Multivariate patterns between brain network properties, polygenic scores, phenotypes, and environment in preadolescents

Jungwoo Seo[1], Eunji Lee[2], Bo-gyeom Kim[2], Gakyung Kim[1], Yoonjung Yoonie Joo[3], Jiook Cha[1, 2, 4, 5]

1. Department of Brain and Cognitive Sciences, College of Natural Sciences, Seoul National University, Seoul, South Korea
2. Department of Psychology, College of Social Sciences, Seoul National University, Seoul, South Korea
3. Institute of Data Science, Korea University, Seoul, South Korea
4. Institute of Psychological Science, Seoul National University, Seoul, South Korea
5. Graduate School of Artificial Intelligence, Seoul National University, Seoul, South Korea

Correspondence to:

Jiook Cha, PhD

Gwanak-ro 1, Building 16, Suite M512, Gwanakgu, Seoul, 08826, South Korea, connectome@snu.ac.kr

Abstract

The brain network is an infrastructure for cognitive and behavioral processes. Genetic and environmental factors influence the development of the brain network. However, little is known about how specific genetic traits and children's brain network properties are related. Furthermore, insight into the holistic relationship of brain network properties with genes, environment, and phenotypic outcomes in children is still limited. To fill these knowledge gaps, we investigated the multivariate associations between the brain network properties and three domains using a large youth sample (the ABCD study, N=9,393, 9-10 years old): (i) genetic predisposition of various traits, (ii) phenotypic outcomes, and (iii) environmental factors. We constructed structural brain networks using probabilistic tractography and estimated nodal and global network measures such as degree and network efficiency. We then conducted sparse canonical correlation analysis with brain network measures and polygenic scores of 30 complex traits (e.g., IQ), phenotypic traits (e.g., cognitive ability), and environmental variables. We found multivariate associations of brain network properties with (i) genetic risk for psychiatric disorders, (ii) genetic influence on cognitive ability, and (iii) the phenotype of cognitive ability-psychopathology in preadolescents. Our subsequent mediation analysis using the latent variables from the canonical correlation analysis showed that the influence of genetic factors for cognitive ability on the cognitive outcomes was partially mediated by the brain network properties. Taken together, this study shows the key role of the development of the brain structural network in children in cognitive development with its tight, likely causal, relationship with genetic factors. These findings may shed light on future

studies of the longitudinal deviations of those gene-environment-brain network relationships in normal and disease conditions.

Introduction

Childhood and adolescence are critical periods for brain development (Bethlehem et al., 2022). Proper brain development during this time is vital for cognitive and behavioral maturation (Bunge & Wright, 2007; Luna et al., 2010) and mental health (Fornito et al., 2015; Paus et al., 2008). Such development is influenced by genetic and environmental factors. Therefore, understanding the connections between the brain, cognitive-behavioral traits in children, and the impact of genetics and the environment on brain development is crucial in developmental and clinical neuroscience. The Adolescent Brain and Cognitive Development (ABCD) study (Jernigan et al., 2018) provides a rich dataset, encompassing genetic, neuroimaging, environmental, and phenotypic data for over 10,000 children aged 9-10 years old at baseline. This dataset opens an unprecedented opportunity to explore the connections between genes, environment, brain, and phenotypic outcomes more robustly with less concern about sampling bias (Marek et al., 2022) and age-confounding effects.

The brain is a complex network of tissues that communicate through white matter bundles. Diffusion MRI and tractography enable the reconstruction of the structural brain network (Jeurissen et al., 2019; Sotiropoulos & Zalesky, 2019). Graph theory (Rubinov & Sporns, 2010) allows us to estimate the brain network properties embedded in the whole brain's complex connectivity. The heritability of the brain network properties in youth ranges from 25% to 70% (Koenis et al., 2015; van den Heuvel, van Soelen, et al., 2013), indicating a strong genetic influence on shaping brain networks. However, little is known about which specific genetic traits and children's brain network properties are related. The recent development of the polygenic score approach (Torkamani et al., 2018) has paved the way to quantify the

genetic propensity of specific traits, such as bipolar disorder, and explore which genetic predispositions are related to the brain network properties.

Brain network properties are associated with cognitive abilities (Bathelt et al., 2018; Kim et al., 2016; Koenis et al., 2015; Ma et al., 2017; Suprano et al., 2020), psychiatric disorders (Alexander-Bloch et al., 2010; Collin et al., 2017; Rudie et al., 2012), and environmental factors such as socioeconomic status (D. J. Kim et al., 2019; Tooley et al., 2020) throughout the developmental period. Despite valuable insights from previous studies, a comprehensive understanding of the holistic relationship of brain network properties with genes, environment, and phenotypic outcomes in children is still limited due to the predominant use of univariate approaches. No single measurement is enough to characterize complex genes, brain networks, environment, and phenotypes information. Rather, combining a range of variables can capture complex traits more appropriately. Multivariate analysis, such as canonical correlation analysis (CCA), is useful for investigating holistic relationships underlying a set of variables simultaneously. CCA reveals multivariate associations linking sets of variables from two domains by maximizing the canonical correlation between them (Hotelling, 1936; Wang et al., 2020). For this reason, CCA has been extensively employed in studying links between the brain, cognition, genes, and environment (Alnaes et al., 2020; Fernandez-Cabello et al., 2022; Modabbernia et al., 2021; Smith et al., 2015; Wang et al., 2020).

This study aims to investigate the multivariate associations between structural brain network properties and three different domains in preadolescents: (i) genetic predisposition of various traits, (ii) phenotypic outcomes, and (iii) environmental factors. To achieve this, we leveraged the largest available childhood dataset (i.e., ABCD study)

and analysis techniques such as the polygenic score approach and sparse canonical correlation analysis (Witten et al., 2009).

Materials and Methods

ABCD participants

We used genetic, neuroimaging, environmental, and phenotypic data from the Adolescent Brain Cognitive Development (ABCD) study release 2.0 and 3.0 (http://abcdstudy.org). The ABCD study, which is the largest longitudinal investigation of brain development and child health in the United States, recruited multiethnic children (N=11,875) aged 9-10 years from 21 research sites with self-reported ethnicities that comprised of 52.3% Caucasians, 20.3% Mexican Americans, 14.7% African Americans, and 12.5% Asian-Americans and others. All participants and their parents or legal guardians provided informed consent and assent forms before participating in the study.

Genotype Data

The saliva DNA samples of study participants were collected, and 733,293 single nucleotide polymorphisms (SNPs) were genotyped at Rutgers University Cell and DNA Repository (RUCDR) with Affymetrix NIDA Smoke Screen Array. We excluded SNPs with genotype call rate <95%, sample call rate <95%, and minor allele frequency (MAF) <1%. The genotypes were imputed using the Michigan Imputation Server (Das et al., 2016) using the 1000 Genome phase3 version5 panel (Genomes Project et al., 2015) with Eagle v2.4 phasing (Loh et al., 2016). Then, the imputed variants with INFO score > .3 that did not meet our quality control criteria (i.e., call rate <95%, MAF <1%, and Hardy–Weinberg equilibrium p-value <1e-6)

were additionally filtered out. To address potential bias derived from genetically diverse and related family members in the ABCD study, we employed PC-Air (Conomos et al., 2015) and PC-Relate (Conomos et al., 2016) to obtain genetically unrelated individuals beyond 4th-degree relatives (i.e., kinship coefficient >0.022) and to remove outliers beyond 6 SD limits from the center of ancestrally informative principal component (PC) space. After quality control procedures, we included a total of 11,301,999 variants in 10,199 unrelated multiethnic participants, among whom 7,893 participants were of European ancestry.

Polygenic Scores (PGSs)

We used publicly available European-based GWAS summary statistics to calculate PGSs for 30 distinct traits: Attention-deficit/hyperactivity disorder (ADHD) (Demontis et al., 2019), cognitive performance (CP) (Lee et al., 2018), educational attainment (EA) (Lee et al., 2018), major depressive disorder (MDD) (Wray et al., 2018), insomnia (Jansen et al., 2019), snoring (Jansen et al., 2019), intelligence quotient (IQ) (Savage et al., 2018), post-traumatic stress disorder (PTSD) (Nievergelt et al., 2019), depression (DEP) (Howard et al., 2019; Shen et al., 2020), body mass index (BMI) (Akiyama et al., 2017; Locke et al., 2015), alcohol dependence (ALCDEP) (Walters et al., 2018), autism spectrum disorder (ASD) (Grove et al., 2019), automobile speeding propensity (ASP) (Akiyama et al., 2017), bipolar disorder (BIP) (Stahl et al., 2019), cannabis during lifetimes (Cannabis) (Pasman et al., 2019), ever smoker (Karlsson Linner et al., 2019), shared effects on five major psychiatric disorder (CROSS) (Cross-Disorder Group of the Psychiatric Genomics,

2013), alcoholic drinks consumption per week (Drinking) (Karlsson Linner et al., 2019), eating disorder (ED) (Watson et al., 2019), neuroticism (Nagel et al., 2018), obsessive-compulsive disorder (OCD) (International Obsessive Compulsive Disorder Foundation Genetics & Studies, 2018), first principal components of four risky behaviors (Risky Behav) (Karlsson Linner et al., 2019), general risk tolerance (RiskTol) (Karlsson Linner et al., 2019), schizophrenia (SCZ) (Bipolar et al., 2018; Lam et al., 2019), worrying (Nagel et al., 2018), anxiety (Otowa et al., 2016), subjective well-being (SWB) (Okbay et al., 2016), general happiness (UK Biobank GWAS. Neale Lab. http://www.nealelab.is/ukbiobank/), and general happiness for health (happiness-health) (UK Biobank GWAS. Neale Lab. http://www.nealelab.is/ukbiobank/) and meaningful life (happiness-meaning) (UK Biobank GWAS. Neale Lab. http://www.nealelab.is/ukbiobank/).

The GWAS summary statistics were used as input for PRS-CS (Ge et al., 2019), a Bayesian regression method, to estimate the posterior effect sizes of SNPs. The final scores were calculated using PLINK v1.9. To optimize the scores, we followed the suggestion of the original PRS-CS paper and chose the optimal global shrinkage hyperparameter (phi, φ) from among four possible values: 1, 1e-2, 1e-4, and 1e-6. The validation procedure was carried out within 14 PGSs (i.e., DEP, MDD, ADHD, general happiness, happiness-health, happiness-meaning, SWB, insomnia, snoring, BMI, PTSD, CP, EA, IQ) that had related measures in the ABCD study. For each PGS, we performed linear regression of the phenotype variable with each of the four scores and covariates (sex, age, and the first ten genetic PCs), and then, based on R^2 and beta coefficient of PGS, selected one of the four PGSs. The remaining 16 PGSs was automatically validated by PRS-CS-auto (Ge et al., 2019),

which select the optimal value of global shrinkage parameter employing a Bayesian approach. Finally, to minimize the bias from population stratification, we residualized the final PGSs with the first ten genetic PCs.

Environmental Factors

To investigate the relationship between children's brain network properties and their environment, we examined 121 environmental and culture-related variables for analysis. The variables consist of 80 variables representing the family history of various problems (e.g., alcohol problem, drug use problem, depression, suicide-related problem, etc.), 19 variables related to the residential history derived area deprivation index (ADI), 11 parent characteristics and family culture related variables, and others (e.g., perinatal condition, school environment, neighborhood safety, etc.).

To retain the sample as much as possible, we imputed missing values. For categorical variables, we replaced the missing values with the most frequent value (mode imputation); for continuous variables, with medians. Next, variables with near zero variance were eliminated. We finally used 75 environmental variables for the statistical analysis.

Phenotype Data

To investigate the relationship between children's brain network properties and their mental health, physical health, and neurocognitive capacity, we examined 174 phenotypic variables. These included 89 mental health and abnormal behaviors

related variables (e.g., KSAD diagnosis, Child Behavior Checklist), 74 physical health-related variables (e.g., physical disorders, medical history), and 10 NIH Toolbox cognitive assessment variables and a delay discounting related variable (i.e., cash choice task score). We also preprocessed the missing values with mode imputation for categorical variables and median imputation for continuous variables. Among phenotype data, variables having near zero variance were removed. We finally used 117 phenotype variables for the statistical analysis after variable selection based on the variance.

Structural Brain Network Construction

Detailed procedures for acquiring and preprocessing MRI data are described in (Kim et al., 2022). To estimate brain structural networks from neuroimaging, individual connectome data was generated. This was achieved by applying MRtrix3 (Tournier et al., 2019) to the preprocessed dMRI data to estimate whole-brain white matter tracts and generate individualized connectomes. Probabilistic tractography was performed using constrained-spherical deconvolution (CSD) (Calamante et al., 2010; Tournier et al., 2007) with random seeding across the brain and target streamline counts of 20 million. Initial tractograms were filtered using spherical-deconvolution informed filtering (2:1 ratio)(Smith et al., 2013), resulting in a final streamline count of 10 million. An 84x84 whole-brain connectome matrix was generated for each participant using the T1-based parcellation and segmentation from FreeSurfer with Desikan-Killiany atlas (Desikan et al., 2006)(68 nodes for the cortical region and 16 nodes for the subcortical region). This approach ensured that

individual participants' connectomes were based on their neuroanatomy. The computation was conducted on supercomputers at Argonne Leadership Computing Facility Theta and Texas Advanced Computing Center Stampede2.

Brain Network Measures (BNMs)

We used the connectome matrix to construct an undirect weighted graph representing the structural brain network. Nodes and edges in the graph represent parcellated gray matter regions and connections between them, respectively. Connection strength was quantified by the streamline count. To account for the potential false positive connections generated by probabilistic tractography and their impact on network topology, we eliminated extremely weak connections (streamline counts less than 3). After thresholding, we excluded individuals with at least one isolated node, assuming all brain regions are communicable via at least one path.

We calculated 18 different types of brain network measurements (BNMs) representing different aspects of brain network's property (Rubinov & Sporns, 2010; van den Heuvel & Sporns, 2011). We calculated eight global graph metrics (including network density, modularity, normalized modularity, normalized average clustering coefficient, normalized characteristic path length, global efficiency, normalized global efficiency, small worldness) and five nodal graph metrics (including degree, strength, clustering coefficient, betweenness centrality, nodal efficiency) to represent brain network's global and regional properties. In addition, to examine rich club organization, we also calculated raw and normalized rich club coefficient (van den Heuvel & Sporns, 2011) of the network and the strength of rich club connection,

feeder connection, and local connection. Brain regions with top 12% degree were defined as rich club nodes following previous work (van den Heuvel, Sporns, et al., 2013). All graph measures were calculated using the package Brain Connectivity Toolbox (https://sites.google.com/site/bctnet/).

Sparse Canonical Correlation Analysis

To examine a latent mode of covariation between structural brain network properties and various polygenic scores, environmental factors, and phenotypic outcomes, we used sparse canonical correlation analysis (Witten et al., 2009) between brain network measures and three types of non-imaging data (i.e., PGSs, environmental variables, phenotype variables) separately. Canonical correlation analysis (CCA) is a multivariate procedure that maximizes the correlation between the linear combinations of two sets of variables. Sparse canonical correlation analysis is one of the popular variants of CCA with L1 regularization for sparse solutions. We used sparse canonical correlation analysis to avoid over-fitting and improve interpretability from sparse solutions.

The most popular algorithm for sparse canonical correlation analysis is penalized matrix decomposition (PMD)(Witten et al., 2009), which solves optimization problem of below equation for given two sets of data matrix $X_{n\times p}$, $Y_{n\times q}$. (n: sample size; p, q: the number of variables of domain X and Y respectively; u, v: canonical weights of domain X and Y respectively; c1, c2: regularization parameter)

$$\max cov(Xu,\ Yv)$$

257 $$s.t.\ \ \| u \|_2 = \| v \|_2 = 1,\ \ \| u \|_1 \leq c_1,\ \ \| v \|_1 \leq c_2$$

258 To interpret Witten's sparse canonical correlation analysis as correlation
259 maximization, we need to assume covariance matrices $X^T X$, $Y^T Y$ are identity
260 matrices (Witten et al., 2009). But in our study with the high dimensional brain
261 datasets, the assumption is hardly satisfied (Fig. S1). For this reason, we interpreted
262 Witten's sparse canonical correlation analysis as a maximizing covariance algorithm
263 between two sets of variables rather than maximizing correlation.

264 To test generalizability of the sparse canonical correlation analysis results,
265 we split the dataset into a training and test set. For sparse canonical correlation
266 analysis with PGS and brain network measures, we split train set (n=5,411) and test
267 set (n=1,145) based on genetic relatedness to avoid including biological family in the
268 same dataset. In addition, we only used participants classified as genetically
269 European ancestry to control genetic confounding effects for sparse canonical
270 correlation analysis with PGSs. For sparse canonical correlation analysis with
271 environmental factors and phenotype data, we used stratified train (80%) - test
272 (20%) split by family history of various problems and KSAD diagnosis respectively.
273 **Table 2** summarizes the demographic information of the samples included in this
274 study.

275 We attempted to control the potential confounding effects by regressing out
276 the variance explained by age, sex, parental education, household income, marital
277 status, race ethnicity, BMI, and ABCD-site out of brain network measures,
278 environmental factors, and phenotype data. For binary variables such as family
279 history and KSADS-COMP, we regressed out with logistic regression. Residualized
280 data was used for the input of sparse canonical correlation analysis. On the other

hand, for polygenic scores, we did not regress out from PGSs because listed covariates may not contaminate genetic information.

We selected optimal L1 regularization parameters from 5-fold cross validation searching from 0.1 to 1 with a step size of 0.05 for both X and Y variables respectively. The optimal L1 parameter combination was selected to maximize the covariance of validation set between canonical variates of the first component (Fig. S2).

For each sparse canonical correlation analysis, we extracted five modes of covariance. To examine the statistical significance of each mode, we used a permutation test. By randomly shuffling the rows of one dataset and remaining the other, we generated 5,000 permutation sets. The p-value of each component was calculated based on the number of permutation sets having greater covariance than that obtained from the original dataset, and FDR-correction was done within each CCA.

$$p_{uncorrected} = \frac{N_{null\ cov > cov}}{N_{null}}$$

Selected variables and their loading depend on the input sample. To find variables reliably related to each mode, we used bootstrap resampling. We randomly resampled 5,000 times with replacement and assessed the 95% confidence interval of each variable's loading and how consistently it was selected. We interpreted the significant modes based on loading patterns of variables whose 95% confidence interval of loading does not cross zero (Xia et al., 2018) and selected more

frequently than expected by chance (i.e., more frequently selected than expected by binomial distribution). Because sparse canonical correlation analysis with bootstrap sample may change the order of components (axis rotation) and signs (reflection) (Misic et al., 2016; Xia et al., 2018), the re-alignment procedure is needed to estimate confidence interval of loading properly. We matched the components and signs based on cosine similarity of weight vectors obtained from original dataset and bootstrap sample. To assess the reproducibility of the findings, we applied the model to the held-out test set and estimated significance of each mode through the permutation test.

Mediation Analysis

After sparse canonical correlation analysis, we raised a hypothesis that the genetic factors of cognitive ability may influence the phenotype of cognitive ability-psychopathology through brain network properties (Results – Mediation Analysis). To test the hypothesis, we examined whether brain network property scores related to the genetic predisposition of cognitive ability (i.e., BNM score of PGS-BNM mode 2) mediated the relationship between polygenic scores for cognitive ability (i.e., PGS score of PGS-BNM mode 2) and the phenotype of cognitive ability-psychopathology (i.e., phenotype score of Phenotype-BNM mode 2). Age, sex, race ethnicity, parental education, household income, marital status, ABCD site, and BMI were used as covariates in the mediation analysis. The two-sided p-values for each path were estimated from 500 bootstrap samples.

Results

Modes of covariation between polygenic scores and structural brain network properties

Using sparse canonical correlation analysis, we investigated the relationship between 30 genome-wide polygenic scores (PGSs) and measures of brain network properties (BNMs). Our analysis revealed a significant mode (mode 1) and a mode that showed marginal significance (mode 2) (mode 1: $p = 0.006$, $cov = 0.846$, $r = 0.107$; mode 2: $p = 0.0545$ $(p_{unc} = 0.021)$, $cov = 0.585$, $r = 0.118$, all p-values were FDR corrected). Detailed results of the permutation test and the loading patterns for all five modes can be found in the supplementary material (Fig. S3, S4).

The first mode of PGSs showed significant positive loadings for PGSs related to psychiatric disorders, such as bipolar disorder and cross disorder **(Fig. 1a)**. On the brain side, normalized clustering and rich club coefficients showed positive loadings, while network density, degree, and raw rich club coefficients, and the connection strengths (among the bilateral rostral anterior cingulate, rostral middle frontal, paracentral, lingual regions, and cerebellums) showed negative loadings **(Fig. 1b, S4)**.

In the second mode of the PGSs, we observed strong positive loadings for PGSs related to cognitive abilities, such as cognitive performance (CP) and IQ. Meanwhile, there were weak negative loadings for material use-related PGSs **(Fig. 1c)**. In terms of corresponding brain network properties, positive loadings were found for connection strength, betweenness centrality, normalized nodal efficiency of middle temporal gyrus, and betweenness centrality of the left inferior parietal area.

On the other hand, the strength of posterior cingulate cortex, thalamus, and caudate, as well as feeder and rich club connection strength had negative loadings **(Fig. 1d, S4)**. To sum up, the polygenic scores for higher cognitive ability were positively linked to network integrity of the middle temporal gyrus and negatively associated with connection strength of the posterior cingulate cortex, thalamus, and caudate.

Modes of covariation between phenotypes and structural brain network properties

Among the five modes of covariation, only the second and the third mode were significant and generalized to the hold-out test sets (mode 2: $p < 0.001$, cov = 1.069, $r = 0.119$; mode 3: $p = 0.016$, cov = 0.640, $r = 0.089$, all p-values were FDR corrected). We thus interpreted the second and third modes only. Detailed results of the permutation test and the loading patterns for all five modes can be found in the supplementary material (Fig. S5, S6).

The second mode of phenotypes depicted the covariation between brain network properties and the shared characteristics of cognitive ability and psychopathology (cognitive ability-psychopathology) **(Fig. 1e, 1f, S6)**. Cognitive test scores (i.e., NIH Toolbox scores) showed positive loadings, whereas psychopathologic traits, such as CBCL and PGBI mania scores, showed negative loadings. Several brain network properties showed loading patterns similar to that of polygenic scores for cognitive ability (i.e., PGS mode 2). Consistent with the loadings observed in PGS mode 2, the connection strength and nodal efficiency of the middle temporal gyrus and betweenness centrality of the left inferior parietal region showed

positive loadings, while the strength of the posterior cingulate cortex, thalamus, and caudate, along with the feeder and rich club connections, showed negative loadings. Conversely, certain network properties were found to be statistically significant only in relation to phenotypes. The nodal efficiency of the superior temporal region and the betweenness centrality of the postcentral and supramarginal regions showed positive loadings. Additionally, the strength of the anterior cingulate, paracentral, and overall subcortical regions, as well as the nodal efficiency of the insula and hippocampus, showed negative loadings. To summarize, phenotypes of cognitive ability-psychopathology were positively associated with the network integrity of the temporal and inferior parietal cortex, while being negatively associated with the connection strength of the cingulate and subcortical regions.

Despite strong negative loadings in the NIH Toolbox scores, the third mode did not show any significant features (Fig. S6).

Modes of covariation between environmental factors and structural brain network properties

Among the five modes analyzed, the second mode showed a marginal level of statistical significance, while the third mode showed a significant relationship (mode 2: $p = 0.092\ (p_{unc} = 0.0402),\ cov = 0.896,\ r = 0.074$; mode 3: $p = 0.016,\ cov = 0.802,\ r = 0.124$, all p-values were FDR corrected). Further information regarding the results of the permutation test and the loading patterns for all five modes can be found in the supplementary material (Fig. S7, S8).

The second mode showed the relationship between environmental factors related to socioeconomic status and brain network properties. Variables representing socioeconomic advantages showed positive loadings, whereas variables representing socioeconomic disadvantages showed negative loadings **(Fig. 2)**. Regarding its relationship with brain network properties, only the nodal efficiency of the right hippocampus showed a significant association with the second mode **(Table 4)**.

The third mode depicted the relationship between perinatal conditions, such as caesarian section and age of parents, and brain network properties **(Fig. 2)**. Perinatal conditions showed positive associations with certain brain network properties **(Table. 4)**. Specifically, the connection strength, betweenness centrality, and clustering coefficient of the entorhinal cortex showed positive loadings. In contrast, the strength of the precentral gyrus, cerebellum, and thalamus, as well as the nodal efficiency of the cerebellum and the degree of the caudate, showed negative loadings.

Mediation Analysis

We discovered that both polygenic scores and phenotypes related to cognitive ability shared covarying brain network properties. Specifically, we observed shared brain network properties, including the connection strength and nodal efficiency of the middle temporal gyrus, betweenness centrality of inferior parietal regions, and the connection strength of the posterior cingulate cortex and thalamus **(Fig. 1d, 1f)**. These findings led us to the hypothesis that the genetic factors

influencing cognitive ability may exert their influence through brain network properties.

To test this hypothesis, we examined whether the BNM score of PGS-BNM mode 2 serves as a mediator between the PGS score of PGS-BNM mode 2 and the phenotype score of phenotype-BNM mode 2 **(Fig. 3)**. The result demonstrated that brain network properties significantly mediated the relationship between polygenic scores for cognitive ability and the phenotype of cognitive ability-psychopathology (indirect effect = 0.013, $p < 0.001$).

Discussion

This study investigated the multivariate relationship of brain network properties with genes, environment, and phenotypic outcomes in children. We discovered multiple modes of covariation among the structural brain network properties and the following factors: (i) the genetic propensity for psychiatric disorders and cognitive capacity, (ii) the phenotypic outcomes related to the cognitive ability and psychopathology, and (iii) the environmental factors, such as socioeconomic status and perinatal conditions. Additionally, our study showed that the brain network mediates the effects of the polygenic scores for cognitive capacity on phenotypic outcomes of cognitive ability and psychopathology. This study offers insight into which genetic factors are related to the brain network properties and the multivariate relationships among variables of brain network properties, genes, environmental factors, and phenotypic outcomes in children. It gives rise to discussions about the underlying biology of brain network and cognitive development and further hypotheses in this field.

In the sparse canonical correlation analysis with phenotypes, we observed that the cognitive intelligence (NIH Toolbox) and psychopathology phenotypes (CBCL) exhibited a single mode of covariation rather than distinct modes. This single mode suggests that children's cognitive ability and abnormal behavior share underlying brain network properties but with opposite signs of correlations. This finding resembles previous findings, for instance, a single mode of positive-negative behavioral and genetic traits covarying with brain connectivity (Smith et al., 2015; Taquet et al., 2021).

Our mediation analysis showed that the brain network mediated genetic predisposition and phenotypic outcomes of cognition and psychopathology during childhood. This mediation may be facilitated by specific brain network properties associated with both the polygenic scores for cognitive abilities and phenotypes related to cognitive ability-psychopathology. Further investigation is necessary to determine the extent of the involvement of these common network properties in the gene-brain network-phenotypic outcome pathway. Mediation through the brain network properties accounts for approximately 5% of the gene-phenotype pathway of intelligence, implying other pathways, such as through cortical surface area or thickness (Lett et al., 2020).

Children with higher polygenic scores for cognitive abilities or phenotype scores of cognitive ability-psychopathology showed greater nodal efficiency in the middle temporal gyrus and inferior parietal region. These regions are constituents of the system proposed by the parieto-frontal integration theory (P-FIT) of intelligence (Jung & Haier, 2007) and play a crucial role in semantic, language, and number processing (Binder et al., 2009; Seghier, 2013; Visser et al., 2012). In contrast, children with lower scores on cognitive ability-psychopathology exhibited greater nodal efficiency in the hippocampus and insula. These regions have been implicated in various cognitive functions, including learning, memory, emotional regulation, and interoception (Jarrard, 1993; Namkung et al., 2017; Phelps, 2004; Rubin et al., 2014; Sweatt, 2004). While previous studies have reported a positive correlation between global efficiency and children's intelligence (Bathelt et al., 2018; Kim et al., 2016; Ma et al., 2017), our results did not. Our study suggests that the relationship between intelligence and network efficiency is more closely linked to the network efficiency of

some brain regions, such as the temporal and parietal regions, rather than the overall global efficiency.

Our results indicate a negative association between the connection strength of the posterior cingulate cortex and subcortical regions and polygenic scores and phenotypes related to cognitive ability. During childhood and adolescence, FA of white matter is known to increase (Tamnes et al., 2018). Unlike studies based on FA, research has reported decreased streamline counts and density during this period, particularly in connections with higher streamline counts and those linked to subcortical regions (Baker et al., 2015; Lim et al., 2015). Considering these documented changes in white matter connectivity, the observed negative association may indicate a progression in white matter development among children with higher cognitive abilities.

Our study suggests that children with higher polygenic scores for psychiatric disorders have lower network density and degree across various brain regions. Individuals with psychiatric disorders are known to have less white matter connection (Chen et al., 2021; Perry et al., 2019; van den Heuvel & Fornito, 2014). These findings suggest that the tendency of less white matter connections in individuals with a higher genetic risk for psychiatric disorders is already observable in 9-10 years old. It is aligned with the findings that youth with higher genetic risks for psychiatric disorders have smaller white matter volume and average connection strength (Fernandez-Cabello et al., 2022; Taquet et al., 2021).

The wiring and maintenance of white matter tracts require significant metabolic support. For this reason, the metabolic budget of the brain is a critical

494 factor in shaping the brain network (Bullmore & Sporns, 2012). Taking into account
495 network density is often used as a proxy for the wiring cost (Achard & Bullmore,
496 2007; Bassett et al., 2009), lower density observed in children with higher genetic
497 risks for psychiatric disorders may imply a link between genetic susceptibility to
498 psychiatric disorders and abnormal brain metabolism, subsequently affecting the
499 development of structural brain networks. Indeed, mitochondrial dysfunction has
500 consistently been proposed as a pathological foundation of psychiatric disorders
501 (Clay et al., 2011; Y. Kim et al., 2019; Rezin et al., 2009; Shao et al., 2008; Zuccoli et
502 al., 2017), and recent studies identified genetic overlaps between psychiatric
503 disorders and metabolic syndromes (Amare et al., 2017; Rodevand et al., 2021).

504 The brain network favors short-range connections over long-distance ones to
505 minimize wiring costs, promoting local segregation within the network (Bullmore &
506 Sporns, 2012). We observed a higher clustering coefficient in children with a higher
507 genetic risk for psychiatric disorders, which may be related to lower brain metabolic
508 support. Our results showed that genetic risks for psychiatric disorders were
509 negatively associated with raw rich club coefficients at k=20-50, but positively
510 associated with normalized rich club coefficients. Network density may partially
511 influence these correlations (van Wijk et al., 2010).

512 Our study indicates that environmental factors, such as socioeconomic status
513 and perinatal conditions, can affect brain network properties in children beyond mere
514 brain morphology (Alnaes et al., 2020). Specifically, we observed a positive
515 correlation between socioeconomic advantage and the nodal efficiency of the right
516 hippocampus. This finding aligns with previous studies showing the impact of
517 socioeconomic disadvantage on various aspects of the hippocampus, including gray

matter volume (Hanson et al., 2011; Jednorog et al., 2012), connectivity (Barch et al., 2016), and network efficiency (D. J. Kim et al., 2019). However, the significant associations between network properties and environmental factors were fewer in number compared to genetic factors. This suggests that the influence of environmental factors on brain network properties may be either less pronounced or more heterogeneous than genetic factors.

This study has several limitations. Firstly, we used cross-sectional data, which limits our ability to investigate the relationships between brain network properties, genes, phenotypes, and the environment over time. Although our findings provide insights into normative brain network development, a longitudinal study design is necessary to fully understand developmental trajectories. Second, polygenic scores only consider common genetic variants that reach a certain level of statistical significance in a genome-wide association study (GWAS). Thus, polygenic scores in this study do not capture the impact of rare variants or complex interactions like gene-gene and gene-environment interactions. Lastly, our statistical models only capture homogeneous linear relationships, which may not be adequate for comprehending the complex interactions (Greene et al., 2022) that occur among genetic profiles, environmental factors, phenotypic traits, and data-collection artifacts within a large population study. The small effect sizes observed in the relationship between brain network properties and genetic, phenotypic, and environmental factors ($r = 0.074$-0.124) may indicate heterogeneous associations depending on covariate profiles. To gain more informative insights into the relationships among genes, environment, brain, and phenotypes, a more sophisticated model that can capture heterogeneous associations based on covariates may be necessary. For

instance, employing conditional (or local) average treatment effect analysis with causal machine learning (Athey et al., 2019) would be beneficial, as it would aid in personal, clinical, and political decision-making.

Data Availability

All original data are publicly available from the NDA (https://nda.nih.gov/abcd/). Mook data, which corresponds to the processed data used in this study, was generated from conditional GAN for tabular data (Xu et al., 2019) and are available from this site (https://github.com/Transconnectome/ABCD-brain-network-SCCA).

Code Availability

Code is available from here: https://github.com/Transconnectome/ABCD-brain-network-SCCA.

Funding

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021R1C1C1006503, 2021K2A9A1A01102014, 2021K1A3A1A21037512, 2021M3E5D2A01022515), by Seoul National University Research Grant in 2021 (No. 200-20210083), by Creative-Pioneering Researchers Program through Seoul National University (No. 200-20220046), by Semi-Supervised Learning Research Grant by SAMSUNG (No.A0426-20220118), by Institute of Information & communications Technology

Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) [NO.2021-0-01343, Artificial Intelligence Graduate School Program (Seoul National University)], and by Identify the network of brain preparation steps for concentration Research Grant by Looxid Labs (No.339-20230001).

Author contributions

- Jungwoo Seo, Conceptualization, Data curation, Formal analysis, Validation, Investigation, Visualization, Writing – original draft, Writing – review and editing
- Eunji Lee, Data curation, Writing – original draft, Writing – review and editing
- Bogyeom Kim, Data curation, Writing – review and editing
- Gakyung Kim, Data curation, Writing – original draft, Writing – review and editing
- Yoonjung Yoonie Joo, Data curation, Writing – review and editing
- Jiook Cha, Conceptualization, Supervision, Project administration, Funding acquisition, Resources, Data curation, Writing – original draft, Writing – review and editing.

Ethics

Competing Interest Disclosures

None of the authors have significant competing financial, professional, or personal interests that might have influenced the performance or presentation of the work described in the manuscript.

583

584

Tables

Table 1. Abbreviation of PGSs and their meaning.

Abbreviation	Meaning
CROSS	Cross disorder between ASD, ADHD, bipolar disorder, MDD, and schizophrenia
Neuroticism	Neuroticism
Risky Behav	The first principal component of the four risky behaviors
RiskTol	General risk tolerance
Worry	Worrying subtype
ASP	Automobile speeding propensity: the tendency to drive faster than the speed limit
ALCDEP	Alcohol dependence
ED	Eating Disorder (Anorexia nervosa)
Drinking	Drinks per week – the average number of alcoholic drinks consumed per week
ASD	Autism Spectrum Disorder
BIP	Bipolar disorder
Cannabis	Cannabis during their lifetime (self-reported)
Smoke	Ever smoker – whether one has ever been a smoker
SCZ	Schizophrenia
OCD	Obsessive-Compulsive Disorder

Anxiety	Anxiety
DEP	Depression
MDD	Major Depression Disorder
GHappi	General Happiness
GHappiHealth	General Happiness - health
GHapiMean	General Happiness - meaningful life
SWB	Subjective Well Being
INSOMNIA	Insomnia
SNORING	Snoring
BMI	Body Mass Index
PTSD	Post-Traumatic Stress Disorder
CP	Cognitive Performance
EA	Education Attainment
IQ	Intelligence Quotient
ADHD	Attention-Deficit/Hyperactivity Disorder

587

588

589 **Table 2. Demographic information of the study participants.**

		PGS – BNM (n=6,549)		ENV – BNM (n=9,393)		Pheno – BNM (n=8,895)	
		train	test	train	test	train	test
N	total	5,406	1,143	7,514	1,879	7,116	1,779
Sex	Male	2,869	600	3,931	965	3,741	905
	Female	2,537	543	3,583	914	3,375	874
Race	White	3,831	865	4,167	1,041	4,086	1015
	Black	34	7	995	234	866	221
	Hispanic	993	154	1,406	401	1,269	323
	Asian	6	1	147	32	143	29
	Other	542	116	799	171	752	191

590

Table 3. p-value, correlation coefficient, covariance of sparse canonical correlation analysis results. p-value of each mode is estimated by permutation test. * p_FDR < 0.05, ** p_FDR < 0.01, *** p_FDR < 0.001, † p_uncorrected < 0.05.

Sparse Canonical Correlation Analysis with Brain Network Metrics		**Mode1**	**Mode2**	**Mode3**	**Mode4**	**Mode5**
Polygenic Scores	p_training	**0.0060****	0.0545 (0.0218)†	0.6633	0.9640	0.3033
	p_test	**0.0060****	**0.0210***	0.4582	0.1587	0.4582
	corr	**0.1071**	**0.1176**	0.1231	0.0831	0.1404
	cov	**0.8458**	**0.5845**	0.4425	0.3517	0.3830
Phenotypes	p_training	**0.0020***	**<0.001*****	**0.0160***	0.2128	0.0678
	p_test	0.1020	**0.0060****	**0.0110***	0.1990	0.5814
	corr	0.1124	**0.1194**	**0.0885**	0.0924	0.1233
	cov	1.2243	**1.0692**	**0.6401**	0.4819	0.4593
Environmental Factors	p_training	0.0920	0.0920 (0.0402)†	**0.0160***	0.5582	0.0920
	p_test	0.1227	**0.0145***	**0.0030****	0.4682	0.3663
	corr	0.0745	**0.0741**	**0.1239**	0.0843	0.1060
	cov	1.1196	**0.8956**	**0.8023**	0.5263	0.5226

Table 4. Sparse canonical correlation analysis results of environmental factors-brain network measures. Only significant brain network measures are presented. The loadings of corresponding environmental factors are visualized in Figure 2.

	Region	Network measure	Loading	Loading 95% CI
Mode 2	Hippocampus	Nodal efficiency R-HI	0.2098	[0.0031, 0.4603]
Mode 3	Caudate	Deg L-CA	-0.3778	[-0.4016, -0.0038]
		Deg R-CA	-0.3753	[-0.3930, -0.0156]
	Entorhinal cortex	Stren L-EC	0.3417	[0.0336, 0.4331]
		Stren R-EC	0.3483	[0.0378, 0.4348]
		BC_L-EC	0.1091	[0.0174, 0.1560]
		Norm_clust_coef L-EC	0.3691	[0.0208, 0.4188]
		Norm_clust_coef R-EC	0.3879	[0.0357, 0.4383]
	Precentral gyrus	Stren L-PrCG	-0.4642	[-0.4978, -0.0116]
		Stren R-PrCG	-0.4356	[-0.4866, -0.0100]
	Cerebellum	Stren L-CER	-0.2761	[-0.3509, -0.0222]
		Stren R-CER	-0.2768	[-0.3465, -0.0171]
		Nodal efficiency L-CER	-0.3695	[-0.4661, -0.0615]
		Nodal efficiency R-CER	-0.3699	[-0.4646, -0.0603]
	Thalamus	Stren L-TH	-0.4175	[-0.5760, -0.0779]
		Stren R-TH	-0.4141	[-0.5797, -0.0651]

	Temporal pole	Stren R-TP	0.1590	[0.0255, 0.2223]
		Norm_clust_coef L-TP	0.3942	[0.0216, 0.4324]

599

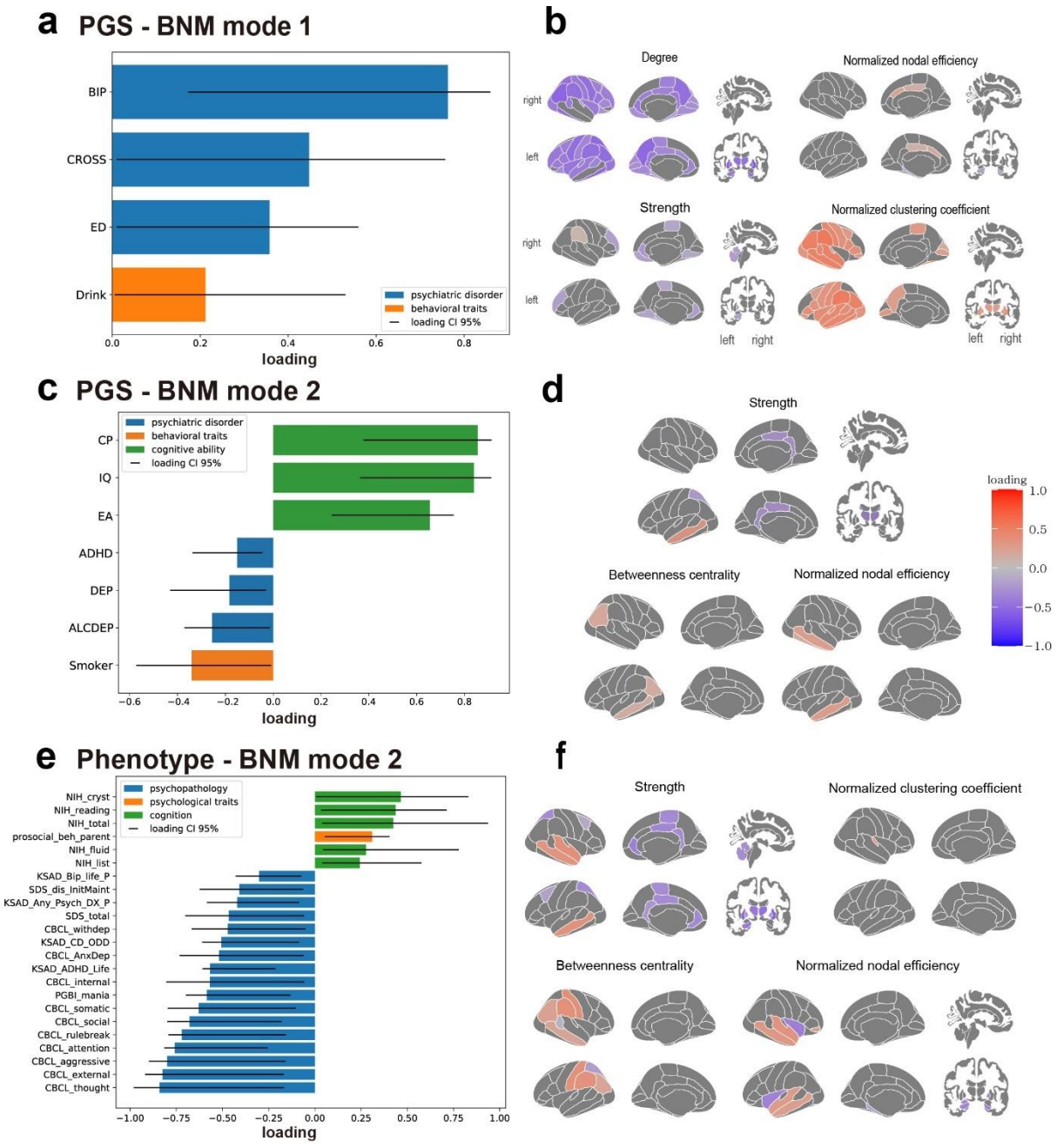
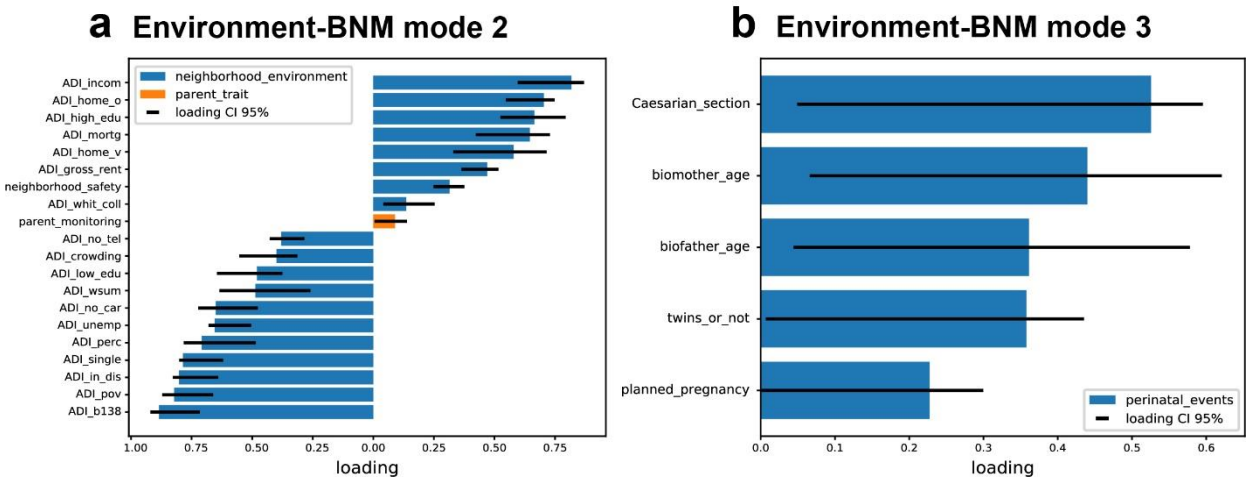


Figure 1 Sparse canonical correlation analysis results of polygenic scores-brain network measures and phenotypes-brain network measures. The results of the first and second modes of polygenic scores (PGSs) - brain network measures (BNMs) and the second mode of phenotypes - brain network measures. (a) The loadings of significant PGS variables in the PGSs-BNMs mode 1. The error bars represent the 95% confidence interval of the loading, estimated from the 5,000 bootstrap samples. Among the 30 PGSs, only significant PGS variables are presented, while the others can be found in supplementary figure 4. The color of each bar represents the category to which

the variable belongs. (Abbreviations - BIP: bipolar disorder; CROSS: cross-disorder; ED: eating disorder; Drink: drink per week) (b) The loadings of significant nodal brain network measures in the PGSs-BNMs mode 1. The loading patterns were visualized with R-package '*ggseg'* (Mowinckel & Vidal-Pineiro, 2020) (c) The loadings of significant PGS variables in the PGSs-BNMs mode 2. (Abbreviations - CP: cognitive performance; EA: educational attainment; DEP: depression; ALCDEP: alcohol dependence; Smoker: ever smoking) (d) The loadings of significant nodal brain network measures in the PGSs-BNMs mode 2. (e) The loadings of significant PGS variables in the phenotypes-BNMs mode 2. The loading patterns of other modes are shown in supplementary figure 6 (abbreviations- NIH: NIH Toolbox score; KSAD: K-SADS (Kiddie Schedule for Affective Disorders and Schizophrenia; SDS: sleep disturbance scale; CBCL: child behavior checklist; PGBI: parent version of general behavior inventory) (f) The loadings of significant nodal brain network measures in the phenotypes-BNMs mode 2.



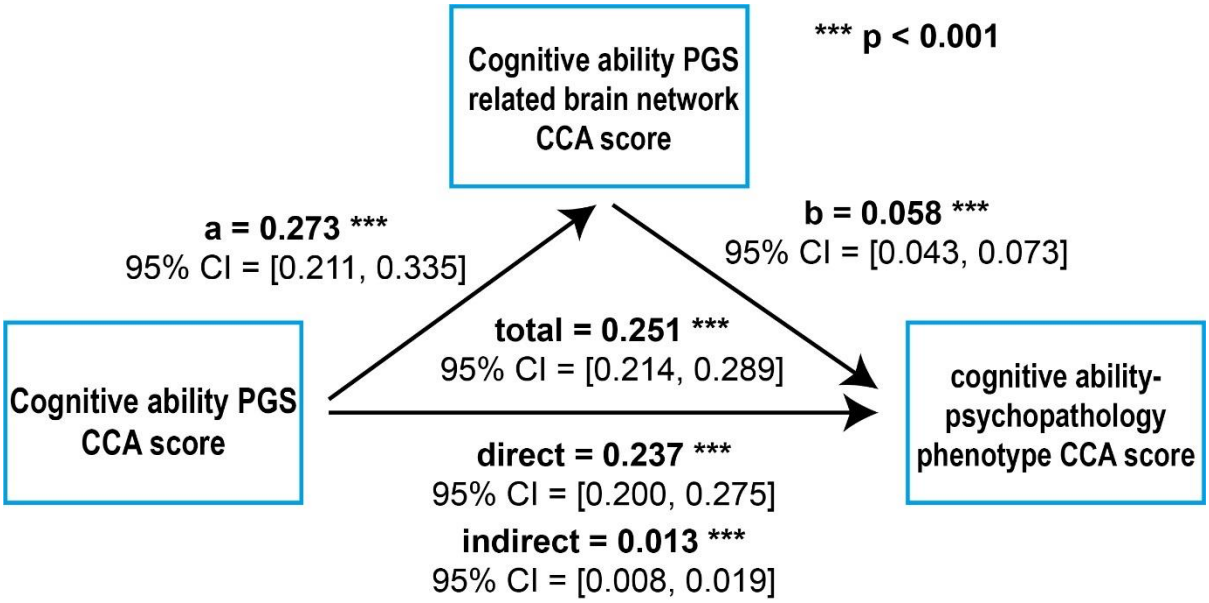
621

622 **Figure 2 Sparse canonical correlation analysis results of environmental factors-brain network**
623 **measures.** (a) The loadings of environmental variables in mode 2. (b) The loadings of environmental
624 variables in mode 3. Only significant variables are presented. The loadings of corresponding brain
625 network measures are presented in Table 4, and the loading patterns of other modes are shown in
626 supplementary figure 8 (abbreviation – ADI: residential history derived area deprivation index)

627

628

629



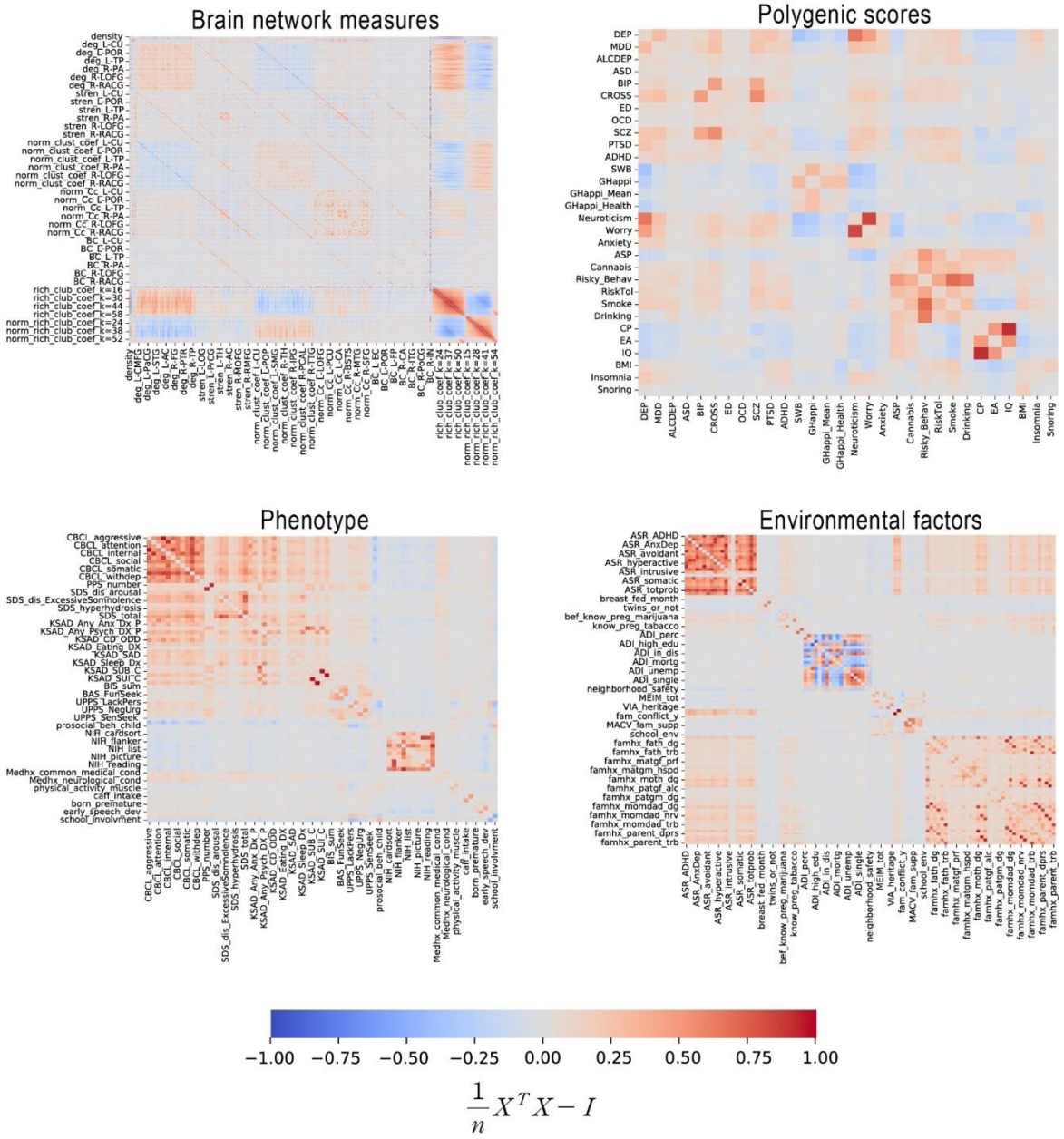
630

631 **Figure 3 Result of mediation analysis.** We used cognitive ability PGS CCA score (i.e., PGS CCA
632 score of PGS-BNM mode 2) as an independent variable, brain network property CCA score related to
633 the cognitive ability PGS (i.e., BNM CCA score of PGS-BNM mode 2) as a mediator, and phenotype
634 of intelligence-psychopathology CCA score (i.e., phenotype CCA score of Phenotype-BNM mode 2)
635 as a dependent variable. The p-value and effect size of all paths were presented.

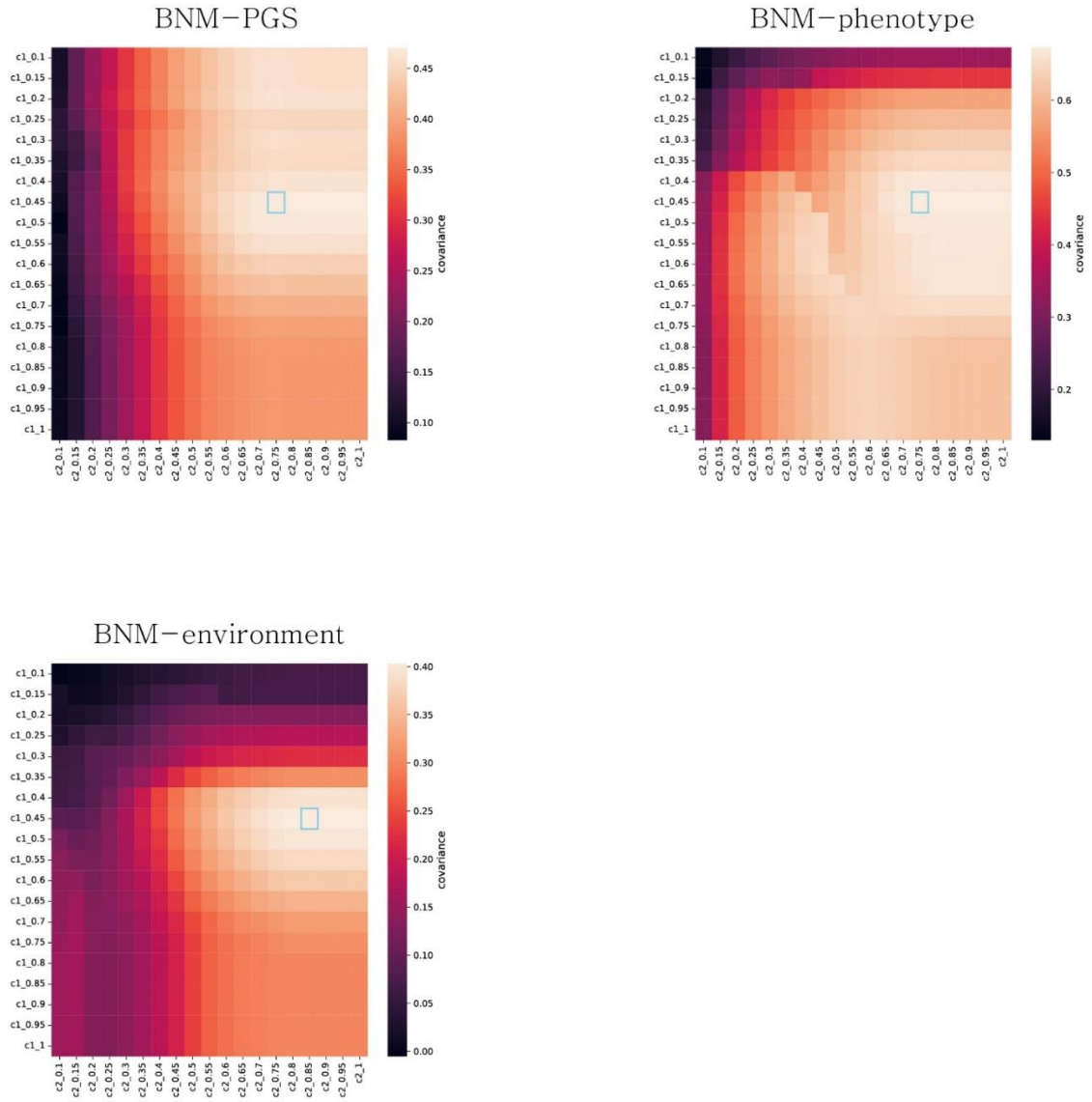
636

637

Supplementary Materials



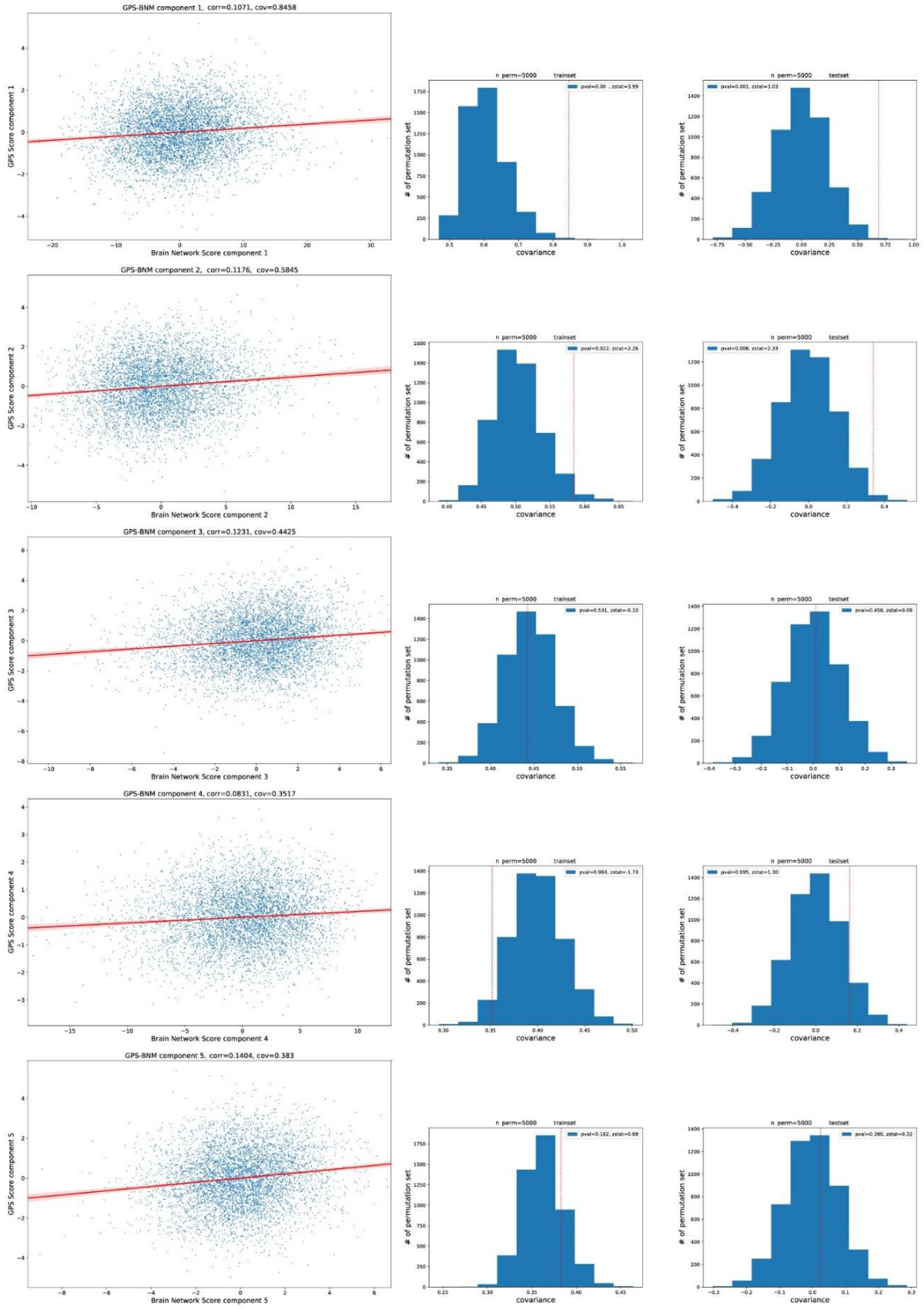
Supplementary Figure 1. Check Witten's assumption. $\frac{1}{\mathbf{n}}\mathbf{X}^{\boldsymbol{T}}\boldsymbol{X} - \boldsymbol{I}$ of the analyzed data is presented. Here, X represents the data from which the variance explained by covariates has been regressed out, as mentioned in the main text.



644

645 **Supplementary Figure 2. Optimal regularization parameters.** We conducted a grid search to
646 optimize the L1 regularization parameters for Sparse Canonical Correlation Analysis (SCCA). The grid
647 search ranged from 0.1 to 1 in sparsity parameters, with increments of 0.05. The objective was to
648 identify the sparsity parameter pair that yielded the highest mean covariance of the first component in
649 the validation set, using 5-fold cross-validation.

650

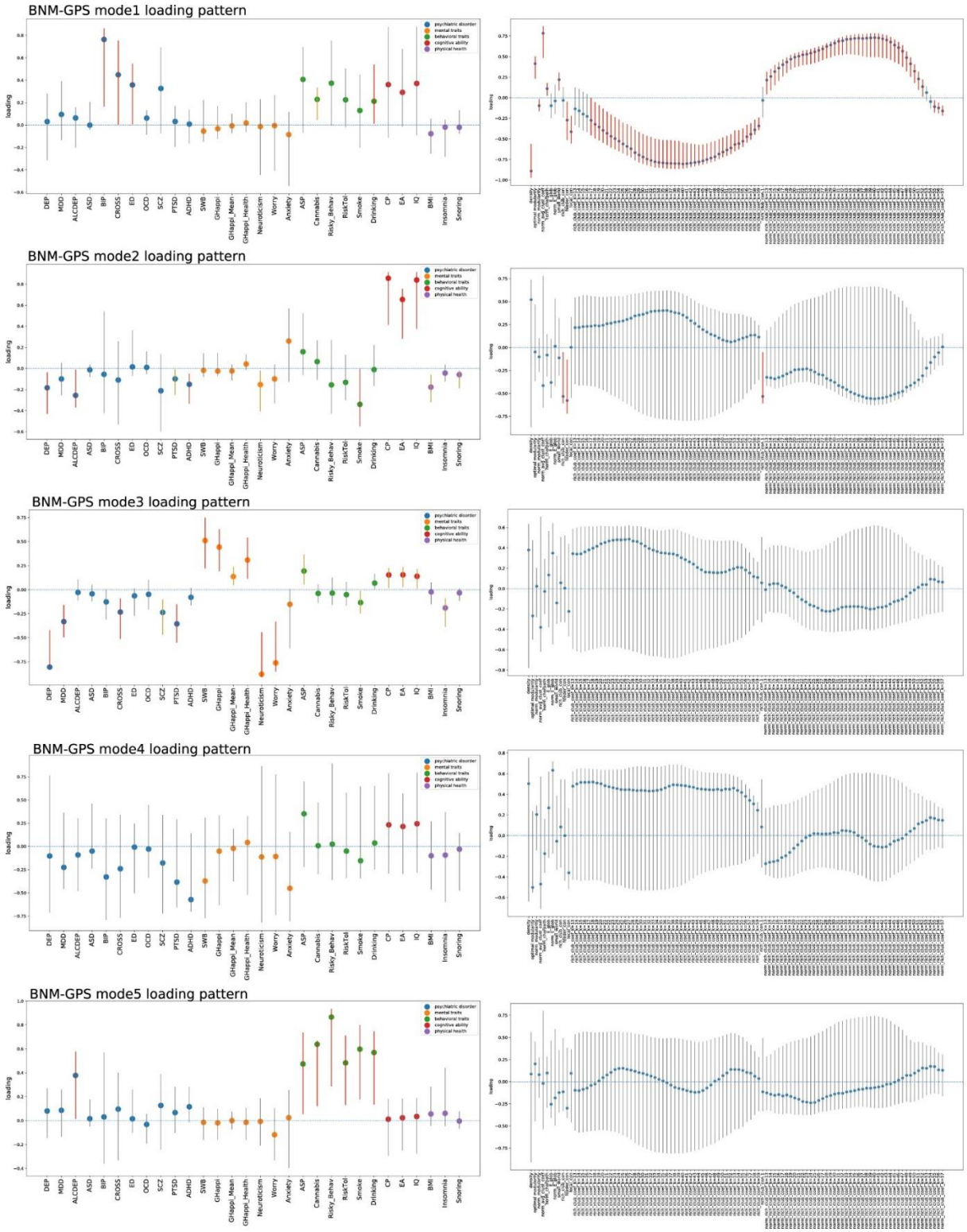


Supplementary Figure 3. Scatter plots of SCCA scores of brain network measures versus SCCA scores of polygenic scores and their permutation test results of training and test set.

654 Each permutation test result shows the null distribution of SCCA estimated by permutation test.

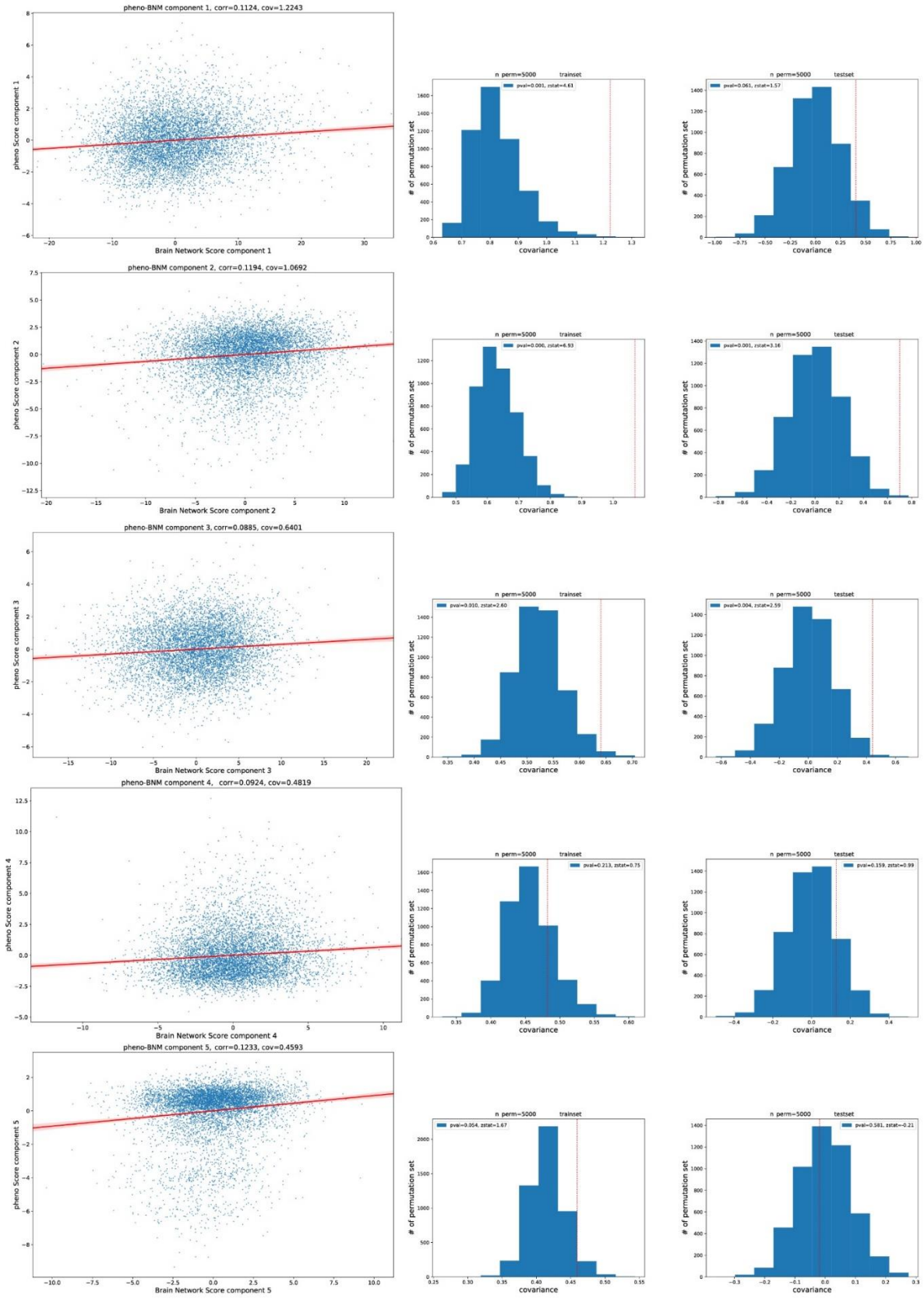
655 Dashed red line refers to the actual covariance of each mode.

656



Supplementary Figure 4. Sparse CCA loading patterns of brain network measures – polygenic scores. The left column presents the loading of each polygenic score variable, while the right column displays the loadings of global brain network measures and rich club coefficients. Each data point

represents the loading value of a specific variable and is color-coded based on its category. The error bars indicate the 95% confidence interval of the loading, which was estimated using bootstrapping. Gray-colored error bars indicate a 95% confidence interval that crosses 0. Yellow-colored error bars indicate a 95% confidence interval that does not cross 0 but corresponds to a variable that is not frequently selected by Sparse CCA. Red-colored error bars represent a 95% confidence interval that does not cross 0 and indicates a variable that is frequently selected. Only the variables represented by red error bars are presented in the main text.

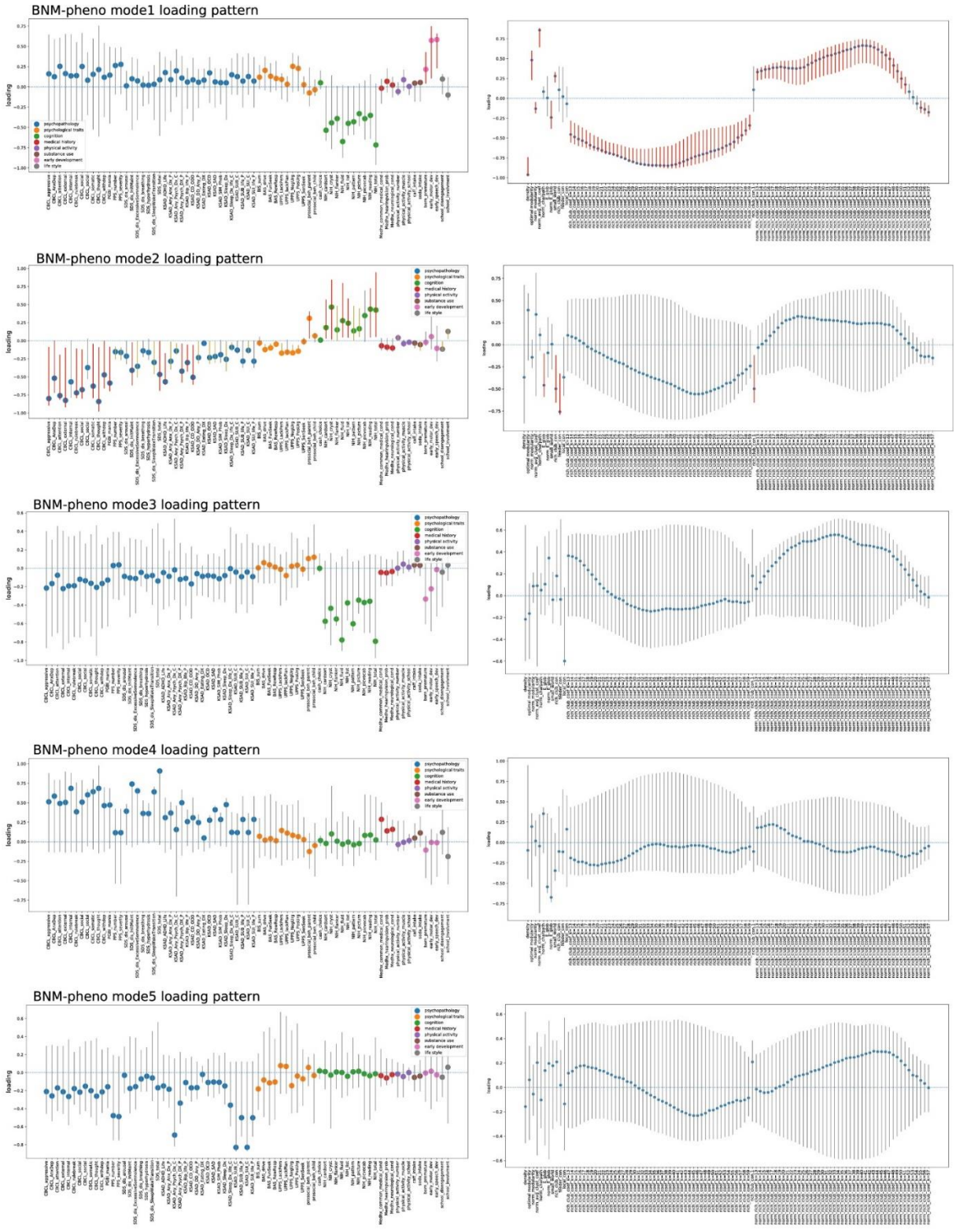


Supplementary Figure 5. Scatter plots of SCCA scores of brain network measures versus SCCA scores of phenotypic measures and their permutation test results of training and test

672 **set.** Each permutation test result shows the null distribution of SCCA estimated by permutation test.

673 Dashed red line refers to the actual covariance of each mode.

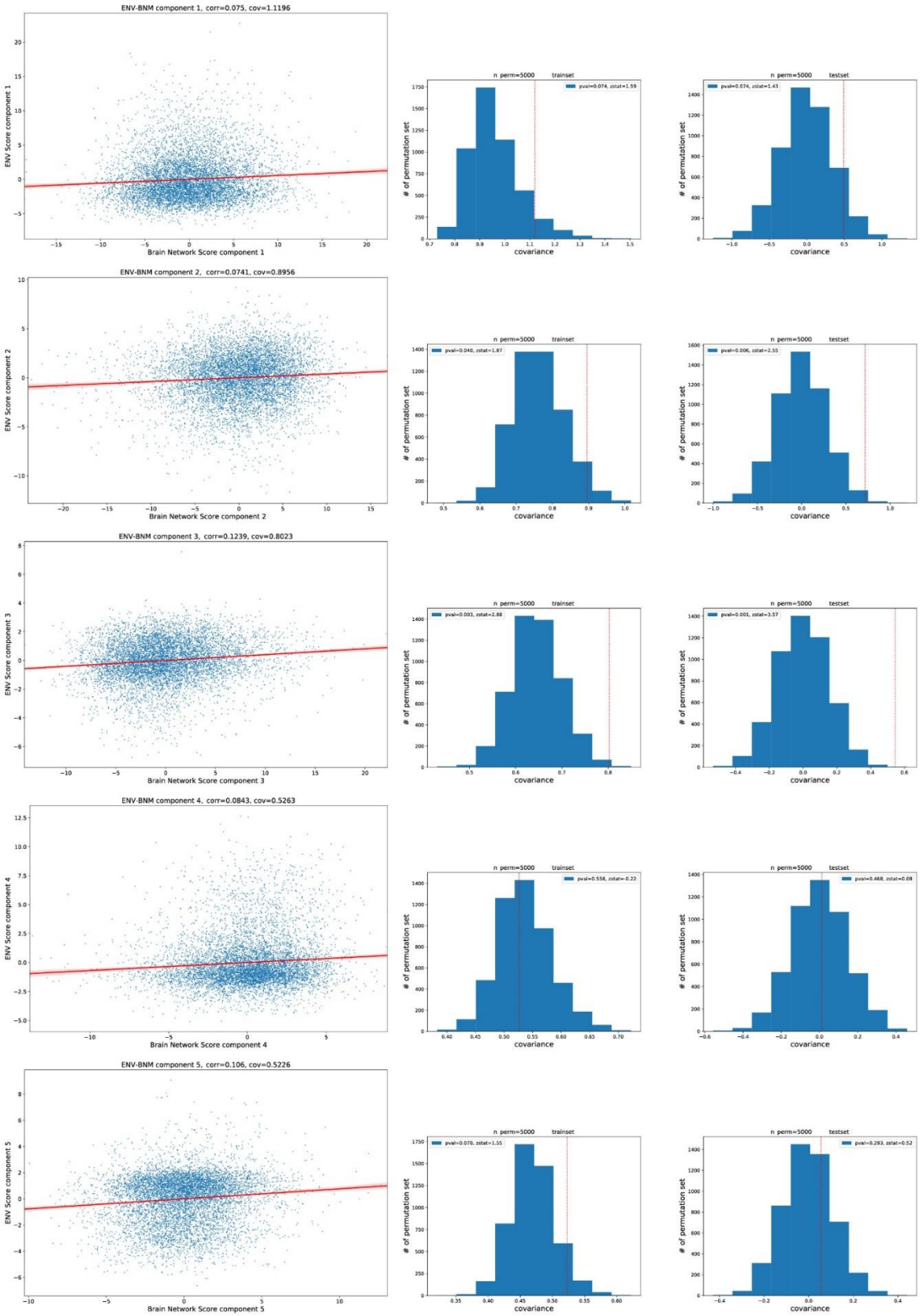
674



675

676 **Supplementary Figure 6. Sparse CCA loading patterns of brain network measures – phenotype**
677 **variables.** The left column presents the loading of each phenotype variable, while the right column
678 displays the loadings of global brain network measures and rich club coefficients. Each data point
679 represents the loading value of a specific variable and is color-coded based on its category. The error

bars indicate the 95% confidence interval of the loading, which was estimated using bootstrapping. Gray-colored error bars indicate a 95% confidence interval that crosses 0. Yellow-colored error bars indicate a 95% confidence interval that does not cross 0 but corresponds to a variable that is not frequently selected by Sparse CCA. Red-colored error bars represent a 95% confidence interval that does not cross 0 and indicates a variable that is frequently selected. Only the variables represented by red error bars are presented in the main text.

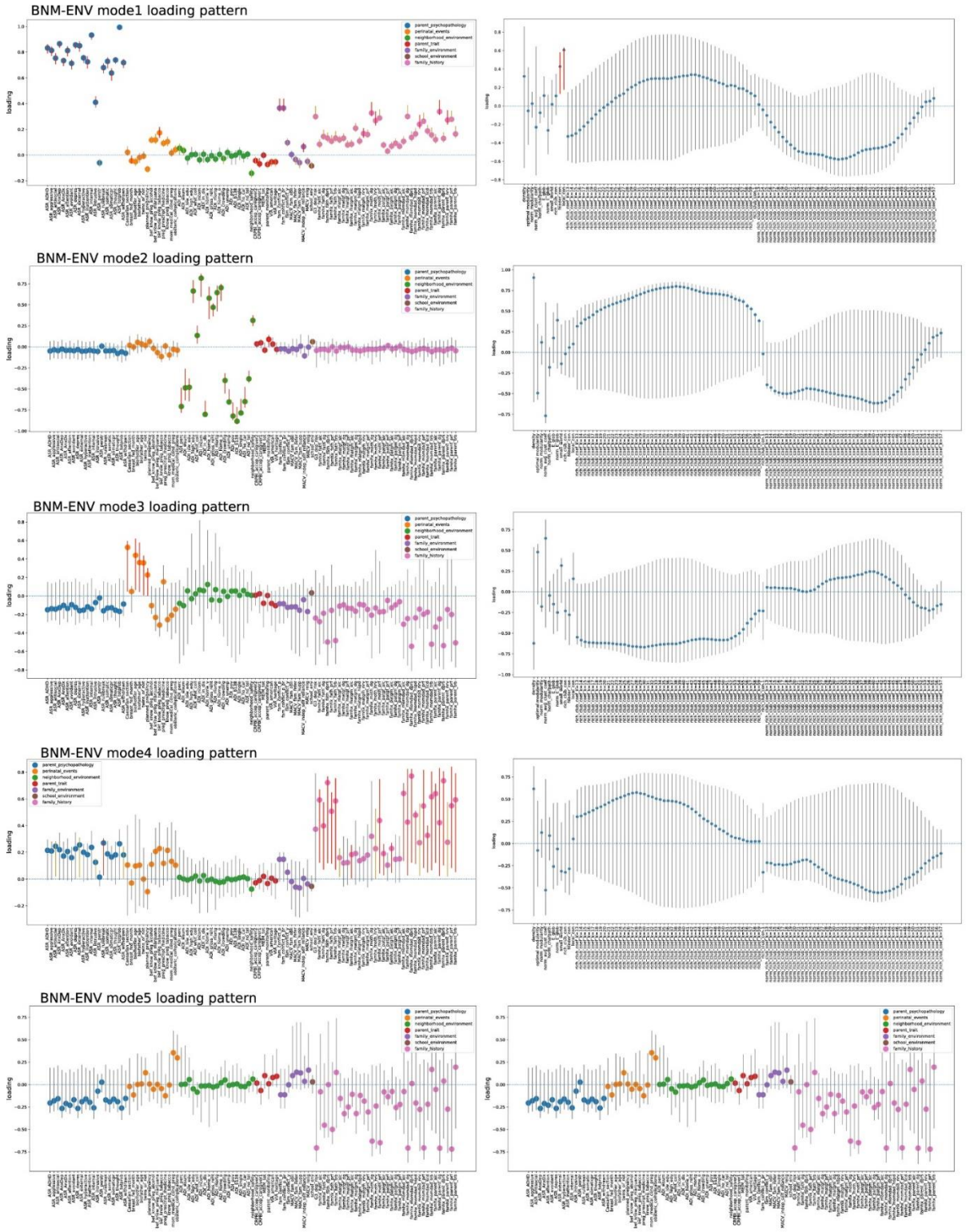


Supplementary Figure 7. Scatter plots of SCCA scores of brain network measures versus SCCA scores of environmental factors and their permutation test results of training and test

690 **set.** Each permutation test result shows the null distribution of SCCA estimated by permutation test.

691 Dashed red line refers to the actual covariance of each mode.

692



Supplementary Figure 8. Sparse CCA loading patterns of brain network measures – environmental factors. The left column presents the loading of each environmental variable, while the right column displays the loadings of global brain network measures and rich club coefficients. Each data point represents the loading value of a specific variable and is color-coded based on its

698 category. The error bars indicate the 95% confidence interval of the loading, which was estimated
699 using bootstrapping. Gray-colored error bars indicate a 95% confidence interval that crosses 0.
700 Yellow-colored error bars indicate a 95% confidence interval that does not cross 0 but corresponds to
701 a variable that is not frequently selected by Sparse CCA. Red-colored error bars represent a 95%
702 confidence interval that does not cross 0 and indicates a variable that is frequently selected. Only the
703 variables represented by red error bars are presented in the main text.

704

References

Achard, S., & Bullmore, E. (2007). Efficiency and cost of economical brain functional networks. *PLoS Comput Biol*, *3*(2), e17. https://doi.org/10.1371/journal.pcbi.0030017

Akiyama, M., Okada, Y., Kanai, M., Takahashi, A., Momozawa, Y., Ikeda, M., Iwata, N., Ikegawa, S., Hirata, M., Matsuda, K., Iwasaki, M., Yamaji, T., Sawada, N., Hachiya, T., Tanno, K., Shimizu, A., Hozawa, A., Minegishi, N., Tsugane, S., . . . Kamatani, Y. (2017). Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. *Nat Genet*, *49*(10), 1458-1467. https://doi.org/10.1038/ng.3951

Alexander-Bloch, A. F., Gogtay, N., Meunier, D., Birn, R., Clasen, L., Lalonde, F., Lenroot, R., Giedd, J., & Bullmore, E. T. (2010). Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. *Front Syst Neurosci*, *4*, 147. https://doi.org/10.3389/fnsys.2010.00147

Alnaes, D., Kaufmann, T., Marquand, A. F., Smith, S. M., & Westlye, L. T. (2020). Patterns of sociocognitive stratification and perinatal risk in the child brain. *Proc Natl Acad Sci U S A*, *117*(22), 12419-12427. https://doi.org/10.1073/pnas.2001517117

Amare, A. T., Schubert, K. O., Klingler-Hoffmann, M., Cohen-Woods, S., & Baune, B. T. (2017). The genetic overlap between mood disorders and cardiometabolic diseases: a systematic review of genome wide and candidate gene studies. *Transl Psychiatry*, *7*(1), e1007. https://doi.org/10.1038/tp.2016.261

Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. *The Annals of Statistics*, *47*(2), 1148-1178, 1131. https://doi.org/10.1214/18-AOS1709

Baker, S. T., Lubman, D. I., Yucel, M., Allen, N. B., Whittle, S., Fulcher, B. D., Zalesky, A., & Fornito, A. (2015). Developmental Changes in Brain Network Hub Connectivity in Late Adolescence. *J Neurosci*, *35*(24), 9078-9087. https://doi.org/10.1523/JNEUROSCI.5043-14.2015

Barch, D., Pagliaccio, D., Belden, A., Harms, M. P., Gaffrey, M., Sylvester, C. M., Tillman, R., & Luby, J. (2016). Effect of Hippocampal and Amygdala Connectivity on the Relationship Between Preschool Poverty and School-Age Depression. *Am J Psychiatry*, *173*(6), 625-634. https://doi.org/10.1176/appi.ajp.2015.15081014

Bassett, D. S., Bullmore, E. T., Meyer-Lindenberg, A., Apud, J. A., Weinberger, D. R., & Coppola, R. (2009). Cognitive fitness of cost-efficient brain functional networks. *Proc Natl Acad Sci U S A*, *106*(28), 11747-11752. https://doi.org/10.1073/pnas.0903641106

Bathelt, J., Gathercole, S. E., Butterfield, S., team, C., & Astle, D. E. (2018). Children's academic attainment is linked to the global organization of the white matter connectome. *Dev Sci*, *21*(5), e12662. https://doi.org/10.1111/desc.12662

Bethlehem, R. A. I., Seidlitz, J., White, S. R., Vogel, J. W., Anderson, K. M., Adamson, C., Adler, S., Alexopoulos, G. S., Anagnostou, E., Areces-Gonzalez, A., Astle, D. E., Auyeung, B., Ayub, M., Bae, J., Ball, G., Baron-Cohen, S., Beare, R., Bedford, S. A., Benegal, V., . . . Alexander-Bloch, A. F. (2022). Brain charts for the human lifespan. *Nature*, *604*(7906), 525-533. https://doi.org/10.1038/s41586-022-04554-y

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex*, *19*(12), 2767-2796. https://doi.org/10.1093/cercor/bhp055

Bipolar, D., Schizophrenia Working Group of the Psychiatric Genomics Consortium. Electronic address, d. r. v. e., Bipolar, D., & Schizophrenia Working Group of the

753 Psychiatric Genomics, C. (2018). Genomic Dissection of Bipolar Disorder and
754 Schizophrenia, Including 28 Subphenotypes. *Cell*, *173*(7), 1705-1715 e1716.
755 https://doi.org/10.1016/j.cell.2018.05.046
756 Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nat Rev*
757 *Neurosci*, *13*(5), 336-349. https://doi.org/10.1038/nrn3214
758 Bunge, S. A., & Wright, S. B. (2007). Neurodevelopmental changes in working memory and
759 cognitive control. *Curr Opin Neurobiol*, *17*(2), 243-250.
760 https://doi.org/10.1016/j.conb.2007.02.005
761 Calamante, F., Tournier, J. D., Jackson, G. D., & Connelly, A. (2010). Track-density imaging
762 (TDI): super-resolution white matter imaging using whole-brain track-density
763 mapping. *Neuroimage*, *53*(4), 1233-1243.
764 https://doi.org/10.1016/j.neuroimage.2010.07.024
765 Chen, T., Chen, Z., & Gong, Q. (2021). White Matter-Based Structural Brain Network of
766 Major Depression. *Adv Exp Med Biol*, *1305*, 35-55. https://doi.org/10.1007/978-981-
767 33-6044-0_3
768 Clay, H. B., Sillivan, S., & Konradi, C. (2011). Mitochondrial dysfunction and pathology in
769 bipolar disorder and schizophrenia. *Int J Dev Neurosci*, *29*(3), 311-324.
770 https://doi.org/10.1016/j.ijdevneu.2010.08.007
771 Collin, G., Scholtens, L. H., Kahn, R. S., Hillegers, M. H. J., & van den Heuvel, M. P. (2017).
772 Affected Anatomical Rich Club and Structural-Functional Coupling in Young Offspring
773 of Schizophrenia and Bipolar Disorder Patients. *Biol Psychiatry*, *82*(10), 746-755.
774 https://doi.org/10.1016/j.biopsych.2017.06.013
775 Conomos, M. P., Miller, M. B., & Thornton, T. A. (2015). Robust inference of population
776 structure for ancestry prediction and correction of stratification in the presence of
777 relatedness. *Genet Epidemiol*, *39*(4), 276-293. https://doi.org/10.1002/gepi.21896
778 Conomos, M. P., Reiner, A. P., Weir, B. S., & Thornton, T. A. (2016). Model-free Estimation of
779 Recent Genetic Relatedness. *Am J Hum Genet*, *98*(1), 127-148.
780 https://doi.org/10.1016/j.ajhg.2015.11.022
781 Cross-Disorder Group of the Psychiatric Genomics, C. (2013). Identification of risk loci with
782 shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet*,
783 *381*(9875), 1371-1379. https://doi.org/10.1016/S0140-6736(12)62129-1
784 Das, S., Forer, L., Schonherr, S., Sidore, C., Locke, A. E., Kwong, A., Vrieze, S. I., Chew, E.
785 Y., Levy, S., McGue, M., Schlessinger, D., Stambolian, D., Loh, P. R., Iacono, W. G.,
786 Swaroop, A., Scott, L. J., Cucca, F., Kronenberg, F., Boehnke, M., . . . Fuchsberger,
787 C. (2016). Next-generation genotype imputation service and methods. *Nat Genet*,
788 *48*(10), 1284-1287. https://doi.org/10.1038/ng.3656
789 Demontis, D., Walters, R. K., Martin, J., Mattheisen, M., Als, T. D., Agerbo, E., Baldursson,
790 G., Belliveau, R., Bybjerg-Grauholm, J., Baekvad-Hansen, M., Cerrato, F., Chambert,
791 K., Churchhouse, C., Dumont, A., Eriksson, N., Gandal, M., Goldstein, J. I., Grasby,
792 K. L., Grove, J., . . . Neale, B. M. (2019). Discovery of the first genome-wide
793 significant risk loci for attention deficit/hyperactivity disorder. *Nat Genet*, *51*(1), 63-75.
794 https://doi.org/10.1038/s41588-018-0269-7
795 Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner,
796 R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006).
797 An automated labeling system for subdividing the human cerebral cortex on MRI
798 scans into gyral based regions of interest. *Neuroimage*, *31*(3), 968-980.
799 https://doi.org/10.1016/j.neuroimage.2006.01.021
800 Fernandez-Cabello, S., Alnaes, D., van der Meer, D., Dahl, A., Holm, M., Kjelkenes, R.,
801 Maximov, II, Norbom, L. B., Pedersen, M. L., Voldsbekk, I., Andreassen, O. A., &
802 Westlye, L. T. (2022). Associations between brain imaging and polygenic scores of
803 mental health and educational attainment in children aged 9-11. *Neuroimage*, *263*,
804 119611. https://doi.org/10.1016/j.neuroimage.2022.119611
805 Fornito, A., Zalesky, A., & Breakspear, M. (2015). The connectomics of brain disorders. *Nat*
806 *Rev Neurosci*, *16*(3), 159-172. https://doi.org/10.1038/nrn3901

Ge, T., Chen, C. Y., Ni, Y., Feng, Y. A., & Smoller, J. W. (2019). Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun*, *10*(1), 1776. https://doi.org/10.1038/s41467-019-09718-5

Genomes Project, C., Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., Marchini, J. L., McCarthy, S., McVean, G. A., & Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68-74. https://doi.org/10.1038/nature15393

Greene, A. S., Shen, X., Noble, S., Horien, C., Hahn, C. A., Arora, J., Tokoglu, F., Spann, M. N., Carrion, C. I., Barron, D. S., Sanacora, G., Srihari, V. H., Woods, S. W., Scheinost, D., & Constable, R. T. (2022). Brain-phenotype models fail for individuals who defy sample stereotypes. *Nature*, *609*(7925), 109-118. https://doi.org/10.1038/s41586-022-05118-w

Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., Pallesen, J., Agerbo, E., Andreassen, O. A., Anney, R., Awashti, S., Belliveau, R., Bettella, F., Buxbaum, J. D., Bybjerg-Grauholm, J., Baekvad-Hansen, M., Cerrato, F., Chambert, K., Christensen, J. H., . . . Borglum, A. D. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet*, *51*(3), 431-444. https://doi.org/10.1038/s41588-019-0344-8

Hanson, J. L., Chandra, A., Wolfe, B. L., & Pollak, S. D. (2011). Association between income and the hippocampus. *PLoS One*, *6*(5), e18712. https://doi.org/10.1371/journal.pone.0018712

Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, *28*, 321-377. https://doi.org/10.1093/biomet/28.3-4.321

Howard, D. M., Adams, M. J., Clarke, T. K., Hafferty, J. D., Gibson, J., Shirali, M., Coleman, J. R. I., Hagenaars, S. P., Ward, J., Wigmore, E. M., Alloza, C., Shen, X., Barbu, M. C., Xu, E. Y., Whalley, H. C., Marioni, R. E., Porteous, D. J., Davies, G., Deary, I. J., . . . McIntosh, A. M. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat Neurosci*, *22*(3), 343-352. https://doi.org/10.1038/s41593-018-0326-7

International Obsessive Compulsive Disorder Foundation Genetics, C., & Studies, O. C. D. C. G. A. (2018). Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Mol Psychiatry*, *23*(5), 1181-1188. https://doi.org/10.1038/mp.2017.154

Jansen, P. R., Watanabe, K., Stringer, S., Skene, N., Bryois, J., Hammerschlag, A. R., de Leeuw, C. A., Benjamins, J. S., Munoz-Manchado, A. B., Nagel, M., Savage, J. E., Tiemeier, H., White, T., andMe Research, T., Tung, J. Y., Hinds, D. A., Vacic, V., Wang, X., Sullivan, P. F., . . . Posthuma, D. (2019). Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat Genet*, *51*(3), 394-403. https://doi.org/10.1038/s41588-018-0333-3

Jarrard, L. E. (1993). On the role of the hippocampus in learning and memory in the rat. *Behav Neural Biol*, *60*(1), 9-26. https://doi.org/10.1016/0163-1047(93)90664-4

Jednorog, K., Altarelli, I., Monzalvo, K., Fluss, J., Dubois, J., Billard, C., Dehaene-Lambertz, G., & Ramus, F. (2012). The influence of socioeconomic status on children's brain structure. *PLoS One*, *7*(8), e42486. https://doi.org/10.1371/journal.pone.0042486

Jernigan, T. L., Brown, S. A., & Dowling, G. J. (2018). The Adolescent Brain Cognitive Development Study. *J Res Adolesc*, *28*(1), 154-156. https://doi.org/10.1111/jora.12374

Jeurissen, B., Descoteaux, M., Mori, S., & Leemans, A. (2019). Diffusion MRI fiber tractography of the brain. *NMR Biomed*, *32*(4), e3785. https://doi.org/10.1002/nbm.3785

Jung, R. E., & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence: converging neuroimaging evidence. *Behav Brain Sci*, *30*(2), 135-154; discussion 154-187. https://doi.org/10.1017/S0140525X07001185

Karlsson Linner, R., Biroli, P., Kong, E., Meddens, S. F. W., Wedow, R., Fontana, M. A.,

Lebreton, M., Tino, S. P., Abdellaoui, A., Hammerschlag, A. R., Nivard, M. G., Okbay, A., Rietveld, C. A., Timshel, P. N., Trzaskowski, M., Vlaming, R., Zund, C. L., Bao, Y., Buzdugan, L., . . . Beauchamp, J. P. (2019). 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*, *51*(2), 245-257. https://doi.org/10.1038/s41588-018-0309-3

Kim, D. J., Davis, E. P., Sandman, C. A., Glynn, L., Sporns, O., O'Donnell, B. F., & Hetrick, W. P. (2019). Childhood poverty and the organization of structural brain connectome. *Neuroimage*, *184*, 409-416. https://doi.org/10.1016/j.neuroimage.2018.09.041

Kim, D. J., Davis, E. P., Sandman, C. A., Sporns, O., O'Donnell, B. F., Buss, C., & Hetrick, W. P. (2016). Children's intellectual ability is associated with structural network integrity. *Neuroimage*, *124*(Pt A), 550-556. https://doi.org/10.1016/j.neuroimage.2015.09.012

Kim, K., Joo, Y. Y., Ahn, G., Wang, H. H., Moon, S. Y., Kim, H., Ahn, W. Y., & Cha, J. (2022). The sexual brain, genes, and cognition: A machine-predicted brain sex score explains individual differences in cognitive intelligence and genetic influence in young children. *Hum Brain Mapp*, *43*(12), 3857-3872. https://doi.org/10.1002/hbm.25888

Kim, Y., Vadodaria, K. C., Lenkei, Z., Kato, T., Gage, F. H., Marchetto, M. C., & Santos, R. (2019). Mitochondria, Metabolism, and Redox Mechanisms in Psychiatric Disorders. *Antioxid Redox Signal*, *31*(4), 275-317. https://doi.org/10.1089/ars.2018.7606

Koenis, M. M., Brouwer, R. M., van den Heuvel, M. P., Mandl, R. C., van Soelen, I. L., Kahn, R. S., Boomsma, D. I., & Hulshoff Pol, H. E. (2015). Development of the brain's structural network efficiency in early adolescence: A longitudinal DTI twin study. *Hum Brain Mapp*, *36*(12), 4938-4953. https://doi.org/10.1002/hbm.22988

Lam, M., Chen, C. Y., Li, Z., Martin, A. R., Bryois, J., Ma, X., Gaspar, H., Ikeda, M., Benyamin, B., Brown, B. C., Liu, R., Zhou, W., Guan, L., Kamatani, Y., Kim, S. W., Kubo, M., Kusumawardhani, A., Liu, C. M., Ma, H., . . . Huang, H. (2019). Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nat Genet*, *51*(12), 1670-1678. https://doi.org/10.1038/s41588-019-0512-x

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T. A., Bowers, P., Sidorenko, J., Karlsson Linner, R., Fontana, M. A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P. N., Walters, R. K., Willoughby, E. A., . . . Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet*, *50*(8), 1112-1121. https://doi.org/10.1038/s41588-018-0147-3

Lett, T., Vogel, B. O., Ripke, S., Wackerhagen, C., Erk, S., Awasthi, S., Trubetskoy, V., Brandl, E. J., Mohnke, S., Veer, I. M., Nothen, M. M., Rietschel, M., Degenhardt, F., Romanczuk-Seiferth, N., Witt, S. H., Banaschewski, T., Bokde, A., Buchel, C., Quinlan, E. B., . . . Consortium, I. (2020). Cortical Surfaces Mediate the Relationship Between Polygenic Scores for Intelligence and General Intelligence. *Cerebral Cortex*, *30*(4), 2708-2719. https://doi.org/10.1093/cercor/bhz270

Lim, S., Han, C. E., Uhlhaas, P. J., & Kaiser, M. (2015). Preferential detachment during human brain development: age- and sex-specific structural connectivity in diffusion tensor imaging (DTI) data. *Cereb Cortex*, *25*(6), 1477-1489. https://doi.org/10.1093/cercor/bht333

Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., Powell, C., Vedantam, S., Buchkovich, M. L., Yang, J., Croteau-Chonka, D. C., Esko, T., Fall, T., Ferreira, T., Gustafsson, S., Kutalik, Z., Luan, J., Magi, R., Randall, J. C., . . . Speliotes, E. K. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nature*, *518*(7538), 197-206. https://doi.org/10.1038/nature14177

Loh, P. R., Danecek, P., Palamara, P. F., Fuchsberger, C., Y, A. R., H, K. F., Schoenherr, S., Forer, L., McCarthy, S., Abecasis, G. R., Durbin, R., & A, L. P. (2016). Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet*, *48*(11), 1443-1448. https://doi.org/10.1038/ng.3679

Luna, B., Padmanabhan, A., & O'Hearn, K. (2010). What has fMRI told us about the development of cognitive control through adolescence? *Brain Cogn*, *72*(1), 101-113. https://doi.org/10.1016/j.bandc.2009.08.005

Ma, J., Kang, H. J., Kim, J. Y., Jeong, H. S., Im, J. J., Namgung, E., Kim, M. J., Lee, S., Kim, T. D., Oh, J. K., Chung, Y. A., Lyoo, I. K., Lim, S. M., & Yoon, S. (2017). Network attributes underlying intellectual giftedness in the developing brain. *Sci Rep*, *7*(1), 11321. https://doi.org/10.1038/s41598-017-11593-3

Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., Donohue, M. R., Foran, W., Miller, R. L., Hendrickson, T. J., Malone, S. M., Kandala, S., Feczko, E., Miranda-Dominguez, O., Graham, A. M., Earl, E. A., Perrone, A. J., Cordova, M., Doyle, O., . . . Dosenbach, N. U. F. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nature*, *603*(7902), 654-660. https://doi.org/10.1038/s41586-022-04492-9

Misic, B., Betzel, R. F., de Reus, M. A., van den Heuvel, M. P., Berman, M. G., McIntosh, A. R., & Sporns, O. (2016). Network-Level Structure-Function Relationships in Human Neocortex. *Cereb Cortex*, *26*(7), 3285-3296. https://doi.org/10.1093/cercor/bhw089

Modabbernia, A., Janiri, D., Doucet, G. E., Reichenberg, A., & Frangou, S. (2021). Multivariate Patterns of Brain-Behavior-Environment Associations in the Adolescent Brain and Cognitive Development Study. *Biol Psychiatry*, *89*(5), 510-520. https://doi.org/10.1016/j.biopsych.2020.08.014

Mowinckel, A. M., & Vidal-Pineiro, D. (2020). Visualization of Brain Statistics With R Packages ggseg and ggseg3d. *Advances in Methods and Practices in Psychological Science*, *3*(4), 466-483. https://doi.org/10.1177/2515245920928009

Nagel, M., Jansen, P. R., Stringer, S., Watanabe, K., de Leeuw, C. A., Bryois, J., Savage, J. E., Hammerschlag, A. R., Skene, N. G., Munoz-Manchado, A. B., andMe Research, T., White, T., Tiemeier, H., Linnarsson, S., Hjerling-Leffler, J., Polderman, T. J. C., Sullivan, P. F., van der Sluis, S., & Posthuma, D. (2018). Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet*, *50*(7), 920-927. https://doi.org/10.1038/s41588-018-0151-7

Namkung, H., Kim, S. H., & Sawa, A. (2017). The Insula: An Underestimated Brain Area in Clinical Neuroscience, Psychiatry, and Neurology. *Trends Neurosci*, *40*(4), 200-207. https://doi.org/10.1016/j.tins.2017.02.002

Nievergelt, C. M., Maihofer, A. X., Klengel, T., Atkinson, E. G., Chen, C. Y., Choi, K. W., Coleman, J. R. I., Dalvie, S., Duncan, L. E., Gelernter, J., Levey, D. F., Logue, M. W., Polimanti, R., Provost, A. C., Ratanatharathorn, A., Stein, M. B., Torres, K., Aiello, A. E., Almli, L. M., . . . Koenen, K. C. (2019). International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci. *Nat Commun*, *10*(1), 4558. https://doi.org/10.1038/s41467-019-12576-w

Okbay, A., Baselmans, B. M., De Neve, J. E., Turley, P., Nivard, M. G., Fontana, M. A., Meddens, S. F., Linner, R. K., Rietveld, C. A., Derringer, J., Gratten, J., Lee, J. J., Liu, J. Z., de Vlaming, R., Ahluwalia, T. S., Buchwald, J., Cavadino, A., Frazier-Wood, A. C., Furlotte, N. A., . . . Cesarini, D. (2016). Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat Genet*, *48*(6), 624-633. https://doi.org/10.1038/ng.3552

Otowa, T., Hek, K., Lee, M., Byrne, E. M., Mirza, S. S., Nivard, M. G., Bigdeli, T., Aggen, S. H., Adkins, D., Wolen, A., Fanous, A., Keller, M. C., Castelao, E., Kutalik, Z., der Auwera, S. V., Homuth, G., Nauck, M., Teumer, A., Milaneschi, Y., . . . Hettema, J. M. (2016). Meta-analysis of genome-wide association studies of anxiety disorders. *Mol Psychiatry*, *21*(10), 1485. https://doi.org/10.1038/mp.2016.11

Pasman, J. A., Verweij, K. J. H., Gerring, Z., Stringer, S., Sanchez-Roige, S., Treur, J. L., Abdellaoui, A., Nivard, M. G., Baselmans, B. M. L., Ong, J. S., Ip, H. F., van der Zee, M. D., Bartels, M., Day, F. R., Fontanillas, P., Elson, S. L., andMe Research, T., de Wit, H., Davis, L. K., . . . Vink, J. M. (2019). Author Correction: GWAS of lifetime

cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal effect of schizophrenia liability. *Nat Neurosci*, *22*(7), 1196. https://doi.org/10.1038/s41593-019-0402-7

Paus, T., Keshavan, M., & Giedd, J. N. (2008). Why do many psychiatric disorders emerge during adolescence? *Nat Rev Neurosci*, *9*(12), 947-957. https://doi.org/10.1038/nrn2513

Perry, A., Roberts, G., Mitchell, P. B., & Breakspear, M. (2019). Correction: Connectomics of bipolar disorder: a critical review, and evidence for dynamic instabilities within interoceptive networks. *Mol Psychiatry*, *24*(9), 1398. https://doi.org/10.1038/s41380-018-0327-7

Phelps, E. A. (2004). Human emotion and memory: interactions of the amygdala and hippocampal complex. *Curr Opin Neurobiol*, *14*(2), 198-202. https://doi.org/10.1016/j.conb.2004.03.015

Rezin, G. T., Amboni, G., Zugno, A. I., Quevedo, J., & Streck, E. L. (2009). Mitochondrial dysfunction and psychiatric disorders. *Neurochem Res*, *34*(6), 1021-1029. https://doi.org/10.1007/s11064-008-9865-8

Rodevand, L., Bahrami, S., Frei, O., Chu, Y., Shadrin, A., O'Connell, K. S., Smeland, O. B., Elvsashagen, T., Hindley, G. F. L., Djurovic, S., Dale, A. M., Lagerberg, T. V., Steen, N. E., & Andreassen, O. A. (2021). Extensive bidirectional genetic overlap between bipolar disorder and cardiovascular disease phenotypes. *Transl Psychiatry*, *11*(1), 407. https://doi.org/10.1038/s41398-021-01527-z

Rubin, R. D., Watson, P. D., Duff, M. C., & Cohen, N. J. (2014). The role of the hippocampus in flexible cognition and social behavior. *Front Hum Neurosci*, *8*, 742. https://doi.org/10.3389/fnhum.2014.00742

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, *52*(3), 1059-1069. https://doi.org/10.1016/j.neuroimage.2009.10.003

Rudie, J. D., Brown, J. A., Beck-Pancer, D., Hernandez, L. M., Dennis, E. L., Thompson, P. M., Bookheimer, S. Y., & Dapretto, M. (2012). Altered functional and structural brain network organization in autism. *Neuroimage Clin*, *2*, 79-94. https://doi.org/10.1016/j.nicl.2012.11.006

Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C. A., Nagel, M., Awasthi, S., Barr, P. B., Coleman, J. R. I., Grasby, K. L., Hammerschlag, A. R., Kaminski, J. A., Karlsson, R., Krapohl, E., Lam, M., Nygaard, M., Reynolds, C. A., Trampush, J. W., . . . Posthuma, D. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat Genet*, *50*(7), 912-919. https://doi.org/10.1038/s41588-018-0152-6

Seghier, M. L. (2013). The angular gyrus: multiple functions and multiple subdivisions. *Neuroscientist*, *19*(1), 43-61. https://doi.org/10.1177/1073858412440596

Shao, L., Martin, M. V., Watson, S. J., Schatzberg, A., Akil, H., Myers, R. M., Jones, E. G., Bunney, W. E., & Vawter, M. P. (2008). Mitochondrial involvement in psychiatric disorders. *Ann Med*, *40*(4), 281-295. https://doi.org/10.1080/07853890801923753

Shen, H., Gelaye, B., Huang, H., Rondon, M. B., Sanchez, S., & Duncan, L. E. (2020). Polygenic prediction and GWAS of depression, PTSD, and suicidal ideation/self-harm in a Peruvian cohort. *Neuropsychopharmacology*, *45*(10), 1595-1602. https://doi.org/10.1038/s41386-020-0603-5

Smith, R. E., Tournier, J. D., Calamante, F., & Connelly, A. (2013). SIFT: Spherical-deconvolution informed filtering of tractograms. *Neuroimage*, *67*, 298-312. https://doi.org/10.1016/j.neuroimage.2012.11.049

Smith, S. M., Nichols, T. E., Vidaurre, D., Winkler, A. M., Behrens, T. E., Glasser, M. F., Ugurbil, K., Barch, D. M., Van Essen, D. C., & Miller, K. L. (2015). A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nat Neurosci*, *18*(11), 1565-1567. https://doi.org/10.1038/nn.4125

Sotiropoulos, S. N., & Zalesky, A. (2019). Building connectomes using diffusion MRI: why,

how and but. *NMR Biomed*, *32*(4), e3752. https://doi.org/10.1002/nbm.3752

Stahl, E. A., Breen, G., Forstner, A. J., McQuillin, A., Ripke, S., Trubetskoy, V., Mattheisen, M., Wang, Y., Coleman, J. R. I., Gaspar, H. A., de Leeuw, C. A., Steinberg, S., Pavlides, J. M. W., Trzaskowski, M., Byrne, E. M., Pers, T. H., Holmans, P. A., Richards, A. L., Abbott, L., . . . Bipolar Disorder Working Group of the Psychiatric Genomics, C. (2019). Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet*, *51*(5), 793-803. https://doi.org/10.1038/s41588-019-0397-8

Suprano, I., Kocevar, G., Stamile, C., Hannoun, S., Fourneret, P., Revol, O., Nusbaum, F., & Sappey-Marinier, D. (2020). White matter microarchitecture and structural network integrity correlate with children intelligence quotient. *Sci Rep*, *10*(1), 20722. https://doi.org/10.1038/s41598-020-76528-x

Sweatt, J. D. (2004). Hippocampal function in cognition. *Psychopharmacology (Berl)*, *174*(1), 99-110. https://doi.org/10.1007/s00213-004-1795-9

Tamnes, C. K., Roalf, D. R., Goddings, A. L., & Lebel, C. (2018). Diffusion MRI of white matter microstructure development in childhood and adolescence: Methods, challenges and progress. *Dev Cogn Neurosci*, *33*, 161-175. https://doi.org/10.1016/j.dcn.2017.12.002

Taquet, M., Smith, S. M., Prohl, A. K., Peters, J. M., Warfield, S. K., Scherrer, B., & Harrison, P. J. (2021). A structural brain network of genetic vulnerability to psychiatric illness. *Mol Psychiatry*, *26*(6), 2089-2100. https://doi.org/10.1038/s41380-020-0723-7

Tooley, U. A., Mackey, A. P., Ciric, R., Ruparel, K., Moore, T. M., Gur, R. C., Gur, R. E., Satterthwaite, T. D., & Bassett, D. S. (2020). Associations between Neighborhood SES and Functional Brain Network Development. *Cereb Cortex*, *30*(1), 1-19. https://doi.org/10.1093/cercor/bhz066

Torkamani, A., Wineinger, N. E., & Topol, E. J. (2018). The personal and clinical utility of polygenic risk scores. *Nat Rev Genet*, *19*(9), 581-590. https://doi.org/10.1038/s41576-018-0018-x

Tournier, J. D., Calamante, F., & Connelly, A. (2007). Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *Neuroimage*, *35*(4), 1459-1472. https://doi.org/10.1016/j.neuroimage.2007.02.016

Tournier, J. D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C. H., & Connelly, A. (2019). MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *Neuroimage*, *202*, 116137. https://doi.org/10.1016/j.neuroimage.2019.116137

van den Heuvel, M. P., & Fornito, A. (2014). Brain networks in schizophrenia. *Neuropsychol Rev*, *24*(1), 32-48. https://doi.org/10.1007/s11065-014-9248-7

van den Heuvel, M. P., & Sporns, O. (2011). Rich-club organization of the human connectome. *J Neurosci*, *31*(44), 15775-15786. https://doi.org/10.1523/JNEUROSCI.3539-11.2011

van den Heuvel, M. P., Sporns, O., Collin, G., Scheewe, T., Mandl, R. C., Cahn, W., Goni, J., Hulshoff Pol, H. E., & Kahn, R. S. (2013). Abnormal rich club organization and functional brain dynamics in schizophrenia. *JAMA Psychiatry*, *70*(8), 783-792. https://doi.org/10.1001/jamapsychiatry.2013.1328

van den Heuvel, M. P., van Soelen, I. L., Stam, C. J., Kahn, R. S., Boomsma, D. I., & Hulshoff Pol, H. E. (2013). Genetic control of functional brain network efficiency in children. *Eur Neuropsychopharmacol*, *23*(1), 19-23. https://doi.org/10.1016/j.euroneuro.2012.06.007

van Wijk, B. C., Stam, C. J., & Daffertshofer, A. (2010). Comparing brain networks of different size and connectivity density using graph theory. *PLoS One*, *5*(10), e13701. https://doi.org/10.1371/journal.pone.0013701

Visser, M., Jefferies, E., Embleton, K. V., & Lambon Ralph, M. A. (2012). Both the middle temporal gyrus and the ventral anterior temporal area are crucial for multimodal

semantic processing: distortion-corrected fMRI evidence for a double gradient of information convergence in the temporal lobes. *J Cogn Neurosci*, *24*(8), 1766-1778. https://doi.org/10.1162/jocn_a_00244

Walters, R. K., Polimanti, R., Johnson, E. C., McClintick, J. N., Adams, M. J., Adkins, A. E., Aliev, F., Bacanu, S. A., Batzler, A., Bertelsen, S., Biernacka, J. M., Bigdeli, T. B., Chen, L. S., Clarke, T. K., Chou, Y. L., Degenhardt, F., Docherty, A. R., Edwards, A. C., Fontanillas, P., . . . Agrawal, A. (2018). Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat Neurosci*, *21*(12), 1656-1669. https://doi.org/10.1038/s41593-018-0275-1

Wang, H. T., Smallwood, J., Mourao-Miranda, J., Xia, C. H., Satterthwaite, T. D., Bassett, D. S., & Bzdok, D. (2020). Finding the needle in a high-dimensional haystack: Canonical correlation analysis for neuroscientists. *Neuroimage*, *216*, 116745. https://doi.org/10.1016/j.neuroimage.2020.116745

Watson, H. J., Yilmaz, Z., Thornton, L. M., Hubel, C., Coleman, J. R. I., Gaspar, H. A., Bryois, J., Hinney, A., Leppa, V. M., Mattheisen, M., Medland, S. E., Ripke, S., Yao, S., Giusti-Rodriguez, P., Anorexia Nervosa Genetics, I., Hanscombe, K. B., Purves, K. L., Eating Disorders Working Group of the Psychiatric Genomics, C., Adan, R. A. H., . . . Bulik, C. M. (2019). Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nat Genet*, *51*(8), 1207-1214. https://doi.org/10.1038/s41588-019-0439-2

Witten, D. M., Tibshirani, R., & Hastie, T. (2009). A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostatistics*, *10*(3), 515-534. https://doi.org/10.1093/biostatistics/kxp008

Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., Adams, M. J., Agerbo, E., Air, T. M., Andlauer, T. M. F., Bacanu, S. A., Baekvad-Hansen, M., Beekman, A. F. T., Bigdeli, T. B., Binder, E. B., Blackwood, D. R. H., Bryois, J., Buttenschon, H. N., Bybjerg-Grauholm, J., . . . Major Depressive Disorder Working Group of the Psychiatric Genomics, C. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet*, *50*(5), 668-681. https://doi.org/10.1038/s41588-018-0090-3

Xia, C. H., Ma, Z., Ciric, R., Gu, S., Betzel, R. F., Kaczkurkin, A. N., Calkins, M. E., Cook, P. A., Garcia de la Garza, A., Vandekar, S. N., Cui, Z., Moore, T. M., Roalf, D. R., Ruparel, K., Wolf, D. H., Davatzikos, C., Gur, R. C., Gur, R. E., Shinohara, R. T., . . . Satterthwaite, T. D. (2018). Linked dimensions of psychopathology and connectivity in functional brain networks. *Nat Commun*, *9*(1), 3003. https://doi.org/10.1038/s41467-018-05317-y

Xu, L., Skoularidou, M., Cuesta-Infante, A., & Veeramachaneni, K. (2019). Modeling tabular data using conditional gan. *Advances in neural information processing systems*, *32*.

Zuccoli, G. S., Saia-Cereda, V. M., Nascimento, J. M., & Martins-de-Souza, D. (2017). The Energy Metabolism Dysfunction in Psychiatric Disorders Postmortem Brains: Focus on Proteomic Evidence. *Front Neurosci*, *11*, 493. https://doi.org/10.3389/fnins.2017.00493