Shared genetic influences on resting-state functional networks of the brain

JPOFT Guimaraes [a, b, c,*], **E Sprooten** [a, b, c], **CF Beckmann** [a, b, d], **B Franke** [b, c, e], **J Bralten** [b, c]

[a] Department of Cognitive Neuroscience, Radboud University Medical Center, 6525 EN Nijmegen, The Netherlands; [b] Donders Institute for Brain, Cognition and Behaviour, Radboud University, 6525 EN Nijmegen, The Netherlands; [c] Department of Human Genetics, Radboud University Medical Center, 6525 GA Nijmegen, The Netherlands; [d] Centre for Functional MRI of the Brain (FMRIB), University of Oxford, Oxford OX3 9DU, United Kingdom; [e] Department of Psychiatry, Radboud University Medical Center, 6525 GA Nijmegen, The Netherlands

* Correspondent Author

Abstract

The amplitude of activation in brain resting state networks (RSNs), measured with resting-state functional MRI, is heritable and genetically correlated across RSNs, indicating pleiotropy. Recent univariate genome-wide association studies (GWAS) explored the genetic underpinnings of RSNs, but do not describe their pleiotropic nature. In this study, we used multivariate genomic structural equation modelling to model latent factors that capture shared genomic influences. Using summary statistics from GWAS of 21 RSNs in the UK Biobank (N = 21,081) sample, we show that their genetic organization can be best explained by two distinct but correlated genetic factors that divide multimodal association networks and sensory networks. We identify single-nucleotide polymorphisms and genes associated with the joint architecture of resting-state networks. We conclude that multivariate models of genetic structures can help us to learn more about the biological mechanisms involved in brain function.

Introduction

The human brain is a complex system comprised of networks of regions that are interconnected in terms of their function [1–3]. At rest, brain function can be assessed using resting-state functional magnetic resonance imaging (rfMRI), which uses a blood oxygenation level dependent (BOLD) signal to indirectly measure synchronicity in the metabolic activity of brain regions [4,5]. Studies investigating rfMRI show that sets of brain regions are highly synchronized in their spontaneous BOLD activity, forming so-called resting-state networks (RSNs) [1–3]. An extensive body of literature shows that activity in RSNs is phenotypically associated with the incidence of neuropsychiatric disorders [6–10]. More recently, RSNs were also linked to physical factors portrayed by anthropometric, cardiac, and bone density traits [11].

RSN activation is heritable [12,13], as demonstrated by twin and pedigree studies (i.e., broad-sense heritability; $0.23 < h^2 < 0.97$) [12,14–16] as well as based on the effect of single nucleotide polymorphisms (SNPs) in unrelated individuals, i.e. SNP-based heritability ($0.05 < h^2_{SNP} < 0.28$) [13,17]. Most of these studies measured the heritability of functional connectivity based on correlations of BOLD timeseries within and between RSNs. However, RSN activity can also be captured by the amplitude of BOLD fluctuations [18,19], which were reported in Elliot et al. [13] to have on average higher SNP-based heritability estimates than correlation-based measures ($0.14 < h^2_{SNP} < 0.36$) [13]. The genome-wide association study (GWAS) of BOLD amplitude conducted by Elliot et al. [13] led to the discovery of the first genomic loci associated with individual RSNs: seven RSNs covering prefrontal, parietal and temporal cortices were associated with SNPs in the gene *PLCE1*; four RSNs mainly covering prefrontal regions were associated with the same three intergenic variants (rs7080018, rs11596664, rs67221163) in chromosome 10; one genome-wide association with a single sensorimotor RSN involved the intronic variant rs60873293.

Next to an overlap in single genetic variants involved in multiple RSNs, significant genetic correlations between different RSNs have been reported using bivariate GWAS analysis [13] and twin models [15,16], with

observed correlations between 0.37 and 0.79. These results suggest that RSNs are driven by shared genetic variation, indicating the potential for pleiotropy, i.e. the same genetic variants being involved in the etiology of different RSNs. One twin study conducted by Reineberg et al. [15] showed that the heritability of brain connectivity within and across RSNs is represented by three clusters, of which one was defined by low-heritability connections, and two clusters of heritable connections. The latter two can be described as one cluster comprising connections with high heritability in the visual cortex (i.e. "sensory" regions) and a second cluster comprising associations among default mode, frontoparietal, salience, dorsal and ventral attention regions (i.e. "multimodal association" regions, which integrate inputs from multiple sensory modalities). This broad division of the connectome into sensory networks and multimodal association networks is in line with what was previously found on the basis of clustering of BOLD amplitude across RSNs as well [18,19]. Based on these results, we hypothesize that RSNs genetically diverge according to their "sensory" or "multimodal association" functions.

To identify the SNPs and genes driving this observed pleiotropy multivariate methods can be applied [20–22]. Grotzinger et al. [20] used a novel technique called genomic structural equation modelling (genomic SEM) to model a single genetic factor capturing GWAS associations across multiple psychiatric diagnoses. This multivariate GWAS approach led to the discovery of SNPs that were not observed by any of the separate univariate GWASs of any of the disorders [20]. In this way, multivariate GWAS provides a new, statistically powerful opportunity to directly characterize the genomic influence on multiple phenotypes simultaneously. Given the observed pleiotropy between brain activation in RSNs, the same approach can be applied to discover the SNPs most associated with shared genetic effects on brain function.

In the current study, we investigated shared genetic etiologies of multiple RSNs within the brain. We used GWAS summary statistics for the amplitude of 21 RSNs throughout the brain made available by the UK Biobank (N = 21,081) [11,13,23,24]. First, we estimated the SNP-heritability of the selected RSNs, and for heritable RSNs we modelled their shared genetic structure using genomic SEM [20]. Next, we performed

multivariate GWASs to characterize the SNPs associated with these pleiotropic factors. Our multivariate GWAS findings were further interpreted via functional annotation of top GWAS loci and gene-mapping with the Functional Mapping and Annotation (FUMA) tool [25] and gene-wide and gene-set analysis in MAGMA [26]. Finally, we also tested whether the newly found genomic factors were genetically correlated with psychiatric and physical traits.

Results

SNP-based heritability of RSNs. We obtained GWAS summary statistics of BOLD amplitudes of ten multimodal association and eleven sensory RSNs (Fig. 1, Supplementary Fig. 1, and Supplementary Table 1) measured in 21,081 adult individuals from the UK Biobank cohort (see Methods: GWAS sample). Nineteen of the twenty-one RSN amplitudes showed a significant SNP-based heritability (h^2_{SNP}; Adjusted P(FDR) ≤ 0.05; Supplementary Fig. 1 and Supplementary Table 1), with estimates of 0.05-0.17. Two sensory networks involved in secondary visual processing had non-significant h^2_{SNP} estimates (h^2_{SNP} = 0.036 and 0.038) and were excluded from subsequent analyses. On average, multimodal association networks showed a higher h^2_{SNP} than sensory networks (average h^2_{SNP} = 0.13 and 0.07, respectively).

Genetic correlations between RSNs. To test the existence of shared genetic etiologies between the heritable RSNs, we calculated genetic correlations using Linkage Disequilibrium Regression Analysis [27] available within the genomic SEM package [20]. Figure 2 displays the 171 genetic correlations between the heritable RSNs, of which 36 reached Bonferroni-level significance (P(Bonferroni) <= 0.05/171 = 3E-4), and 63 other genetic correlations showed uncorrected p-values < 0.05. The Bonferroni and the 'nominally' significant genetic correlations were predominantly positive (97 out of 99), ranging from 0.29 to 0.84; whereas two nominally-significant correlations had negative coefficients: SN11 with MA4 (r_g = -0.44, p = 0.01) and SN11 with MA5 (r_g = -0.59, p = 0.03). For more details on the genetic correlation values and respective standard errors and p-values, see Supplementary Table 2.

Genomic structural equation modelling. To characterize the common underlying genetic etiologies between the nineteen heritable RSNs, we derived latent genomic factors using genomic SEM [20]. We chose the most optimal model on the basis of exploratory factor analysis (EFA). The EFA results are summarized in Fig. 3, showing that the two-factor model explained 17% more variance than the one-factor model, while the addition of a third factor did not explain substantially more variance [28,29]. These

results indicate that the most optimal model in representing the pleiotropy among these RSNs consists of two factors. Therefore, we continued our main analysis on the two-factor model. For completeness, we also report results of the one-factor model, which also had a reasonable fit, for each step in the supplementary data (Supplementary Tables 3-5, and Supplementary Fig. 2). Factor loadings for the two-factor model are available in Supplementary Table 6.

We used confirmatory factor analysis (CFA) to test the model fit of the two-factor model. The model fit estimates are reported in Table 1, organized in two sets: (i) fit estimates reported for the model retrieved by EFA (see top row in Table 1); and (ii) fit estimates for the corrected model after excluding non-significant factor loadings (see bottom row in Table 1). By comparing chi-square and Akaike information criterion (AIC) statistics between the two sets, we observed that excluding non-significant factor loadings from the model led to lower values retrieved by both statistics, and thus an improved model fit.

In Figure 4, we show the path diagram of the two-factor model, where the pleiotropy among RSNs is represented by two distinct but correlated factors ($r = 0.48$, $p = 1.31E\text{-}5$). The first factor (F1) comprises all ten multimodal association networks (MA1-10) and two sensory networks (SN4 and SN10); whereas the second factor (F2) consists of five sensory networks (SN5-9). The primary visual network SN11 was the only RSN whose amplitude was associated with both factors, reporting positive and negative factor loadings in F2 and F1, respectively. In Supplementary Table 7, we include the nominal and Bonferroni-corrected p-values of the factor loadings shown in Fig. 4. Neither of the factor loadings of the primary visual network SN11 were significant upon Bonferroni correction ($P(\text{Bonferroni}) <= 0.05/19 = 0.0026$), leading to the exclusion of SN11 from the multivariate GWAS model.

Multivariate GWAS Results. We estimated the SNP effects driving the pleiotropy of RSNs using multivariate GWASs of the two latent genetic factors. Both F1 and F2 showed significant SNP-based

heritability (F1: h^2_{SNP} = 0.20, SE = 0.028, P = 3.98E-13; F2: h^2_{SNP} = 0.11 , SE = 0.020, P = 6.87E-08). For F1, 142 SNPs, encompassing 3 genomic loci, showed genome-wide significance (P < 5E-8). Of these SNPs, none had a genome-wide significant Q_{SNP} statistics (P < 5E-8), indicating that these SNPs are driven by multiple RSNs associated with F1, rather than by any specific RSN. All the genomic loci have been reported in the univariate GWASs [13], however, the association with the locus on chromosome 17 has not previously reached genome-wide significance corrected across all the tested traits. For F2, no SNP reached genome-wide significance (Supplementary Fig. 3).

Functional characterization of top GWAS loci. We interpreted our multivariate GWAS results by conducting functional annotation and gene-mapping of genomic loci using FUMA [25]. In addition to 142 genome-wide significant SNPs reported for F1, FUMA analysis identified 131 other SNPs in LD with these genome-wide significant SNPs, making a total 273 candidate SNPs distributed among three genomic loci (see Table 2). With the functional annotation of these candidate SNPs, we mapped a total of 51 genes using positional, eQTL (Adjusted P(FDR) <= 0.05), and chromatin interaction mapping (Adjusted P(FDR) <= 1e-6), as reported in Supplementary Tables 8-10. Supplementary Table 11 contains a list of studies from the GWAS Catalog reporting genome-wide significant SNPs that map to these genomic loci.

Gene-wide and gene-set results. To investigate whether our multivariate SNP-associations aggregated in a biologically meaningful way, we performed gene-wide and gene-set association analyses for F1 and F2 using MAGMA [26]. We found eight genome-wide significant genes associated with F1 (Fig. 5): *FHL5*, *PLCE1, HPS1, ANO1*, *EPN2*, *B9D1*, *MAPK7*, and *HIC1*. Furthermore, F1 was significantly associated with one gene-set: a biological process involving lipoprotein clearance from the blood via receptor-mediated endocytosis (Supplementary Table 12). For F2, we discovered one gene-wide association for *SMYD3*, but no significant gene-sets (Supplementary Table 13). Additionally, we investigated whether the genes associated with F1 and F2 were enriched in 30 general human tissue types analyzed via MAGMA tissue

expression profile analysis. No significant enrichment was found for F1 or F2 genes (Supplementary Fig. 4).

Genetic correlations with neuropsychiatric and physical traits. To examine shared genetic effects between the two multivariate RSN factors and ten pre-selected neuropsychiatric and physical traits we performed genetic correlation analyses with GWAS summary statistics. The genetic correlation results are reported in Fig. 6. No genetic correlation reached significance after multiple comparison correction (Adjusted P(FDR) <= 0.05). One genetic correlation between F2 and BMI showed nominal significance (P <= 0.05). For more details on the genetic correlation values and respective standard errors and p-values, see Supplementary Table 14.

Discussion

We investigated the genomic basis of pleiotropy of brain function in 21 resting-state networks across the brain. We discovered that two latent genetic factors best captured the genomic influence on resting state activity throughout the brain. The first factor was associated with multimodal association networks and two sensory networks; the second factor represented only sensory networks. Further, we found that the first factor was associated with SNPs, genes, and a gene-set with implications for our understanding of the molecular basis of brain function. With the statistical power provided by our multivariate genomic approach, we were able to characterize SNP effects relevant to general brain function over and above individual GWAS of RSN amplitude.

Our genomic factor analyses point to a genetic divergence of multimodal association and sensory functions. This distinction is in line with previous studies using functional connectivity measures [15] and phenotypic analyses of RSN amplitudes [18,19]. Brain regions involved in sensory and multimodal association functions have also been found to differ in cytoarchitectonic properties [30]. For example, sensory cortical areas contain higher concentrations of myelin compared to higher order association areas [31–33]. Furthermore, sensory and multimodal association areas exhibit distinct patterns of gene expression [34]. Together with our findings, the extensive evidence of genetic and brain differences between these two factors may potentially reflect the known differences in the period of maturation between their respective brain regions [35]. In addition, in evolution humans show more pronounced cortical expansion in multimodal association networks than they do in sensory networks compared to other primate species [36,37]. Thus, differences between sensory and multimodal brain networks have been consistently indicated across biological disciplines from neurodevelopment, to neurophysiology, to evolution. With our findings, we suggest that this divergence between the sensory and multimodal association systems may also be represented by partly distinct effects of common genetic variation.

The first factor of general brain function was associated with a total of 273 candidate SNPs distributed among three genomic loci, of which one locus, on chromosome 17, was not previously detected by their univariate GWAS when corrected for multiple testing across traits and SNPs [13]. We also detected eight gene-wide associations, seven of which were not previously detected by the univariate GWAS [13], and we were able to functionally map 47 additional genes relevant to the study of brain physiology. We thus demonstrate that a multivariate genomic approach has additional value in the search for genetic underpinnings of brain function.

The gene *EPN2*, which encodes a protein involved in notch signaling endocytosis pathways, showed up in our gene-wide analysis and has previously been associated with educational attainment [38,39] and schizophrenia [40]; this indicates that notch signaling, known for its role in neurodevelopment and the onset of psychiatric disorders [41], may also have an influence on general brain function in adulthood. We also reported a significant gene-set for the mainly multimodal association factor, namely triglyceride-rich lipoprotein particle clearance, which contains nine genes including *APOE*, the most well-known risk-gene for Alzheimer's disease [42,43]. A possible role of neurodegenerative processes in this factor was also suggested by the gene-wide association with *MAPK7*, which has also been reported in a GWAS of Alzheimer's disease [44]. This hypothesis is also supported by the functional mapping of *LGI1*, which was previously reported in relation to beta-amyloid measurement in cerebrospinal fluid, a biomarker for Alzheimer's disease [45,46]. An eventual link between this factor and ageing-effects potentially reflective of Alzheimer's disease was, however, not corroborated by our genetic correlation analysis with Alzheimer's disease, suggesting that this link is not reflected by association patterns at the genome-wide scale.

Interesting significant gene-wide associations also included *FHL5*, a gene previously associated with migraine [47,48], spatial memory [49], and cerebral blood flow [50], and *SMYD3*, which previously came up in GWASs of cognitive ability [51,52], suicide attempt [53], and bipolar disorder [54]. However, the genes retrieved by our gene-wide analyses were not all related to traits exclusively relevant to the brain, but also to

cardiovascular [55,56], metabolic [57,58], and drug response traits [59,60]. BOLD amplitude, being a blood-based measure, may also be susceptible to genetic effects affecting blood-related traits that are not necessarily specific to the brain. The gene-wide result for *PLCE1* is an example of such an observation, since it was previously reported for 38 other phenotypes, covering brain (e.g. migraine), cardiovascular (e.g. hypertension, blood pressure), and more general metabolic traits (e.g. BMI). The association of *PLCE1* was previously reported with seven individual RSNs in Elliott et al. [13]. The gene encodes a phospholipase enzyme involved in cell growth, cell differentiation, and regulation of gene expression.

Despite known associations of the above rfMRI-associated genes with neuropsychiatric and physical traits and previously reported phenotypic associations between these traits and rfMRI-derived imaging phenotypes [6–11], we did not detect significant genetic correlations between our two genetic factors for brain function and those other neuropsychiatric and physical traits. This suggests that eventual links between general brain function and these phenotypes are not primarily explained by additive effects of common variants across the whole genome.

This study should be viewed in light of several strengths and limitations. Strengths of our study are the use of GWAS results of large resting-state fMRI samples, which provided the power necessary to run this analysis. Furthermore, we used state of the art novel technologies to find shared genetic etiologies in summary statistics including genomic SEM, which provided statistical power-boosting through the joint analysis of GWASs. Our results provide a new, data-driven basis for studying biological pathways relevant to brain function, by integrating multiple data sources spanning genomics, epigenomics, and transcriptomics. However, this characterization was limited by the data sources that are currently available. As more resources become publicly available and integrated in FUMA and equivalent platforms, in the future an even broader genetic mapping of traits will be possible. Another limitation of our study is the fact that our approach focused exclusively on the effects of common SNPs, without including the effects of rare genetic variants or gene-environment interactions and correlations.

Including rare variation in follow-up studies and more extended explicit modelling of gene-environment interplay may provide even more insight into the biological pathways underlying brain function.

In conclusion, we show that pleiotropy in heritable RSNs is best represented by a two-factor model mainly distinguishing the genetic influences on multimodal association from those on sensory networks. GWAS-based analysis of these genetic factors led to the discovery of relevant SNPs, genes, and a gene-set. With our approach, we demonstrate that taking advantage of the pleiotropy of RSNs using multivariate genome-wide approach provides new insights in the genetic and molecular roots of brain function.

Methods & Materials

GWAS Sample. We used GWAS summary statistics from the UK Biobank initiative, publicly available in a second release via **Oxford Brain Imaging Genetics Server** (open.win.ox.ac.uk/ukbiobank/big40/; accessed on 10 June 2020). They contain GWAS results for 3,919 imaging phenotypes of brain structure and function, based on a discovery sample consisting of 22,138 unrelated individuals of UK ancestry, of which 11,624 female (Females: mean age= 63.6 ± 7.3 years; Males: mean= 65.0 ± 7.6 years). In this dataset, 21,081 individuals with available rfMRI data were included in the GWAS on the amplitude of 21 RSNs (i.e., the standard deviation of BOLD signal measured within each RSN). The MRI acquisition and analysis procedures of the brain imaging phenotypes have been described previously [11,61] and accounted for the confounders age, sex, head size, and estimated amount of head motion [62]. Genotypes were imputed with the Haplotype Reference Consortium (HRC) reference panel and a merged UK10K + 1000 Genomes reference panel as described by Bycroft et al. [24]. The GWAS summary statistics come from the study by Smith et al. [63]. This GWAS used a quality control procedure that included thresholding for minor allele frequency (MAF >= 0.001), the quality of the imputation (INFO >= 0.3), and Hardy-Weinberg Equilibrium (HWE –Log10(P) <= 7), while controlling for population structure represented by the first 40 genetic principal components. A total of 20,381,043 SNP associations were reported in the selected GWAS summary statistics, which were estimated via linear association testing in BGENIE software [24].

Description of RSNs. The 21 RSNs covering spontaneous BOLD fluctuations in the brain were labeled based on the clustering analyses conducted in Bijsterbosch et al. [19], which appointed RSNs to one of two distinct system categories: multimodal association and sensory systems. In Supplementary Table 15, we show the system category given to each RSN, the respective UK Biobank label, and its respective two-dimensional anatomical visualization. Visualization of RSNs is also provided by the UK Biobank online resources (fmrib.ox.ac.uk/ukbiobank/group_means/rfMRI_ICA_d25_good_nodes.html).

Genomic Structural Equation Modelling. Taking the GWAS summary statistics of BOLD amplitude in 21 RSNs, we modelled the potentially shared underlying genetic etiologies using a genomic factor analyses in genomic SEM package v0.0.2 in R v3.4.3, developed by Grotzinger et al. [20]. For more details see github.com/MichelNivard/GenomicSEM/wiki/3.-Models-without-Individual-SNP-effects.

First, we conducted a quality control (QC) step on the selected GWAS summary statistics that included (i) selection of SNPs reported in the HapMap3 reference panel; (ii) exclusion of SNPs located in the major histocompatibility complex (MHC) region; (iii) exclusion of SNPs with MAF lower than 1%; (iv) exclusion of SNPs with INFO scores lower than 0.9. This QC step retained a total of 1,171,392 autosomal SNPs.

Secondly, we calculated the SNP-based heritability of the 21 RSN amplitudes with LD-Score regression (LDSC v1.0.0) [27]. The univariate LDSC calculates SNP-based heritability estimates of traits, based on SNP effect sizes in relation to each SNP's linkage disequilibrium (LD) [27]. Only RSNs with FDR-corrected significant (Adjusted P(FDR)<0.05) SNP-based heritability were taken forward to the next genomic SEM steps.

In the following step, the covariance matrices estimating the pleiotropy among heritable RSN amplitudes were retrieved using the multivariate extension of LDSC distributed by the genomic SEM package. We obtained (i) a genetic covariance matrix quantifying the genetic overlap among the RSNs; (ii) the respective matrix containing the standardized genetic covariance values (i.e. genetic correlations); (iii) a sampling covariance matrix informative of the standard errors associated with the genetic covariance measures.

To determine the number of factors in the model, and which imaging phenotype loaded on which factor, we conducted an exploratory factor analysis (EFA) with maximum likelihood estimation. Before running EFA, the LDSC-derived covariance matrix was smoothed to the nearest positive, as part of the

default genomic SEM pipeline. We tested EFA with one factor and repeated the same step for an increasing number of factors up to six. We selected the highest number of factors leading to an explained variance increase (r^2) of equal or more than 10% [29]. For all the modelling results, positive or negative factor loadings with magnitudes equal or higher than 0.35 were assigned to a given factor, identical to Grotzinger et al. [20].

For the most optimal model, we ran confirmatory factor analysis (CFA) using the genomic SEM package, in order to estimate the factor loadings of the variables included in the model and evaluate the respective model fit. Both the genetic and sampling covariance matrices were analyzed using weighted least squares estimation, providing fit statistics and inferred factor loadings. We retained factor loadings at a Bonferroni significance level across the factor loadings within the model (P(Bonferroni) <= 0.05/Number of factor loadings).

Multivariate GWAS. A multivariate GWAS was conducted on the factors of the most optimal model (see above), in order to discover the SNPs driving their pleiotropy. Only SNPs reported in the 1000 Genomes phase 3 reference panel were taken forward in this step, and SNPs were excluded in case of MAF lower than 1% or INFO score lower than 0.6, as in Grotzinger et al. [20]. This analysis leads to the multivariate effect sizes and p-values for each SNP, reflecting the contribution of 8,135,328 autosomal SNPs to each factor. Additionally, for each SNP, the results included a heterogeneity statistic (Q_{SNP}) and respective p-value addressing whether the SNP effect was mediated by the common factor(s) (null hypothesis), or is specific to one of the traits (P <= 5E-8). The SNP-based heritability of genetic factors represented in each model was also estimated, using LDSC [27], following the same procedure used for the 21 RSN amplitudes (see above).

Functional Mapping Analysis. Functional annotation and gene-mapping of genomic risk loci of our multivariate GWAS results was performed using FUMA version v1.3.6 [25], an online platform used to

prioritize, annotate, and interpret GWAS summary results (access via fuma.ctglab.nl). For each multivariate GWAS, FUMA annotates SNPs that reach independent genome-wide significance (P < 5E-8), or that reach nominal significance (P < 0.05) and are in LD (r2>=0.6) with any of the independent genome-wide significant SNPs within a 250 kb window. After determining the independent significant SNPs, the lead SNP of each genomic locus is chosen according to a more stringent LD squared coefficient r^2 <= 0.1 [25]. For each independent significant SNP, FUMA retrieved information regarding the type of variant and the nearest gene, while providing for each genomic locus a GWAS Catalog list of published studies reporting genome-wide associations with SNPs located in that same locus.

Gene-mapping was performed by (i) selecting genes located within 10 kb of each SNP, (ii) annotating SNPs based on their expression quantitative trait loci (eQTL) enrichment in the data resources listed in Supplementary Table 16, and (iii) the chromatin interactions depicted in the HI-C data resources reported in Supplementary Table 17. Only FDR-corrected significant gene associations were reported based on eQTL mapping (Adjusted P(FDR) <= 0.05) and chromatin interaction mapping (Adjusted P(FDR) <= 1E-6), as recommended in FUMA [25].

Gene-wide and gene-set analyses. To test for aggregated association of multiple SNPs within genes, we performed gene-wide analyses on the multivariate GWAS results. We then performed gene-set analysis for curated gene-sets and GO terms from MsigDB c2 and c5, respectively, testing for the presence of pathways associated with these factors. Furthermore, we performed tissue gene expression analysis in the genomic factors. These analyses were all performed using the MAGMA v1.07 software [26] as embedded within the FUMA platform [25], for details see the Supplementary Methods.

Genetic correlations with other traits. To examine shared genetic effects between our RSN genomic factors and other traits, we performed genetic correlation analyses with GWAS summary statistics from ten selected traits. We followed the same QC and bivariate genetic analysis procedures used for RSN

amplitudes (see above). Based on literature associating RSN function with many neuropsychiatric disorders, we selected GWAS summary statistics reported for Alzheimer's disease [43] and for five major psychiatric disorders reported by the Psychiatric Genomics Consortium: attention deficit/hyperactivity disorder [64], autism spectrum disorder [65], bipolar disorder [54], major depressive disorder [66], and schizophrenia [60]. We also included in our analysis GWAS summary statistics reported for body-mass index [67], height [68], bone density [69], and diastolic blood pressure [70] traits, which were previously linked to RSN activation [11]. Significant genetic correlations were determined by FDR multiple comparison correction (Adjusted P(FDR) <= 0.05). For detailed information about these GWAS summary statistics, consult Supplementary Table 18.

Statistical Analysis. The significance of SNP-based heritability of BOLD amplitude in the 21 RSNs was determined via FDR correction accounting for multiple comparison bias. Bonferroni correction was used to determine the significance of factor loadings tested in CFA, correcting for the total number of factor loadings tested in each model. Genes discovered via gene mapping using eQTL enrichment and chromatin interactions were selected if their associated p-values reached the FDR-corrected threshold assigned to eQTL enrichment and chromatin interaction-based associations, respectively. Significance of findings for the gene-wide, gene-set, and tissue expression profiles was determined via a Bonferroni correction accounting for the number of tested genes, gene-sets, and tissue samples, respectively. FDR correction was used to determine the significance of the genetic correlations involving the latent genomic factors with psychiatric and physical health traits.

Acknowledgments

The research leading to these results received funding from the Radboud University Medical Center PhD program and the European Community's Horizon 2020 research and innovation programme under grant agreement no. 847879 (PRIME). This work is part of the research programme *Computing Time National Computing Facilities Processing Round pilots 2018* with project reference EINF-446, which is (partly) financed by the Dutch Research Council (NWO). This work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative. B.F. received additional funding from a grant for the Dutch National Science Agenda (NWA) for the NeurolabNL project (grant 400 17 602). E.S. is funded by a NARSAD Young Investigator Award (GRANT ID: 25034), a Hypatia Tenure Track Grant and Christine Mohrmann Fellowship (Radboudumc). Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: D.C. Van Essen and K. Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. C.F.B. gratefully acknowledges funding from the Netherlands Organisation for Scientific Research Innovation program (NWO-Vidi 864.12.003), the Wellcome Trust UK Strategic Award [098369/Z/12/Z], and the NWO Gravitation Programme Language in Interaction (grant 024.001.006). J.B. gratefully acknowledges funding from the NWO Innovation program (Veni 09150161910091). We would like to thank Andrew D. Grotzinger (University of Texas at Austin) for the support provided in conducting the genomic SEM approach.

Conflicts of interest

B.F. has received educational speaking fees from Medice. C.B. is director and shareholder of SBGneuro Ltd.

Contributions

The conception of the idea motivating this study was elaborated by JG, ES and JB. JG performed the analyses, which were supervised by JB and ES. JG, ES and JB wrote the paper with contributions from the remaining coauthors.

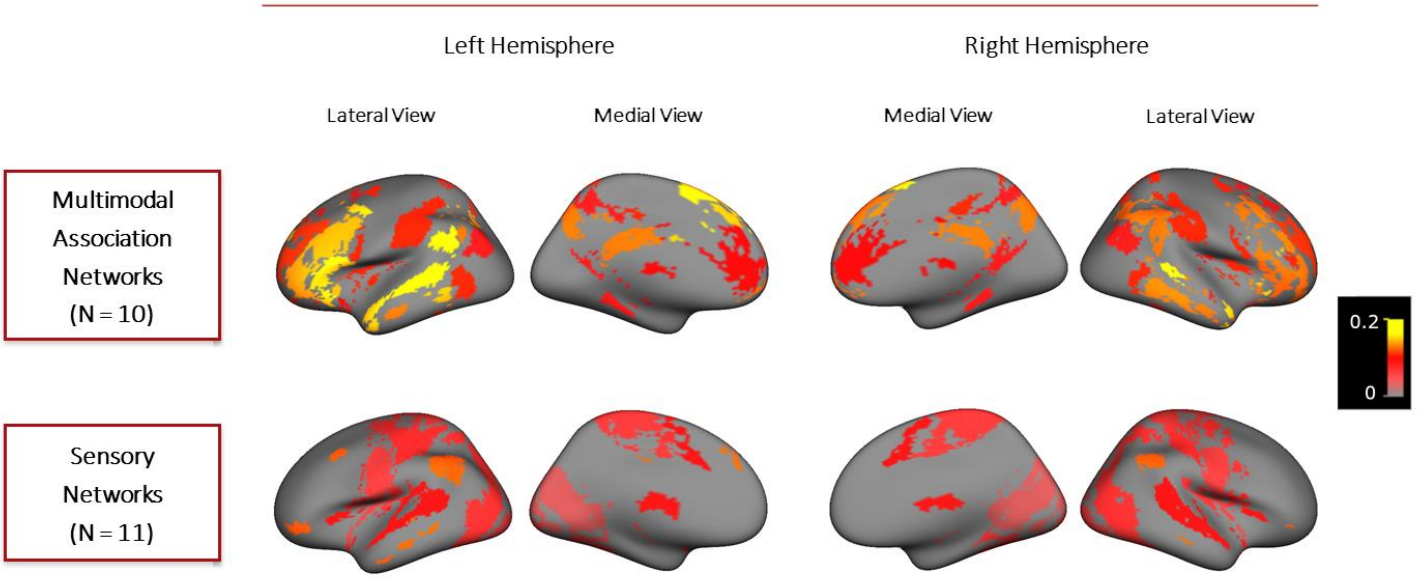


Figure 1. SNP-based heritability results obtained for RSNs of multimodal association and sensory networks.

Cortical surface maps displayed show the multimodal association (top) and sensory networks (bottom). Both the medial and lateral views of RSNs in the left and right hemispheres are displayed from left to right. RSNs are color-coded according to their SNP-based heritability - proportion of variance in the trait explained by SNP effects – whose scale is displayed on the right.

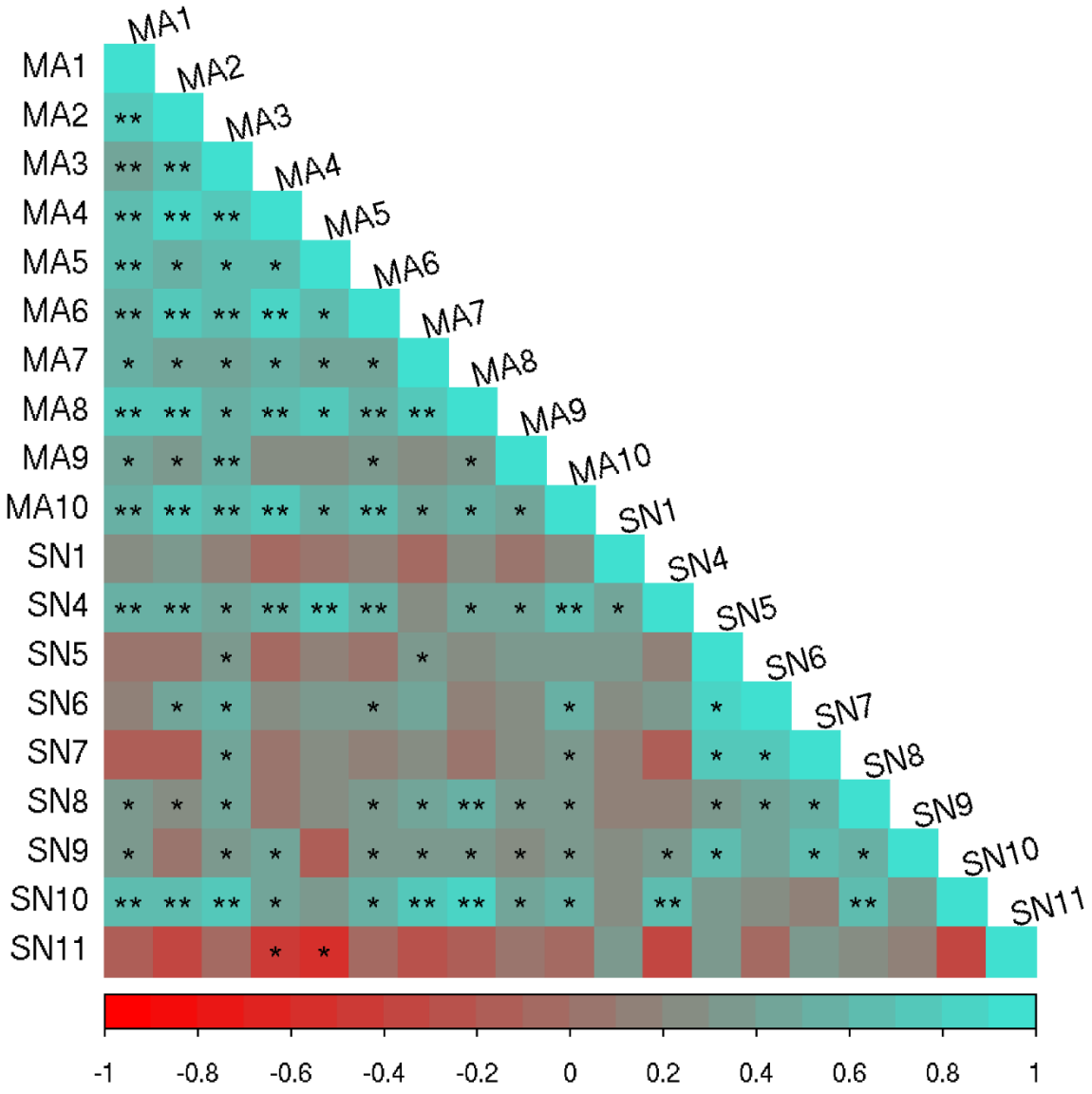
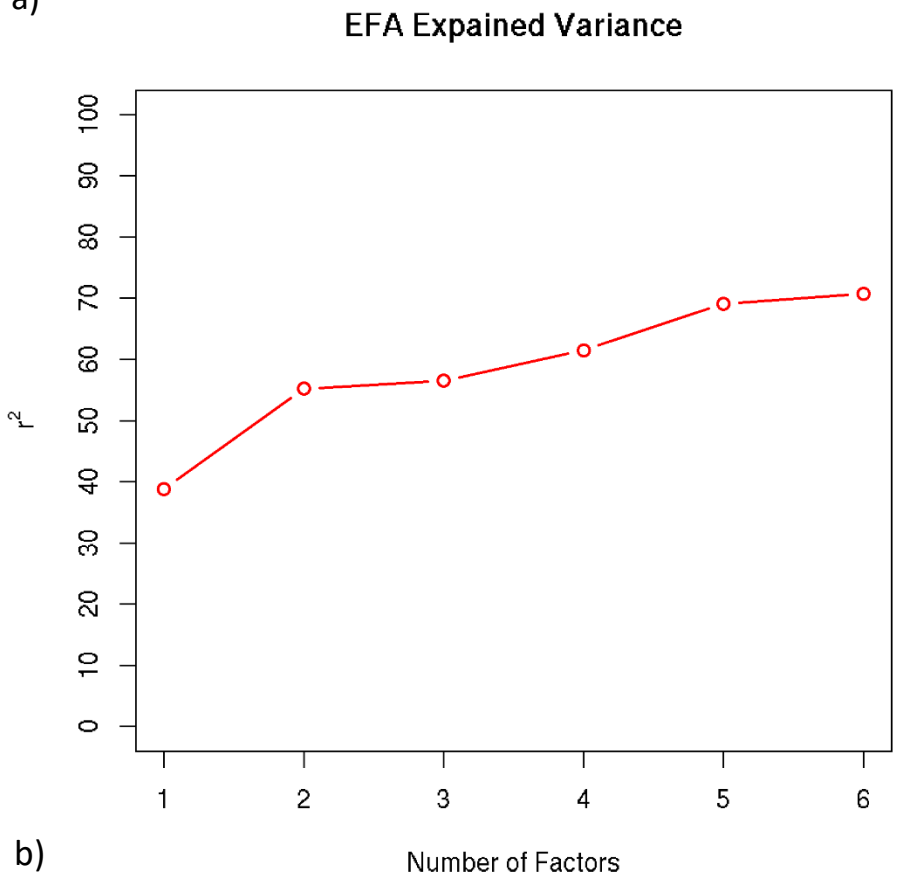


Figure 2. Genetic correlation matrix of the nineteen heritable RSN amplitudes.

Genetic correlations among multimodal association (MA) and sensory (SN) network amplitudes are represented according to the color bar displayed below, with their size being proportional to the magnitude of the correlation. SN2 and SN3 were not included due to non-significant SNP-based heritability. Genetic correlations scoring nominal and Bonferroni-corrected significance are respectively labeled with one (*) and two (**) asterisks.

a)



b)

Number of factors	1	**2**	3
Cumulative explained variance	41%	**58% (+17%)**	61% (+3%)

Figure 3. Summary of exploratory factor analysis. Plot displaying the percentage of cumulative explained variance (r^2) from up to six-factor models tested using EFA (a); Cumulative explained variance by the one, two and three factor models tested using EFA. The added explained variance corresponding to an additional factor in the model is shown in parenthesis (b).

Table 1. Summary of confirmatory factor analysis.

Model	Number of factors	Included RSNs	Chi-square statistic	Degrees of freedom	AIC	CFI	SRMR
EFA-based Model	2	18	897	133	973	0.81	0.12
Model for Multivariate GWAS	2	17	575	118	645	0.84	0.12

The first column distinguishes the two stages composing our CFA approach: the first stage, i.e. EFA-based model, tested the model design for the two-factor model indicated by our EFA approach; the second stage, i.e. modelling for Multivariate GWAS, only kept RSNs that showed Bonferroni-corrected factor loadings in the first stage. The models were tested with a distinct number of factors (second column), with a given number of RSNs (third column). For each model, we display the chi-square statistic, degrees of freedom, Akaike information criterion (AIC), CFI (comparative fit index), and standardized root mean square residual (SRMR), from the fourth to the eighth columns.

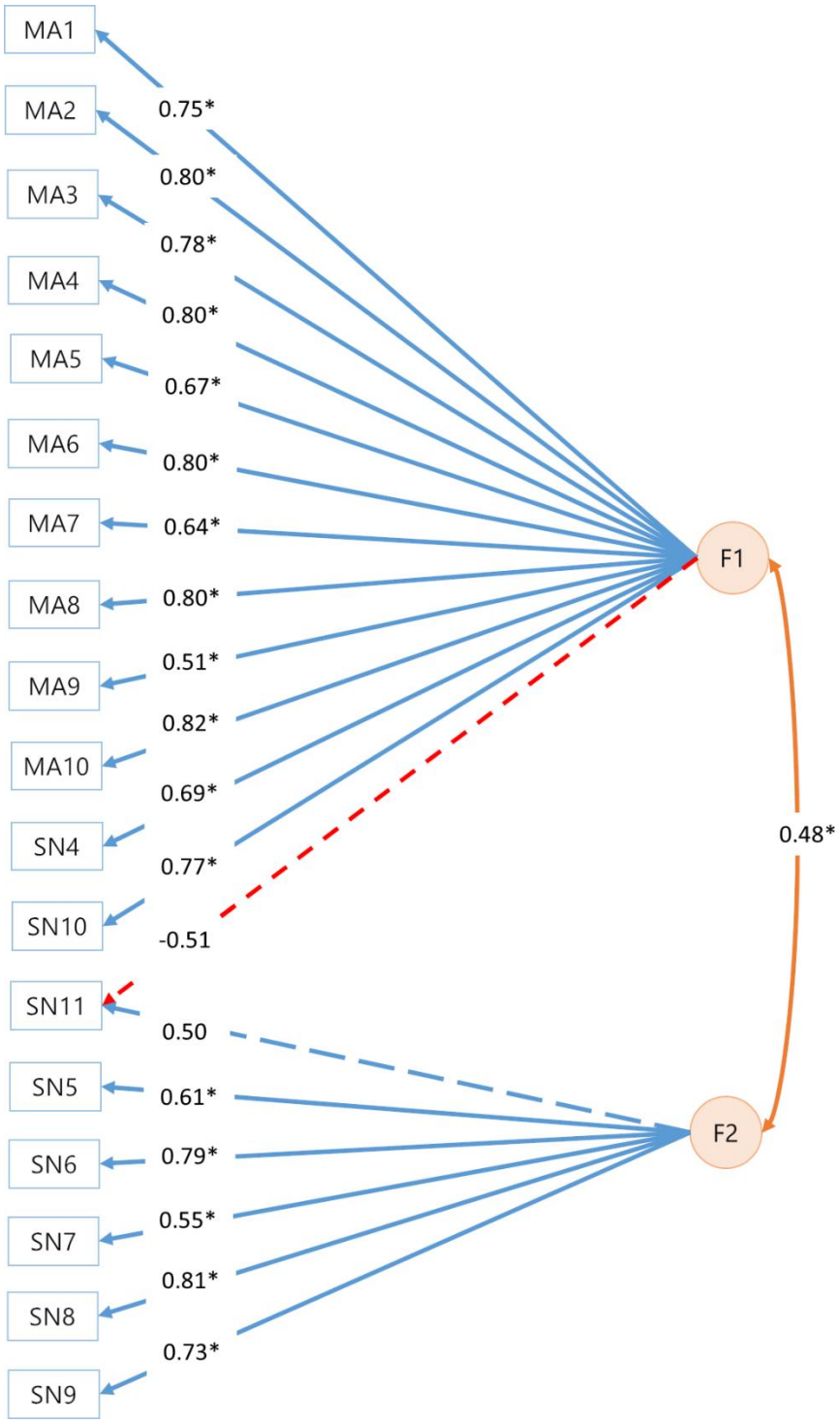


Figure 4. Path diagram of two-factor model. Orange circles represent the two latent genetic factors of the two-factor confirmatory factor analysis (CFA). Factor 1 (F1) and 2 (F2) are connected by a double-headed arrow, which represents the correlation between the two factors. F1 and F2 are associated with eighteen RSNs represented by blue rectangles, with loadings represented by a blue (positive) or red (negative) arrow. Factor correlation and loadings reaching Bonferroni-corrected significance (P(Bonferroni) <= 0.05/19 = 0.0026) are indicated by an asterisk, whereas those that do not reach Bonferroni-corrected significance are represented by dashed lines.

Table 2. LD-independent significant SNPs for latent factor F1.

Genomic Locus	rsID	Chromosome	Position	P-value	Variant Type	Nearest Gene
1	rs7069316	10	96000282	9.83e-09	Intronic	*PLCE1*
	rs10786156	**10**	**96014622**	**1.93e-11**	**Intronic**	*PLCE1*
	rs11187842	10	96052511	2.46e-08	Intronic	*PLCE1*
2	rs11596664	10	134280157	3.89e-09	Intergenic	*C10orf91*
	rs7907962	10	134287486	1.92e-09	Intergenic	*C10orf91*
	rs4880380	10	134288177	3.04e-09	Intergenic	*C10orf91*
	rs7080018	**10**	**134301505**	**8.69e-11**	**Intergenic**	***RP11-432J24.5***
	rs34102287	10	134331173	2.21e-10	Intronic	***RP11-432J24.5***
3	**rs1969161**	**17**	**19194812**	**8.84e-09**	**Intronic**	***EPN2***

For each genomic locus, the respective independent genome-wide significant SNPs are displayed. The lead SNP of each genomic locus is highlighted in bold, and information regarding the chromosome, position, genome-wide p-value, variant type, and nearest gene is provided.

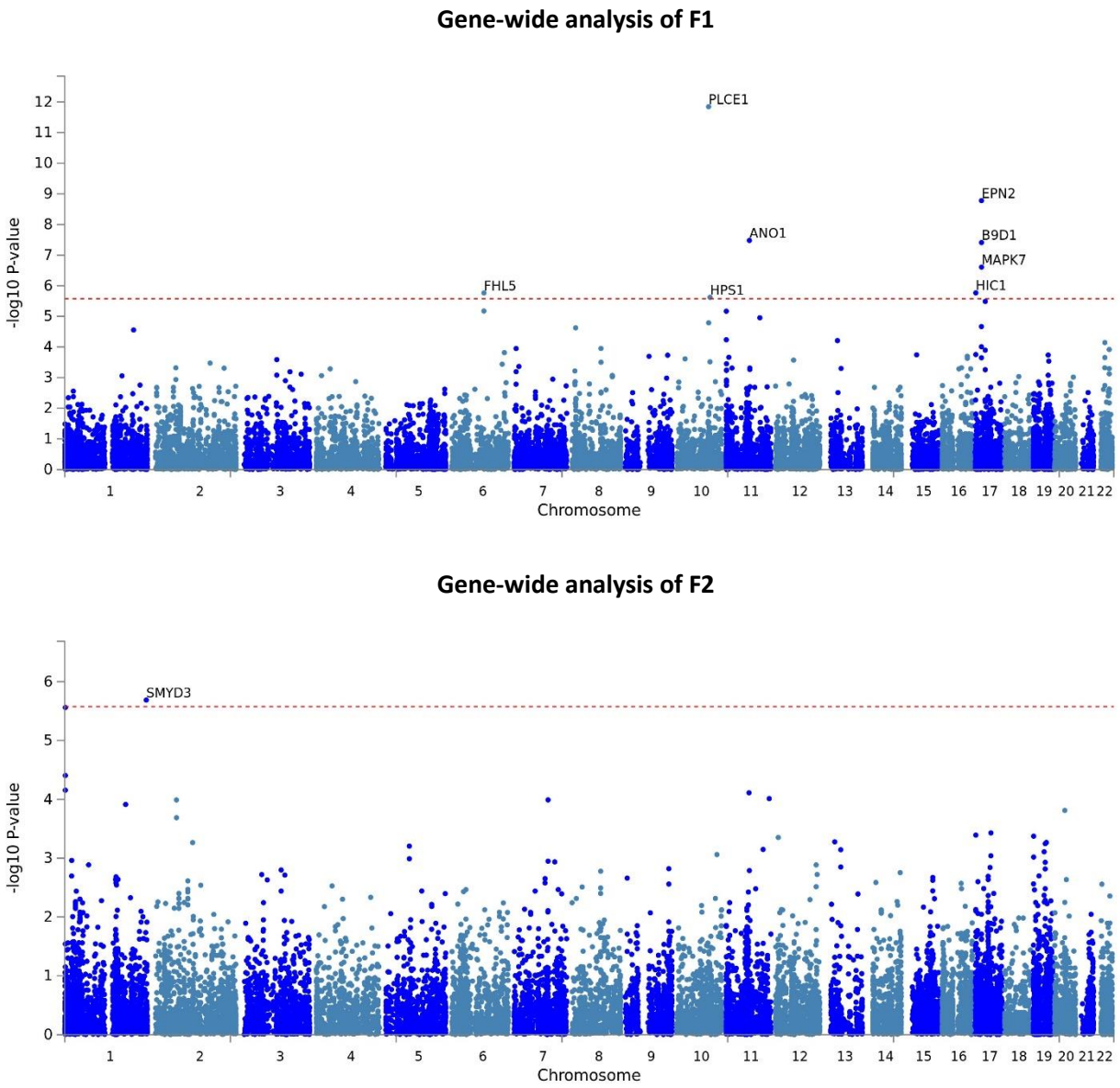


Figure 5. Manhattan plot of MAGMA gene analysis findings of latent factors F1 and F2.

Gene-wide p-values of associations in F1 (top), which comprises genetic effects shared among all ten multimodal association networks (MA1-10) and two sensory networks (SN4 and SN10); and F2 (bottom), consisted of five sensory networks (SN5-9).

In each plot, genes located across the 22 autosomes labeled along the x-axis are represented by blue dots, whose position along the y-axis represents the log p-value scored by their gene-wide association with the latent factor F1. The red-dashed horizontal line marks the Bonferroni-corrected significance for the number of genes being tested (P(Bonferroni) <= 2.64e-6).

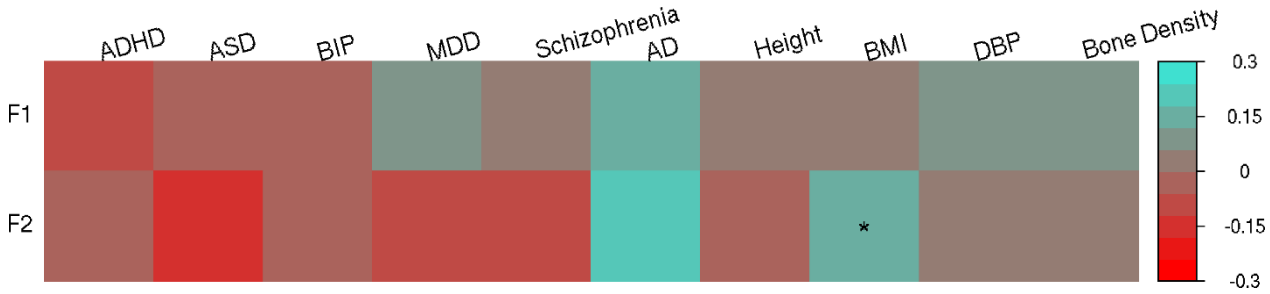


Figure 6. Genetic correlation matrix comparing the two factors of general brain function with neuropsychiatric and physical health traits.

Genetic correlations of two genetic factors (F1 and F2) with ten neuropsychiatric and physical traits: attention deficit/hyperactivity disorder (ADHD), autistic spectrum disorder (ASD), bipolar disorder (BIP), major depressive disorder (MDD), schizophrenia, Alzheimer's disease (AD), height, body-mass index (BMI), diastolic blood pressure (DBP) and bone density.

Genetic correlations at nominal and FDR-corrected significance are respectively labeled with * and **.

References

1. Greicius, M. D., Krasnow, B., Reiss, A. L. & Menon, V. Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *PNAS* **100**, 253–258 (2003).
2. Beckmann, C. F., DeLuca, M., Devlin, J. T. & Smith, S. M. Investigations into resting-state connectivity using independent component analysis. *Philos Trans R Soc Lond B Biol Sci* **360**, 1001–1013 (2005).
3. Fox, M. D. *et al.* The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *PNAS* **102**, 9673–9678 (2005).
4. Biswal, B., Yetkin, F. Z., Haughton, V. M. & Hyde, J. S. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magn Reson Med* **34**, 537–541 (1995).
5. Logothetis, N. K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**, 150–157 (2001).
6. Cortese, S., Aoki, Y. Y., Itahashi, T., Castellanos, F. X. & Eickhoff, S. B. Systematic Review and Meta-analysis: Resting-State Functional Magnetic Resonance Imaging Studies of Attention-Deficit/Hyperactivity Disorder. *J Am Acad Child Adolesc Psychiatry* **60**, 61–75 (2021).
7. Lau, W. K. W., Leung, M.-K. & Lau, B. W. M. Resting-state abnormalities in Autism Spectrum Disorders: A meta-analysis. *Sci Rep* **9**, 3892 (2019).
8. Mulders, P. C., van Eijndhoven, P. F., Schene, A. H., Beckmann, C. F. & Tendolkar, I. Resting-state functional connectivity in major depressive disorder: A review. *Neurosci Biobehav Rev* **56**, 330–344 (2015).
9. Wojtalik, J. A., Smith, M. J., Keshavan, M. S. & Eack, S. M. A Systematic and Meta-analytic Review of Neural Correlates of Functional Outcome in Schizophrenia. *Schizophr Bull* **43**, 1329–1347 (2017).
10. Badhwar, A. *et al.* Resting-state network dysfunction in Alzheimer's disease: A systematic review and meta-analysis. *Alzheimers Dement (Amst)* **8**, 73–85 (2017).

11. Miller, K. L. *et al.* Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat Neurosci* **19**, 1523–1536 (2016).
12. Glahn, D. C. *et al.* Genetic control over the resting brain. *PNAS* **107**, 1223–1228 (2010).
13. Elliott, L. T. *et al.* Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature* **562**, 210–216 (2018).
14. Ge, T., Holmes, A. J., Buckner, R. L., Smoller, J. W. & Sabuncu, M. R. Heritability analysis with repeat measurements and its application to resting-state functional connectivity. *PNAS* **114**, 5521–5526 (2017).
15. Reineberg, A. E., Hatoum, A. S., Hewitt, J. K., Banich, M. T. & Friedman, N. P. Genetic and Environmental Influence on the Human Functional Connectome. *Cereb Cortex* **30**, 2099–2113 (2020).
16. Teeuw, J. *et al.* Genetic and environmental influences on functional connectivity within and between canonical cortical resting-state networks throughout adolescent development in boys and girls. *NeuroImage* **202**, 116073 (2019).
17. Feng, J. *et al.* Partitioning heritability analyses unveil the genetic architecture of human brain multidimensional functional connectivity patterns. *Hum Brain Mapp* **41**, 3305–3317 (2020).
18. Zhang, Z. *et al.* Resting-state brain organization revealed by functional covariance networks. *PLoS One* **6**, e28817 (2011).
19. Bijsterbosch, J. *et al.* Investigations into within- and between-subject resting-state amplitude variations. *NeuroImage* **159**, 57–69 (2017).
20. Grotzinger, A. D. *et al.* Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav* **3**, 513–525 (2019).
21. Turley, P. *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat Genet* **50**, 229–237 (2018).

22. Zhu, X. *et al.* Meta-analysis of Correlated Traits via Summary Statistics from GWASs with an Application in Hypertension. *Am J Hum Genet* **96**, 21–36 (2015).
23. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* **12**, e1001779 (2015).
24. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
25. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* **8**, 1826 (2017).
26. Leeuw, C. A. de, Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput Biol* **11**, e1004219 (2015).
27. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236–1241 (2015).
28. Karlsson Linnér, R. *et al.* Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nat Genet* **51**, 245–257 (2019).
29. Levey, D. F. *et al.* GWAS of Depression Phenotypes in the Million Veteran Program and Meta-analysis in More than 1.2 Million Participants Yields 178 Independent Risk Loci. *medRxiv* 2020.05.18.20100685 (2020) doi:10.1101/2020.05.18.20100685.
30. Mesulam, M. M. From sensation to cognition. *Brain* **121**, 1013–1052 (1998).
31. Van Essen, D. C. & Glasser, M. F. In vivo architectonics: A cortico-centric perspective. *NeuroImage* **93**, 157–164 (2014).
32. Glasser, M. F., Goyal, M. S., Preuss, T. M., Raichle, M. E. & Van Essen, D. C. Trends and properties of human cerebral cortex: Correlations with cortical myelin content. *NeuroImage* **93**, 165–175 (2014).

33. Marques, J. P., Khabipova, D. & Gruetter, R. Studying cyto and myeloarchitecture of the human cortex at ultra-high field with quantitative imaging: R1, R2* and magnetic susceptibility. *NeuroImage* **147**, 152–163 (2017).

34. Hawrylycz, M. J. *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).

35. Fuhrmann, D., Knoll, L. J. & Blakemore, S.-J. Adolescence as a Sensitive Period of Brain Development. *Trends Cogn Sci* **19**, 558–566 (2015).

36. Buckner, R. L. & Krienen, F. M. The evolution of distributed association networks in the human brain. *Trends Cogn Sci* **17**, 648–665 (2013).

37. Wei, Y. *et al.* Genetic mapping and evolutionary analysis of human-expanded cognitive networks. *Nat Commun* **10**, 4839 (2019).

38. Kichaev, G. *et al.* Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *Am J Hum Genet* **104**, 65–75 (2019).

39. Lee, J. J. *et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet* **50**, 1112–1121 (2018).

40. Goes, F. S. *et al.* Genome-wide association study of schizophrenia in Ashkenazi Jews. *Am J Med Genet B Neuropsychiatr Genet* **168**, 649–659 (2015).

41. Lasky, J. L. & Wu, H. Notch Signaling, Brain Development, and Human Disease. *Pediatr Res* **57**, 104–109 (2005).

42. Burke, J. R. & Roses, A. D. Genetics of Alzheimer's disease. *Int. J. Neurol* **25–26**, 41–51 (1991).

43. Jansen, I. E. *et al.* Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet* **51**, 404–413 (2019).

44. Nazarian, A., Yashin, A. I. & Kulminski, A. M. Genome-wide analysis of genetic predisposition to Alzheimer's disease and related sex disparities. *Alzheimers Res Ther* **11**, 5 (2019).

45. Blennow, K. & Hampel, H. CSF markers for incipient Alzheimer's disease. *Lancet Neurol* **2**, 605–613 (2003).

46. Sunderland, T. *et al.* Decreased β-Amyloid1-42 and Increased Tau Levels in Cerebrospinal Fluid of Patients With Alzheimer Disease. *JAMA* **289**, 2094–2103 (2003).

47. Gormley, P. *et al.* Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. *Nat Genet* **48**, 856–866 (2016).

48. Adewuyi, E. O. *et al.* Shared Molecular Genetic Mechanisms Underlie Endometriosis and Migraine Comorbidity. *Genes* **11**, 268 (2020).

49. Greenwood, T. A. *et al.* Genome-wide Association of Endophenotypes for Schizophrenia From the Consortium on the Genetics of Schizophrenia (COGS) Study. *JAMA Psychiatry* **76**, 1274–1284 (2019).

50. Ikram, M. A. *et al.* Heritability and genome-wide associations studies of cerebral blood flow in the general population. *J Cereb Blood Flow Metab* **38**, 1598–1608 (2018).

51. Davies, G. *et al.* Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat Commun* **9**, 2098 (2018).

52. Warrier, V. *et al.* Genome-wide meta-analysis of cognitive empathy: heritability, and correlates with sex, neuropsychiatric conditions and cognition. *Mol Psychiatry* **23**, 1402–1409 (2018).

53. Mullins, N. *et al.* GWAS of Suicide Attempt in Psychiatric Disorders and Association With Major Depression Polygenic Risk Scores. *AJP* **176**, 651–660 (2019).

54. Stahl, E. A. *et al.* Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet* **51**, 793–803 (2019).

55. Ehret, G. B. *et al.* Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* **478**, 103–109 (2011).

56. German, C. A., Sinsheimer, J. S., Klimentidis, Y. C., Zhou, H. & Zhou, J. J. Ordered multinomial regression for genetic association analysis of ordinal phenotypes at Biobank scale. *Genet Epidemiol* **44**, 248–260 (2020).

57. Shin, S.-Y. *et al.* An atlas of genetic influences on human blood metabolites. *Nat Genet* **46**, 543–550 (2014).

58. Hübel, C. *et al.* Genomics of body fat percentage may contribute to sex bias in anorexia nervosa. *Am J Med Genet B Neuropsychiatr Genet* **180**, 428–438 (2019).

59. Takeuchi, F. *et al.* A Genome-Wide Association Study Confirms VKORC1, CYP2C9, and CYP4F2 as Principal Genetic Determinants of Warfarin Dose. *PLoS Genet* **5**, e1000433 (2009).

60. Pardiñas, A. F. *et al.* Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet* **50**, 381–389 (2018).

61. Alfaro-Almagro, F. *et al.* Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *NeuroImage* **166**, 400–424 (2018).

62. Alfaro-Almagro, F. *et al.* Confound modelling in UK Biobank brain imaging. *NeuroImage* **224**, 117002 (2021).

63. Smith, S. M. *et al.* Enhanced Brain Imaging Genetics in UK Biobank. *bioRxiv* 2020.07.27.223545 (2020) doi:10.1101/2020.07.27.223545.

64. Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat Genet* **51**, 63–75 (2019).

65. Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet* **51**, 431–444 (2019).

66. Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet* **50**, 668–681 (2018).

67. Pulit, S. L. *et al.* Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Hum Mol Genet* **28**, 166–174 (2019).

68. Wood, A. R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* **46**, 1173–1186 (2014).

69. Morris, J. A. *et al.* An atlas of genetic influences on osteoporosis in humans and mice. *Nat Genet* **51**, 258–266 (2019).

70. Evangelou, E. *et al.* Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat Genet* **50**, 1412–1425 (2018).