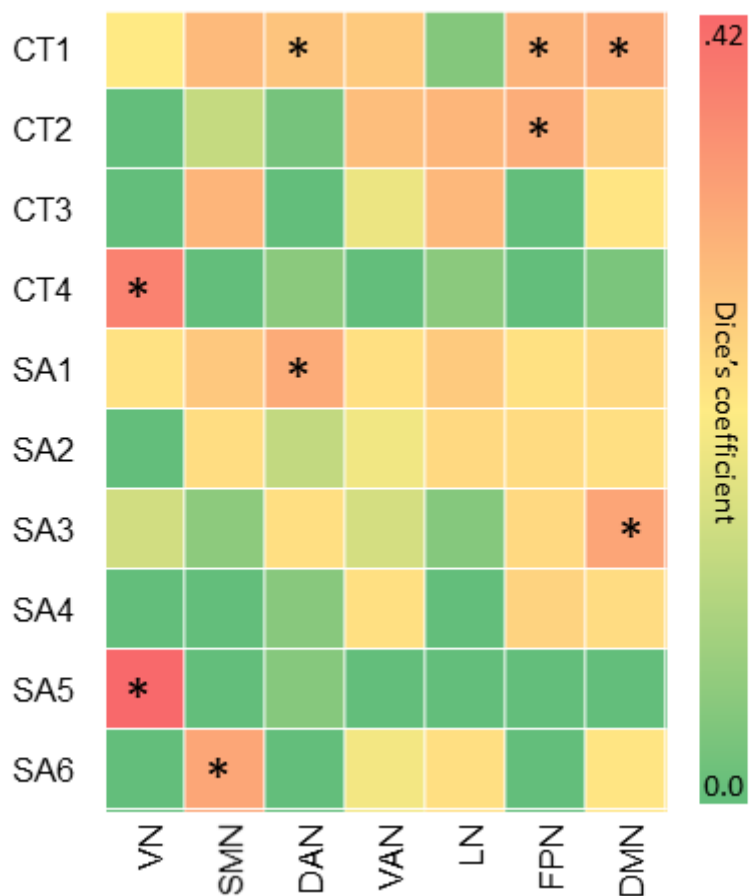**Percent Volume Overlap with 7 Canonical Resting Networks**

Figure 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. <u>Abbreviations</u>: 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.
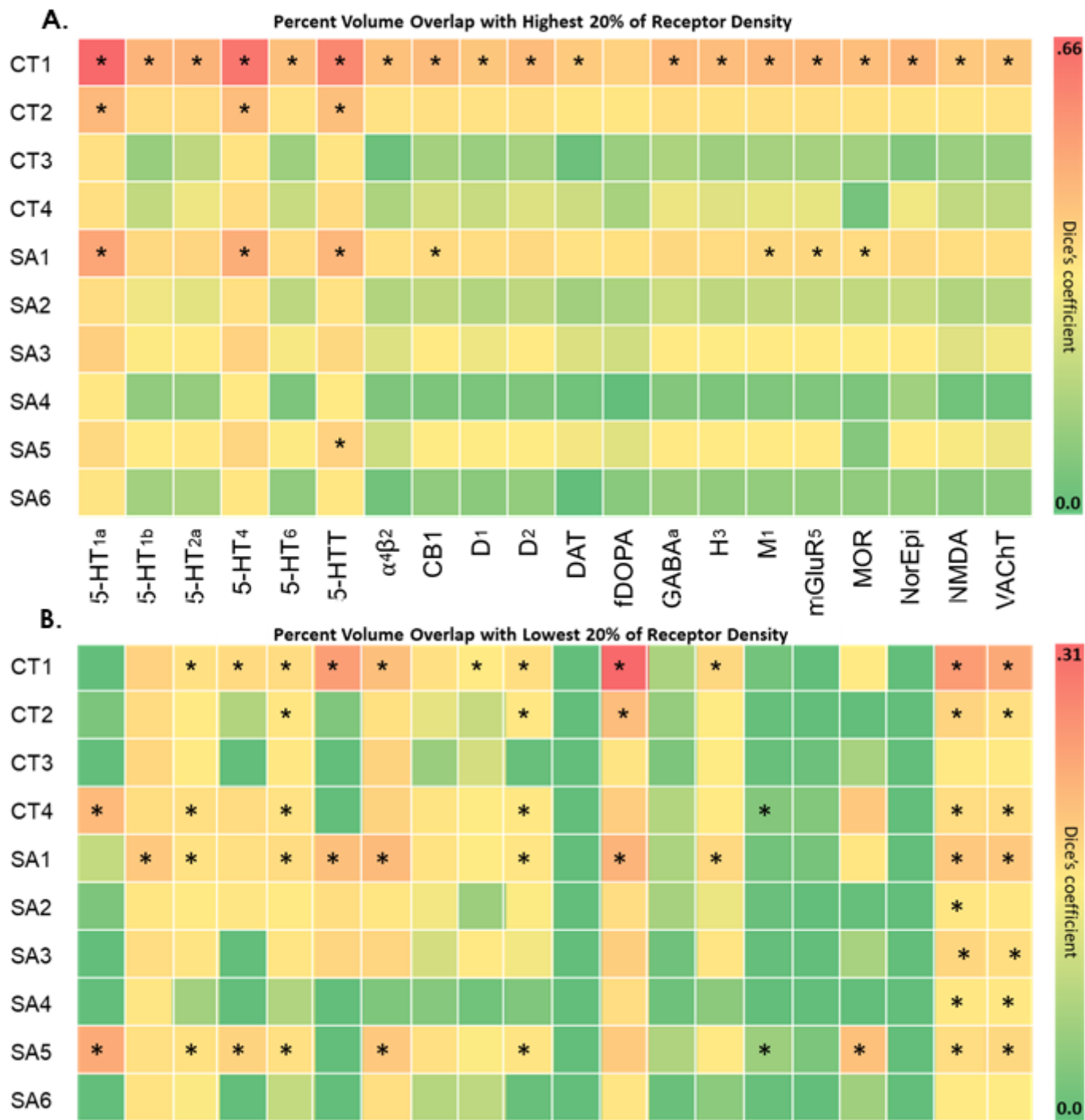
Figure 5

Receptor Density. Dice's coefficients and corresponding p-values between each 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 (2022) 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 X 10 GIBNs) are displayed in heatmaps. <u>Abbreviations</u>: SA=surface area, CT=cortical thickness, GIBN=genetically informed brain network.
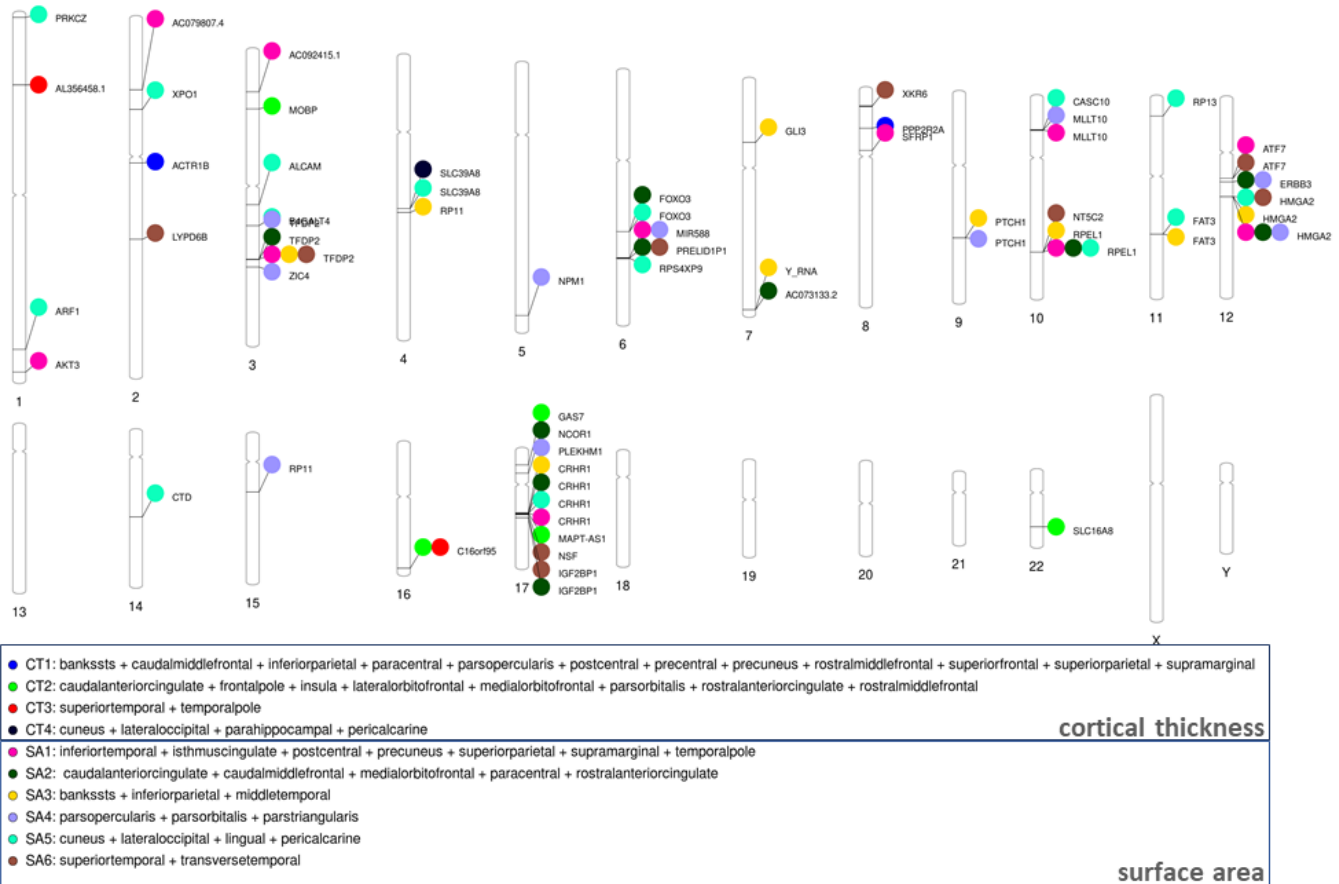

Figure 6

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

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SUPPLEMENTARYTABLES8822Final.docx