Specificity and overlap in the genetic architectures of functional and structural connectivity within cerebral resting-state networks

E.P. Tissink[1], J. Werme[1], S.C. de Lange[1,2], J.E. Savage[1], Y. Wei[1,3], C.A. de Leeuw[1], M. Nagel[1], D. Posthuma[1,4], M.P. van den Heuvel[1,4*]

[1] Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Vrije Universiteit Amsterdam, Amsterdam Neuroscience, 1081 HV Amsterdam, The Netherlands

[2] Department of Sleep and Cognition, Netherlands Institute for Neuroscience, The Netherlands

[3] School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing, 100876, China

[4] Department of Clinical Genetics, Section Complex Trait Genetics, Amsterdam Neuroscience, Vrije Universiteit Medical Center, Amsterdam UMC, Amsterdam, The Netherlands

[*] Corresponding author: martijn.vanden.heuvel@vu.nl

Abstract

The functional connectivity and dynamics of resting-state networks (RSN-FC) are vital for cognitive functioning. RSN-FC is heritable and partially translates to the anatomical architecture of white matter, but the genetic component of structural connections of RSNs (RSN-SC) and their potential genetic overlap with RSN-FC remains unknown. Here we perform genome-wide association studies ($N_{discovery}$=24,336; $N_{replication}$=3,412) and in silico annotation on RSN-SC and RSN-FC. We identify the first genes for visual network-SC, that are involved in axon guidance and synaptic functioning and show that genetic variation in RSN-FC impacts biological processes related to brain disorders that have previously been associated with FC alterations in those same RSNs. Correlations of the genetic components of RSNs are mostly observed within the functional domain, whereas less overlap is observed within the structural domain and between the functional and structural domains. This study advances the understanding of the complex functional organization of the brain and its structural underpinnings from a genetics viewpoint.

Introduction

Structural (SC) and functional connectivity (FC) are vital for healthy cognitive behaviour[1]. Brain regions that show temporally synchronized activity form functionally specialized resting-state networks (RSNs)[2], including primary networks (such as the visual or somatomotor network) and higher-order cognitive networks (such as the frontoparietal network, salience network, or default mode network)[3]. Many psychiatric and neurological disorders have been associated with disruptions within specific RSNs[4] and improving our understanding of the biological principles underlying the concept SC and FC of RSNs (RSN-SC/FC) could help elucidate the neural basis of human cognition and disorders associated with disruptions in brain connectivity.

Studies have shown that genetic factors significantly contribute to RSN-FC (H^2 = 20-40%)[5–10]. Genome-wide association studies (GWAS) on FC graph theory measures[11] and extrinsic and intrinsic functional organization[12] of RSNs have identified the first genetic variants and genes that make up this genetic component (mean h^2_{SNP} = 13.3%[11]), and show genetic overlap between FC and psychiatric disorders[13]. RSNs were traditionally discovered based on FC[2] and correlate with the structural connectivity (SC) architecture of white matter in the brain[14–16] to varying degrees across RSNs[17]. The genetic architecture of RSNs-SC has not been investigated to date, but the substantial heritability of multiple properties of major white matter tracts (mean h^2_{SNP} 25.18% - 34.9%)[18–20] suggests the importance of genetic factors for the anatomical backbone of RSNs. Describing the genetic architecture of both RSN-FC and RSN-SC as well as annotation and interpretation of the genetic signal can give insight into a biological substrate relevant to a wide variety of neurological and psychiatric disorders[21] and additionally enables us to estimate to which degree RSN-SC relates to RSN-FC based on a shared genetic source.

In this study, we aim to characterise the genetic architecture of RSNs, both structurally and functionally. Large-scale (discovery N_{FC} = 24,336 and N_{SC} = 23,985; replication N_{FC} = 3,408 and N_{SC} = 3,412) GWAS are performed on the SC and FC within seven well-known RSNs[2]. We estimate and partition the SNP-based heritability and examine the convergence of the polygenic signal from these GWAS onto genes and biological pathways, with the purpose of aiding the biological interpretation of

results and providing meaningful starting points for functional follow-up experiments[22]. We examine genetic correlations both between different RSNs, as well as across structural and functional domains. These genetic correlation analyses are extended to the locus level to facilitate the prioritisation of possible pleiotropic loci for future studies[23]. Altogether, we focus on the translation of RSN-associated genetic loci into biological interpretation and provide insights into the genetic specificity and overlap of RSN-FC and RSN-SC.

Results

GWAS of RSN-SC and RSN-FC identify six genome-wide significant loci

Following previously described procedures[24], we started our analysis by grouping cortical areas into seven RSN as defined by Yeo et al[2] (visual, somatomotor, limbic, dorsal attention, ventral attention, frontoparietal, and default-mode network; Supplementary Figure 1) and calculating the mean functional and structural connectivity within the RSNs in UK Biobank subjects (discovery N_{FC} = 24,336 and N_{SC} = 23,985; replication N_{FC} = 3,408 and N_{SC} = 3,412). RSN functional connectivity was measured as the average correlation between the activation signals of brain regions within each RSN over time, RSN structural connectivity was measured as the average fractional anisotropy (FA) of white matter tracts between brain regions within each RSN (see Methods). Discovery GWAS were performed for the FC and SC within every RSN and identified 518 genome-wide significant SNPs ($p < 5\times10^{-8}/16 = 3.13\times10^{-9}$) located in six genomic loci: three for visual network-SC, one for limbic network-FC, and a shared locus for frontoparietal network-FC and somatomotor network-FC (Supplementary Table 1). These loci seem to show RSN specific genetic effects rather than simply being driven by overall connectivity, given that none of these six loci showed a genome-wide significant association with global FC or SC.

SNP-based heritability (h^2_{SNP}) estimates for RSN-SC (M = 13.59%, SD = 1.79%) were moderately higher than those observed for RSN-FC (M = 6.71%, SD = 3.36%; Supplementary Table 2). We did not find evidence for enrichments of h^2_{SNP} in functional genomic categories after Bonferroni-

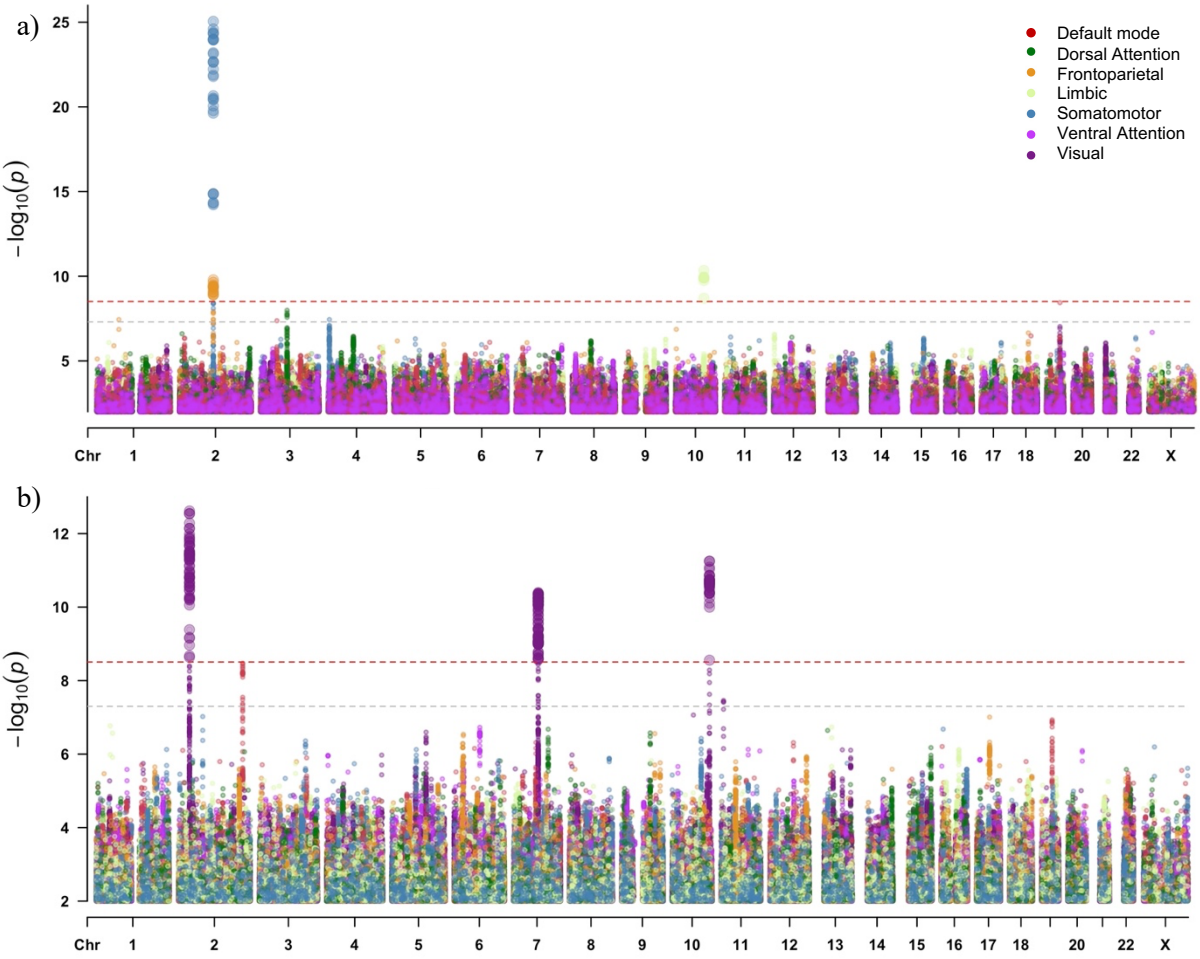


Figure 1. Multitrait Manhattan plots of SNP-based GWAS for a) RSN-FC and b) RSN-SC. The light grey dashed horizontal line indicates traditional genome-wide significance ($p < 5\times10^{-8}$), whereas the red dashed horizontal line indicates genome-wide significance after additional correction for the number of traits tested ($p < 3.13\times10^{-9}$). SNPs with $p > 0.01$ are omitted for visualisation purposes. Manhattan plots per RSN are provided as Supplementary Figure 3a (FC) and 4a (SC).

correction (Supplementary Methods 1.1 and Supplementary Table 3). The LD Score Regression (LDSC) intercept approached one for all phenotypes, indicating limited bias from population stratification. The robustness of discovery GWAS results is illustrated by polygenic score prediction and lead SNP validation (Supplementary Methods 1.3-1.4) in a replication sample (Supplementary Results 2.2-2.3).

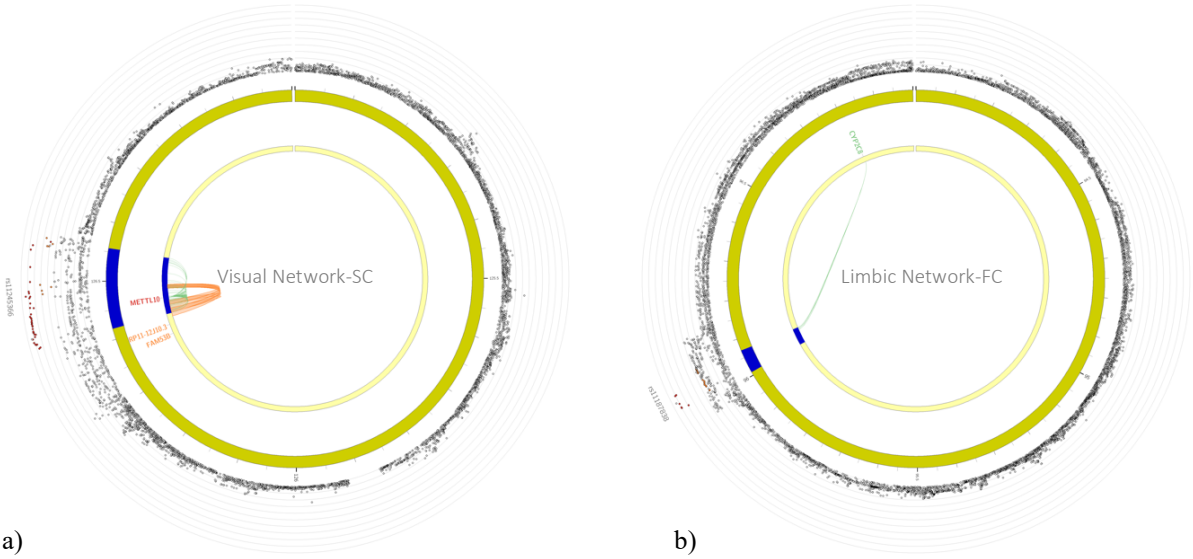
Axon guidance and synaptic functioning genes implicated in visual network-SC GWAS

We continued by examining the possible functional consequences of the SNPs involved in RSN-FC and RSN-SC. SNPs in linkage disequilibrium (LD; $r^2 \geq 0.6$) with the Bonferroni-corrected genome-

wide significant SNPs from the GWAS which also had suggestive *p*-values ($< 1\times10^{-5}$) and a minor allele frequency (MAF) > 0.005 were annotated in FUMA v1.3.7[25]. A detailed overview of the functional annotation of all candidate SNPs is displayed in Supplementary Table 4, whereas the mapped genes that resulted from positional, expression quantitative trait loci (eQTL) and chromatin interaction mapping in FUMA are listed in Supplementary Table 5.

For visual network-SC, an exonic nonsynonymous (ExNS) SNP located in exon 1 of *AC007382.1* (rs711244, $p = 1.42\times10^{-12}$, CADD = 10.39) was among the candidate SNPs in the locus on chromosome 2. The function of *AC007382.1* is unknown, but it has been associated with amygdala volume previously[26]. Within the loci on chromosome 10 and 7, exonic synonymous SNPs were found in exon 7 and exon 12 of *FAM175B* and *SEMA3A* respectively. The transcript of *FAM175B* is a component of the BRISC enzyme complex that deubiquitinates Lys-63 linked chains in order to control protein function[27]. Experimental studies have suggested that such deubiquitination can regulate synaptic transmission and synaptic plasticity[28]. *SEMA3A* contained multiple intronic SNPs associated with visual network-SC with high CADD scores (11 SNPs with CADD > 12.37), which are usually considered reducing organismal fitness and correlating with molecular functionality and pathogenicity[29]. The product of *SEMA3A* is known as a key regulator of axon outgrowth during the establishment of correct pathways in the developing nervous system[30].

We additionally mapped 46 visual network-SC candidate SNPs to *METTL10,* because of their established eQTL associations in fetal and adult cerebral cortex tissue as well as through chromatin interaction mapping. *METTL10* encodes a methyltransferase that catalyses the trimethylation of eEF1A at Lys-318 – a key regulator of ribosomal translation[31]. Visual network-SC SNPs were also mapped to the *METTL10-FAM53B* readthrough (*RP11-12J10.3*) and *FAM53B* gene, because of known chromatin interaction in fetal and adult cerebral cortex tissue (Figure 2a). FAM53B is required for Wnt signaling, a pathway important for cell regeneration[32]. Lastly, positional mapping of candidate SNPs within a 10kb window of a gene resulted in the identification of *VIT, STRN, and HEATR5B* genes for visual network-SC (Supplementary Table 5).



a) b)

Figure 2. a) Visual network-SC SNPs were mapped to *METTL10*, *FAM53B* and *METTL10-FAM53B* readthrough (*RP11-12J10.3*) through chromatin interaction mapping (orange). *METTL10* was additionally mapped by 46 SNPs because of their eQTL associations in cerebral cortex tissue. b) FUMA gene mapping, based on established eQTL associations (green) in human temporal cortex, link eight limbic network-FC SNPs on chromosome 10 to *CYP2C8.*

Annotation of specific and shared loci across RSN-FC

We observed two ExNS SNPs in exon 19 (rs2274224, $p = 1.771\times10^{-10}$) and 25 (rs2274223, $p = 1.22\times10^{-5}$) of the *PLCE1* gene to be associated with limbic network-FC. The *PLCE1* gene encodes for the phospholipase C ϵ1, which mediates the production of two second messengers that regulate cell growth, differentiation, and gene expression[33]. The high CADD scores (17.35 and 17.48 respectively) suggest deleteriousness of these two ExNS SNPs. Additionally, four intergenic SNPs within the same locus were located near the *NOC3L* gene.

On chromosome 10, eight SNPs associated with limbic network-FC were eQTLs for the *CYP2C8* gene (Figure 2b). Expression of *CYP2C8* results in an enzyme important for drug metabolism[34]. One of CYP2C8 substrates, the non-selective monoamine oxidase inhibitor phenelzine, is known to target the nervous system and is clinically prescribed as treatment for major depressive disorder[35]. A large body of research has verified the association between major depressive disorder and

changes in limbic network functional connectivity, as well as with other RSNs (see Kaiser et al[36] for a meta-analysis).

The annotation of SNPs in the locus that was shared between frontoparietal and somatomotor network-FC revealed only intergenic candidate SNPs (enrichment = 2.15, $p = 5.09\times10^{-9}$), which convolutes biological interpretation but is a common observation for complex traits[37]. The nearest genes to the candidate SNPs in this locus were *PAX8* and *IGKV1OR2-108* (respectively 29 and 53 kb distance). *PAX8* encodes a transcription factor that is considered to regulate the expression of genes important for thyroid development[38] and the production of thyroid hormone[39]. FC within both the somatomotor and frontoparietal network is reduced in individuals with subclinical[40] and clinical hypothyroidism[41].

Default mode network-FC genes associated with Alzheimer's disease

We next performed gene-based GWAS for the FC and SC within every RSN using MAGMA (Supplementary Table 5). We detected two Bonferroni-corrected genome-wide significant genes additional to the FUMA mapped genes by combining information from neighbouring variants within a single gene in MAGMA (Figure 3, Supplementary Table 6). Visual network-FC was associated with *APOC1* ($z = 5.15$, $p = 1.31\times10^{-7}$), and for default mode network-FC *APOE* was found to be associated ($z = 5.13$, $p = 1.43\times10^{-7}$). *APOC1* and *APOE* are both located within the 19q13.2 locus and are well-known risk factors for Alzheimer's disease[42]. Additionally, gene-set analysis results are provided in Supplementary Methods 1.2 and Supplementary Results 2.1.

In order to determine whether there is genetic overlap between Alzheimer's disease[43] and default mode network-FC, we performed local genetic correlation (r_g) analysis using LAVA (see Methods; Supplementary Table 7). For default mode network-FC, we detected two loci on chromosome 12 (BP 64,403,858-66,114,643) and 19 (BP 45,040,933-45,893,307) which showed significant local r_g at $p < (0.05/71=) 7.04\times10^{-4}$ with Alzheimer's disease (Supplementary Figure 5). Given the negligible heritability of global FC in these loci (univariate $p = 0.27$ and $p = 0.01$ respectively, whereas $p = 1.30\times10^{-5}$ and $p = 1.62\times10^{-8}$ for default mode network-FC) we conclude that these local genetic associations with Alzheimer's disease are not driven by total brain connectivity. The locus on

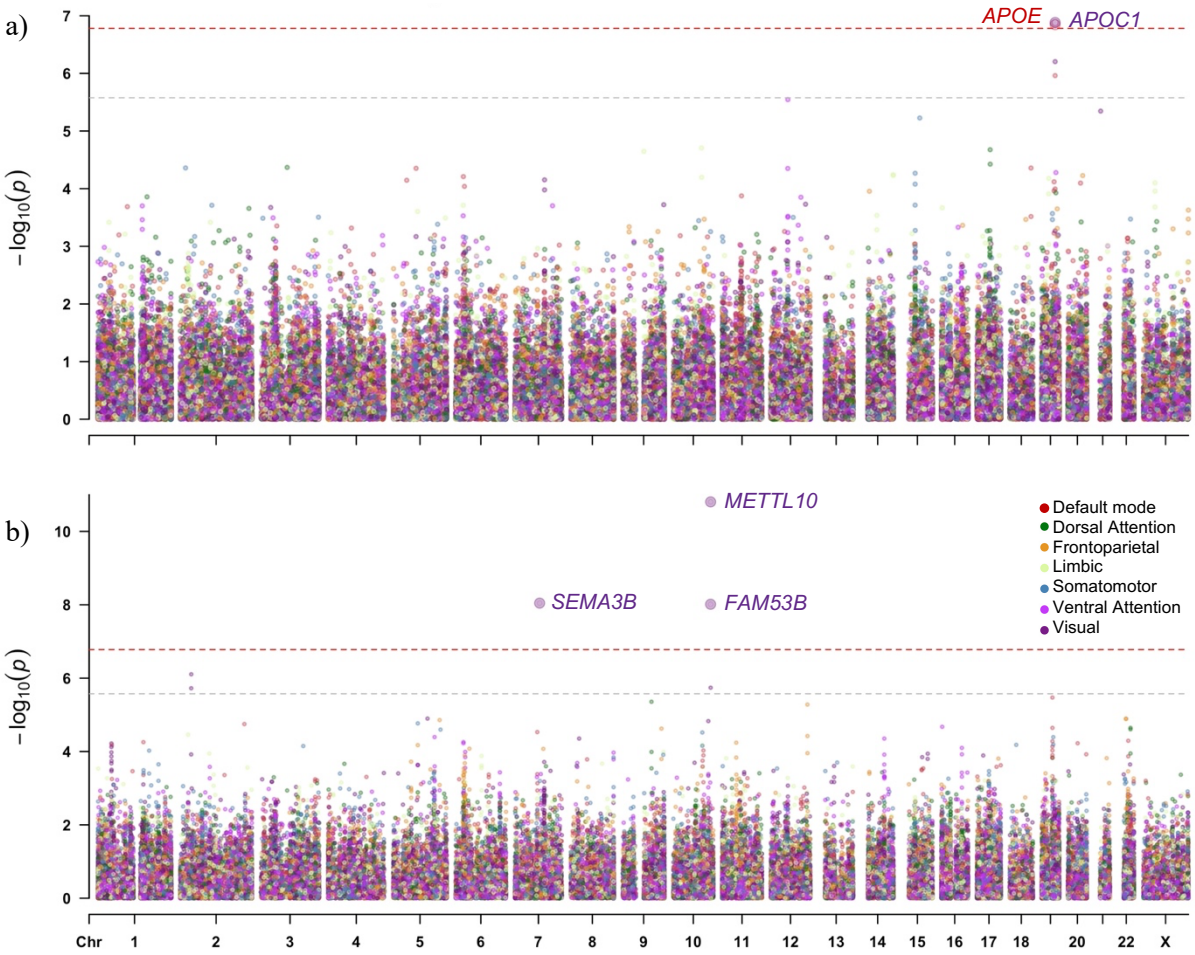


Figure 3. Multitrait Manhattan plots of gene-based GWAS for a) FC and b) SC within RSNs. The light grey dashed horizontal line indicates significance after correcting for the number of genes tested per trait ($p < 2.65\times10^{-6}$), whereas the red dashed horizontal line indicates significance after an additional correction for the number of traits tested ($p < 1.66\times10^{-7}$). Manhattan plots per RSN are provided as Supplementary Figure 3b (FC) and 4b (SC).

chromosome 12 showed a positive r_g (ρ) between Alzheimer's disease and default mode network-FC (BP 64,403,858-66,114,643, ρ = 0.69, 95% CI = 0.35 – 1.00, $p = 3.25\times10^{-4}$). Interestingly, this locus has been identified in a previous GWAS for hippocampal atrophy, a biological marker of Alzheimer's disease[44]. Negative r_g between Alzheimer's disease and FC within DMN was observed in the locus on chromosome 19 (BP 45,040,933-45,893,307, ρ = -0.56, 95% CI = -0.82 – -0.38, $p = 9.23\times10^{-9}$), indicating that lower default mode network-FC was associated with higher genetic risk of Alzheimer's disease. Note that this larger defined locus showed weak heritability ($p = 0.014$) for visual network-FC despite the significance of *APOC1* in the gene-based GWAS, which would make genetic correlation

estimates with Alzheimer's disease unreliable and uninterpretable[23]. Therefore, Alzheimer's disease seems to show genetic overlap specifically with default mode network-FC.

Examining overlap between structure and function per RSN through genetic correlations

As SC strength has been noted to correlate with FC strength on the phenotypic level[16], we sought to investigate the correlations between FC and SC within each RSN on a genetic level. Genome-wide genetic correlations (r_g) were estimated in LDSC using SNP-based summary statistics (Figure 4). We observed no nominally significant genome-wide r_g's between SC and FC in any of the RSNs (Supplementary Table 8). Genome-wide r_g estimates ranged from -0.19 (SE = 0.15, p = 0.19) in the dorsal attention network (DAN) and 0.23 (SE = 0.23, p = 0.30) in the frontoparietal network (FPN).

Strongly localized or opposing local r_g's possibly may go undetected, since genome-wide r_g's are an average of the shared genetic association signal across the genome. We examined whether such relationships between SC and FC within any given RSN exist by performing local r_g analysis using LAVA[23], though we did not identify any significant r_g on a locus level either (Supplementary Table 9).

Genome-wide and local genetic correlations within the functional and structural domain

We examined the shared genetic signal across RSNs within the same domain by conducting genome-wide r_g analyses using LDSC (Figure 5; Supplementary Table 8). For functional connectivity, a positive Bonferroni significant genome-wide r_g was observed between the default mode and ventral attention

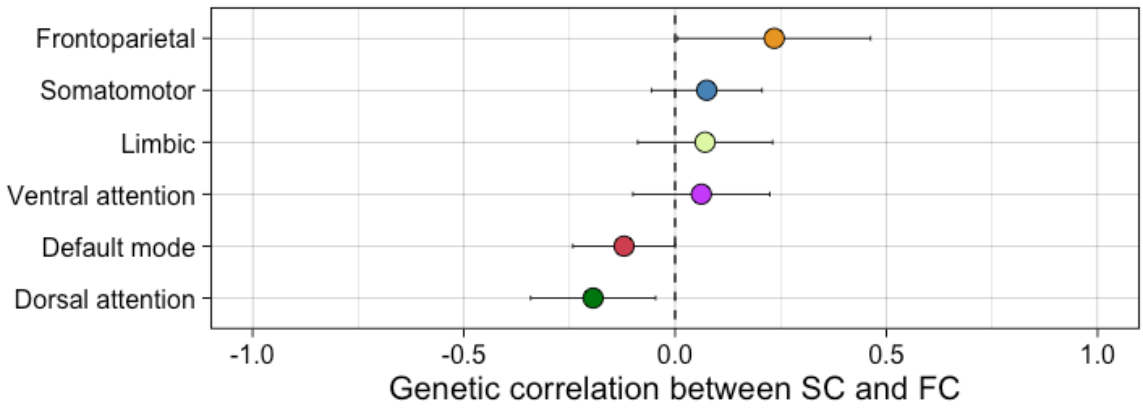


Figure 4. Global r_g (±SE) between FC and SC within the same RSN as performed in LDSC do not show estimates significantly different from zero (Supplementary Table 8). Additional estimation of local r_g did not yield significant overlapping loci between SC and FC within each RSN either (Supplementary Table 9).

network (r_g = 0.52, SE = 0.16, p = 1.00×10^{-3}). This association was not driven by global FC as neither default mode nor ventral attention network-FC were genetically correlated with global FC (r_g = 0.19, SE = 0.18, p = 0.29; r_g = 0.26, SE = 0.19, p = 0.18 respectively). Note that this positive r_g does not imply simultaneous functional activation of these two RSNs or their involvement in similar cognitive tasks (which would contradict previous research[45]), but suggests that variants that influence default mode network-FC generally tend to influence ventral attention network-FC in the same direction.

For structural connectivity, we observed multiple significant genome-wide r_g's ($p < 1.19\times10^{-3}$) between RSNs, though many of these were also correlated with global SC (Supplementary Table 6). To determine whether the correlations between the structural RSNs could be accounted for by global SC, we used genomic SEM to compute residual r_g estimates between the structural RSNs while taking global SC into account (see Methods). As none of the residual r_g estimates remained significant, we conclude that global SC likely accounts for the observed relations between the RSN-SC.

We extended our investigation into shared genetic signal across RSNs beyond the global to the local scale. Eighteen loci showed Bonferroni corrected significant r_g's when comparing RSNs within the functional domain (Table 1). These were all highly positive (mean ρ = 0.84, SD = 0.09) and were not confounded by global FC. When comparing RSNs within the structural domain, local r_g analysis

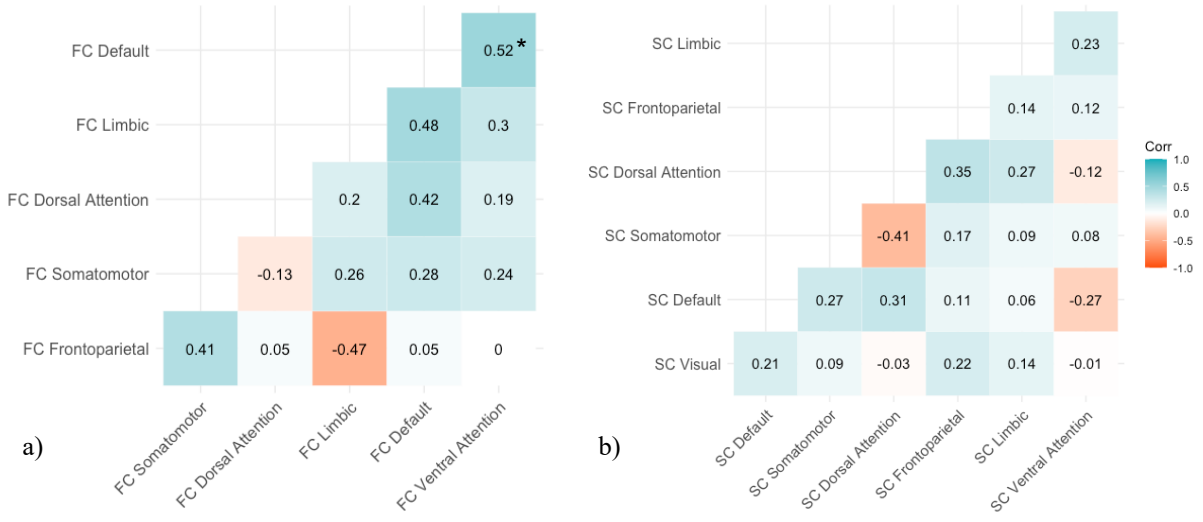


Figure 5. Genome-wide r_g between (a) RSN-FC and (b) RSN-SC. If one of the two RSNs showing significant LDSC r_g showed additional significant r_g with global FC/SC, we instead report the residual r_g (r_g between the two RSNs while taking global FC/SC into account in Genomic SEM; see Methods and Figure 6). The significant r_g that survived correction for multiple testing ($p < 1.19\times10^{-3}$) is indicated with an asterisk (*).

with LAVA revealed only one positively correlated locus between SC within DAN and FPN (15:39238841:40604780, local r_g (ρ) = 0.85, $p = 9.51\times10^{-7}$; Table 1). A complete overview of LAVA local r_g results can be found in Supplementary Table 9.

Table 1.

Loci with Bonferroni-corrected significant ($p < (0.05/774=) 6.46\times10^{-5}$) r_g (ρ with lower and upper limit of 95% confidence interval) between RSN-FC or RSN-SC as performed in LAVA. Within these loci, global FC or SC did not show significant univariate h^2 or r_g with either of the two RSNs. See Supplementary Table 9 for all local r_g summary statistics. SMN = somatomotor network, VN = visual network, DMN = default mode network, FPN = frontoparietal network, VAN = ventral attention network, DAN = dorsal attention network, LN = limbic network.

Chr	Start	Stop	RSN 1	RSN 2	ρ	95% CI		p-value
1	2,215,496	2,983,519	FC SMN	FC VN	0.77	0.47	1.00	3.39×10^{-5}
1	18,427,821	19,238,924	FC DMN	FC FPN	0.72	0.45	1.00	9.48×10^{-6}
1	211,082,893	212,3475,82	FC VAN	FC SMN	1.00	0.74	1.00	1.75×10^{-7}
2	113,930,669	115,203,835	FC FPN	FC SMN	0.88	0.64	1.00	3.42×10^{-7}
2	207,726,595	208,674,588	FC FPN	FC VN	0.97	0.72	1.00	1.02×10^{-6}
5	4,636,543	5,828,694	FC DMN	FC DAN	0.73	0.47	1.00	2.72×10^{-5}
5	68,006,994	71,468,651	FC VAN	FC SMN	0.79	0.53	1.00	2.05×10^{-5}
5	75,959,516	77,290,255	FC DMN	FC DAN	0.91	0.65	1.00	3.42×10^{-6}
6	10,416,551	11,790,671	FC VAN	FC SMN	0.83	0.55	1.00	2.49×10^{-5}
7	50,894,509	51,951,647	FC LN	FC VN	0.88	0.57	1.00	5.12×10^{-6}
8	64,215,359	66,018,204	FC DMN	FC VN	0.86	0.59	1.00	1.09×10^{-5}
9	93,441,051	94,175,374	FC FPN	FC SMN	0.90	0.61	1.00	1.73×10^{-5}
9	93,441,051	94,175,374	FC FPN	FC VN	0.87	0.62	1.00	4.58×10^{-6}
10	89,971,629	91,021,321	FC VAN	FC VN	0.96	0.67	1.00	1.23×10^{-6}
15	39,238,841	40,604,780	SC DAN	SC FPN	0.85	0.53	1.00	9.51×10^{-7}
17	13,648,447	14,508,610	FC DMN	FC LN	0.89	0.69	1.00	3.50×10^{-9}
18	2,839,843	3,722,828	FC DMN	FC DAN	0.70	0.45	1.00	2.66×10^{-5}
19	17,045,964	17,750,518	FC LN	FC DAN	0.73	0.47	1.00	2.34×10^{-6}
19	17,045,964	17,750,518	FC DMN	FC DAN	0.79	0.53	1.00	1.43×10^{-5}

Discussion

Mapping the genetic components of resting state networks (RSNs) may provide insight into the aetiology of brain function and brain disorders. RSNs are typically defined using functional connectivity (FC), and structural connectivity (SC) correlates to FC in varying degrees across RSNs[17]. The genetic component of RSN-SC has been less studied and as one of the fundamental goals in

246 neuroscience is to understand the relationship between structure and function within the brain, the aim
247 of this study was to gain more insight into the genetic underpinnings of structural and functional
248 connectivity (SC; FC) within a framework that respects the brain's hierarchical functional architecture.
249 With the use of GWAS and in silico annotation we identify the first genes for visual network-SC, that
250 are involved in axon guidance and synaptic functioning. We further observe that genetic variation in
251 RSN-FC (e.g. limbic network-FC and default mode network-FC) impacts biological processes related
252 to brain disorders (major depressive disorder and Alzheimer's disease respectively) that have previously
253 been associated with FC alterations in those same RSN. The genetic component of RSNs overlaps
254 mostly within the functional domain, whereas less overlap is observed within the structural domain and
255 between the functional and structural domains.

256 For FC within RSNs (RSN-FC), we detect biologically interpretable results that are specific to
257 default mode and limbic network-FC. For default mode network-FC, we observe *APOE* as a genome-
258 wide significant gene. The default mode network is hypothesized to relate to Alzheimer's disease
259 through the role of default model network-FC in memory consolidation[46] and through the spreading of
260 cortical atrophy over time, which follows the pattern of default mode network regions[47]. Here, we
261 complement earlier phenotypic observations that link Alzheimer's disease to default mode network-
262 FC[48] by now also showing genetic correlations in two loci between Alzheimer's disease and default
263 mode network-FC. Functional follow up would be necessary to investigate how the variants and genes
264 in these loci affect default mode network-FC. The limbic network is commonly known for its
265 involvement in emotion regulation, episodic memory, and action–outcome learning[49] and has been
266 associated with mood disorders, such as major depression disorder and bipolar depression[50]. The genes
267 *PLCE1*, *NOC3L* and *CYP2C8* were related to limbic network-FC, all of which have been noted to have
268 a relationship with major depressive disorder[35,51,52]. A previous study investigating the role of *PLCE1*
269 in major depressive disorder patients has demonstrated an association with antidepressant remission in
270 female patients, together with other genes within the calcium/calmodulin-dependent protein kinase
271 (CaMK) pathway[51]. *NOC3L* eQTLs in the cerebellum and nucleus accumbens have previously been
272 demonstrated to associate with depression severity and antidepressant response[52], and one of the
273 substrates of CYP2C8 is clinically prescribed as treatment for major depressive disorder (phenelzine)[35].

These results seem to suggest that major depressive disorder and antidepressant response involve processes that are impacted by genetic variation in limbic network-FC.

In addition to RSN-FC specific effects, we find evidence of shared genetic signal in FC across different RSNs using several approaches. Specifically, we observe a genetically correlated and common genome-wide significant locus for both somatomotor and frontoparietal network-FC near *PAX8*. *PAX8* regulates multiple genes involved in the production of thyroid hormone[39], an interesting result considering that both somatomotor and frontoparietal network-FC have been linked to (subclinical) hypothyroidism[40,41]. Additionally, we detect genetically correlating loci between all RSN-FC and a genome-wide genetic correlation between ventral attention and default mode network-FC. The ventral attention network supports salience processing[53], whereas the default mode network includes areas widespread over the brain and supports emotional processing, self-referential mental activity, and recollection of prior experiences[54]. Increased FC within these two RSNs has been associated with bulimia nervosa[55] and contributes to episodic memory retrieval[53]. Altogether, the shared genetic underpinnings of different RSN-FC that we present here could give a possible explanation how multiple disorders are associated with more than one RSN.

We report considerable heritability estimates for RSN-SC (ranging from 10.00% to 15.40%) and identify nine genes that suggest a role for synaptic transmission in the genetics of visual network-SC. For example, *STRN* encodes for a calmodulin-binding protein that is mostly found in dendritic spines playing a role in Ca2+-signaling[56], the transcript of *FAM175B* is a component of a deubiquitylation enzyme complex that has been suggested play a role in synaptic transmission and synaptic plasticity[28], and *SEMA3A* is known as an axonal guidance gene during development[30]. The SEMA3A protein has been shown to be upregulated in schizophrenia patients and is suggested to contribute to the developmentally induced impairment of synaptic connectivity in the disorder[57]. Visual network functional hyperconnectivity has been observed in schizophrenia[58,59] and related to visual hallucinations[59], but future studies should investigate the equivalent SC component in more detail given our findings.

When investigating the genetic relationship between SC and FC within each RSN, we find no significant genome-wide or local genetic correlations. Since the estimation of genetic correlations is

dependent on sample size and the heritability estimates of both traits[60], studies with increased power are needed to examine the robustness of these results. Future studies could additionally incorporate recent insights that indirect structural connections supporting direct functionally connected regions complicate simple structure to function mapping[61]. Our study focussed on direct structural connections within RSNs. The possibility that the genetics of RSN-FC overlap with that of indirect pathways that structurally connect brain regions within RSNs via a route beyond the borders of that RSN could therefore be subject to future research.

Several limitations must be considered while interpreting our results. It is known that rsfMRI measures can be noisy and subject to motion distortion, which raises the possibility of differences in measurement error between RSN-FC and RSN-SC. However, given our stringent pre-processing and quality control to enable noise minimization and additional use of rsfMRI-specific covariates in GWAS, we were able to find heritability estimates for RSN-FC that are concordant with previous studies[13]. Second, even though UK Biobank provides genetic and uniform MRI data at unprecedented sample sizes, it is evident that even larger sample sizes are needed for discovering the often small genetic effects of polygenic traits[62]. The null results observed for some RSN-FC/SC GWAS, partitioned heritability and gene-set analyses might be explained by the multiple comparison correction for the number of phenotypes analysed, in conjunction with insufficient statistical power. Third, some other sample characteristics, such as the European ancestry, age-class and socioeconomic status of subjects, may limit the generalizability of our findings. While we corrected for age and Townsend deprivation index (a proxy of socio-economic status) in our GWAS to reduce this bias, larger and more diverse imaging-genetics datasets are undoubtedly needed.

This study examines the specificity and overlap in genetic architecture of RSNs – structurally and functionally. We observe several genetic effects that seem to be specific to certain RSNs and highlight relevant biological processes for brain connectivity and related brain disorders. The complexity of structure-function coupling within RSNs is illustrated by the observation that, despite genetic overlap of RSNs within the functional domain, genetic overlap is less apparent within the structural domain and between the functional and structural domains. Altogether, this study advances

329 the understanding of the complex functional organisation of the brain and its structural underpinnings
330 from a genetics viewpoint.
331
332 ## Methods
333 A flowchart that describes all Methods used in this manuscript is displayed in Supplementary Figure 1.
334
335 ### Sample
336 The UK Biobank (UKB) is a resource with genomic and imaging data of volunteer participants[63]. The
337 National Research Ethics Service Committee North West–Haydock ethically approved this initiative
338 (reference 11/NW/0382) and data were accessed under application #16406. Combined SNP-genotypes
339 and neuroimaging data of N = 40,682 participants have been available since January 2020. From all
340 new subjects ID's in the latest neuroimaging release (January 2020), we randomly assigned 5,000
341 subjects to a holdout set for validation. Subsetting the total sample to subjects with all neuroimaging
342 data necessary to construct our phenotypes as described below, resulted in N_{FC} = 37,017 and N_{SC} =
343 36,645. We only included subjects for which the projected ancestry principal component score was
344 closest to and < 6 SD from the average principal component score of the European 1000 Genomes
345 sample based on Mahalanobis distance. This procedure has been described in previous publications by
346 our group[64] and the number of non-European exclusions are displayed in Supplementary Table 10.
347 Other exclusion criteria were withdrawn consent, UKB-provided relatedness, discordant sex or sex
348 aneuploidy (Supplementary Table 10). Further quality control on genomic and neuroimaging data is
349 described below and resulted in the sample sizes and sample characteristics as displayed in
350 Supplementary Table 11.
351
352 ### Genotype data
353 The genotype data used in this study were obtained from the UK BiobankTM Axiom and the UK
354 BiLEVE Axiom arrays. These Affymetrix arrays cover 812,428 unique genetic markers and overlap
355 95% in SNP content. This number of SNPs was increased to 92,693,895 by imputation carried out by
356 UKB. Variants were imputed using the Haplotype Reference Consortium and the UK10K haplotype

panel as reference. We applied our in-house quality control pipeline in addition to quality control performed by UKB. This procedure excluded SNPs with low imputation scores (INFO<0.9), low minor allele frequency (MAF<0.005) or high missingness (>0.05), multiallelic SNPs, indels, and SNPs without unique rs-identifiers. A total of 9,380,668 SNPs passed quality control and were converted to hard call SNPs using a certainty threshold of 0.9 for further analyses.

Neuroimaging data

Pre-processing & connectome reconstruction

The UKB scanning protocol and processing pipeline is described in the UKB Brain Imaging Documentation[65]. For this study, we made use of the available resting-state functional brain images (rsfMRI) and multiband diffusion brain images (DWI) together with T1 surface model files and structural segmentation from FreeSurfer[66]. These three types of data were used as input for the structural and functional pipeline of CATO (Connectivity Analysis TOolbox)[67]. Prior to this, UKB performed pre-processing on DWI and rsfMRI data as described in the UKB Brain Imaging Documentation[65].

In CATO's structural pipeline, additional pre-processing of DWI files was performed in FSL[68] by computing a DWI reference image based on the corrected diffusion-unweighted (b0) volumes, computing the registration matrix between DWI reference image and the anatomical T1 image, and registering the Freesurfer segmentation to the DWI reference image. The surface was parcellated based on the Cammoun sub-parcellations of the Desikan-Killiany atlas including 250 cortical regions[69]. We reconstructed the diffusion signal with diffusion tensor imaging (DTI), a deterministic method that is robust and relatively simple compared with more advanced diffusion reconstruction methods[67]. In CATO, the Fiber Assignment by Continuous Tracking (FACT) algorithm[70] is used to reconstruct fibers and fractional anisotropy (FA) was used as weights of reconstructed fibers. FA is a robust measure of white matter integrity and has been found to be sensitive to changes in connectivity[18] and correlates with axon density, size and myelination[71]. The structural connectivity matrix was built out of all fiber segments that connected two regions in the atlas. Additional filters were applied, namely a minimal FA of 0.1, minimal length of 30 mm and having 2 or more number of streamlines.

The functional pipeline in CATO consisted of similar steps. First, we computed a rsfMRI reference image by averaging all rsfMRI frames in FSL and subsequently registered this reference image and the T1 image in FreeSurfer. Second, we parcellated the surface based on the same atlas as in the structural pipeline (to enable structure-function comparison in downstream analyses) and we registered the T1 parcellation to the rsfMRI image. Third, motion metrics were estimated, and time-series were corrected for covariates (linear trends and first order drifts of motion parameters and the mean signal intensity of voxels in white matter and cerebrospinal fluid and of all voxels in the brain) by regression. Fourth, time-series were passed through band-pass filtering (frequencies 0.01 to 0.1) and scrubbing (max FD = 0.25, max DVARS = 1.5, min violations = 2, backward neighbours = 1, forward neighbours = 0). Fifth, the functional connectivity matrix was computed by the Pearson's correlation coefficient of the average signal intensity of every pair of brain regions across the frames that survived filtering.

Quality control

The UKB scanning and pre-processing protocol includes filters for outliers based on manual QC and an advanced classifier described elsewhere[72]. We excluded a small number of subjects that UKB identified as outliers and placed in an "unusable" folder. The UKB main documentation[65] suggests a second set of UKB data fields that can be used as outlier criteria. Outlier subjects are defined as subjects that score for any of the values > 3 interquartile ranges above the upper quartile or below the lower quartile. Outlier criteria included measures that describe the discrepancy between the T1-weighted, rsfMRI and DWI images and the population average template after LINEAR and NON-LINEAR alignment, the amount of nonlinear warping necessary to map a subject to the standard template, the signal to noise ratio in rsfMRI, the mean rfMRI head motion averaged across space and time points and the total number of outlier slices in DWI volumes. We extended this recommended list with connectome specific measures, including the average prevalence of all connections present and absent in the reconstructed brain network of a subject (low average prevalence scores indicate the presence of odd connections and high values indicate the absence of common connections), the sum of number-of-

411 streamlines and average FA of all connections in the reconstructed brain network of a subject. The
412 number of exclusions can be viewed in Supplementary Table 10.
413
414 *Phenotype reconstruction*
415 In this study, the phenotypes of interest were the functional and structural connectivity (FC;SC) within
416 seven resting-state networks (RSNs) that previously have been identified[2] and are commonly used in
417 (clinical) neuroimaging studies: the default mode network, ventral attention network, dorsal attention
418 network, visual network, limbic network, somatomotor network and frontoparietal network. Each of
419 the 250 cortical regions of the reconstructed structural and functional connectomes were assigned the
420 ratio to what extent they belonged to each of these seven RSNs, using a mask created and validated
421 elsewhere (see Supplementary Information of Wei *et al*[24]). Each connection was then weighted by
422 multiplying the ratios of the two regions involved in the particular RSN. FC and SC within the RSNs
423 were respectively calculated as the mean correlation and mean fractional anisotropy of the connections
424 within the RSN. We also computed two global FC and SC phenotypes as the mean correlation and mean
425 fractional anisotropy of all available connections, to be able to correct for connectivity that is non-
426 specific to RSNs in downstream analyses.
427
428 **Statistical analyses**
429 *SNP-based GWAS*
430 To identify common genetic variants involved in FC within each of the seven RSN, we performed seven
431 SNP-based GWAS in PLINK2[73]. Also, for the SC within each of the seven RSN, a SNP-based GWAS
432 was performed. It is common practice to include a global FC or SC estimate as covariate in GWAS to
433 capture associations that are driven by the level of connectivity within an RSN irrespective of the level
434 of connectivity throughout the whole brain. It has become apparent that this risks the introduction of
435 collider bias (inducing false-positives)[74]. Here we build upon recent developments in statistical genetics
436 that have provided multiple methods that allow for post-GWAS analyses conditional on global
437 connectivity. Therefore, we used the global FC and global SC phenotypes to run two additional SNP-
438 based GWAS, for which the summary statistics were used for conditional downstream analyses. The

total amount of GWAS was therefore sixteen. In order to correct for population stratification during GWAS, a principal component analysis was performed in FlashPCA2[75] using only independent ($r^2 <$ 0.1), common (MAF > 0.01) and genotyped SNPs or SNPs with very high imputation quality (INFO=1). The first 30 principal components were used as covariates in all GWAS, together with sex, age, genotype array, Townsend deprivation index (a proxy of socio-economic status), general neuroimaging confounders as well as FC/SC specific covariates (recommended by Alfaro-Almagro and colleagues[76]). The general set included handedness, scanning site, the use of T2 FLAIR in Freesurfer processing, intensity scaling of T1, intensity scaling of T2 FLAIR, scanner lateral (X), transverse (Y) and longitudinal (Z) brain position, and Z-coordinate of the coil within the scanner. FC-specific and SC-specific covariates were respectively intensity scaling and echo time of rsfMRI, and intensity scaling of DWI. For reasons of collinearity, we ran principal component analysis on all covariates (excluding the population stratification principal components) and retained those principal components that explained > 99% of variance. Rare variants (MAF < 0.005) and SNPs with high missingness (>5%) were excluded from GWAS and male X variants were counted as 0/1. The genome-wide significance threshold was $\alpha = (0.05/1{,}000{,}000/16 =) 3.13\times10^{-9}$ according to the Bonferroni correction for multiple testing.

SNP-based heritability

SNP-based (h^2_{SNP}; or narrow-sense) heritability represents the proportion of phenotypic variance that can be explained by common additive variation. In contrast, broad-sense heritability captures the total genetic contribution to the phenotype and is often based on family studies[77]. We applied Linkage Disequilibrium Score regression (LDSC) on the SNP-based GWAS summary statistics of all sixteen phenotypes to estimate h^2_{SNP} using precomputed LD scores from 1000 Genomes EUR, as provided by the LDSC developers.

Functional annotation

FUMA is a web-based platform that can be used to functionally map and annotate SNPs that appear significant during GWAS. We uploaded summary statistics to FUMA if GWAS identified at least one

genome-wide significant SNP. Candidate SNPs were defined as all SNPs in LD r^2>0.6 with an independent genome-wide significant SNPs (r^2<0.6). Annotation was subsequently performed using ANNOVAR[78], RegulomeDB[79] score and ChromHMM[80]. Lead SNPs were defined as independent SNPs r^2<0.1. Genomic loci were constructed by taking all independent significant SNPs $r^2 < 0.1$ with LD blocks within 250 kb distance and independent significant SNPs $r^2 \geq 0.1$. Within every locus, SNPs were mapped to genes using three methods: positional mapping, eQTLs mapping or chromatin interaction mapping. SNPs were positionally mapped to genes if their physical distance was <10 kb. Mapping based on eQTLs relied on known associations between SNPs and the gene-expression of genes within a 1Mb window, from BRAINEAC[81] (frontal, occipital, temporal, cerebral cortex), GTEx v8[82] cerebral cortex and xQTLServer[83] dorsolateral prefrontal cortex. Chromatin interaction mapping was based on established 3D DNA-DNA interactions between SNP and gene regions from Hi-C databases in cortex tissue (PsychENCODE[84], Giusti-Rodriguez *et al*[85], and GSE87112[86]). To restrict chromatin interaction mapping to plausible biological interactions, we only included interactions where one region overlapped with an enhancer (as predicted by the Roadmap Epigenomics project[87] in cortex tissue) and the other region overlapped with a promoter (250 bp upstream to 500 bp downstream of the transcription start site as well as predicted by the Roadmap Epigenomics project in cortex tissue). A FDR threshold of 1×10^{-5} was used, as recommended in previous literature[86].

Gene-based GWAS

Performing GWAS on the level of genes has been suggested to be more powerful than GWAS on the level of SNPs[88]. Therefore, the sixteen SNP-based GWAS summary statistics were used to perform sixteen gene-based GWAS in MAGMA (Multi-marker Analysis of GenoMic Annotation) v1.08[88]. A mean SNP-wise model was applied (with the UKB European population serving as an ancestry reference group) to test the joint association of all SNPs within 18,850 genes with RSN-FC/RSN-SC. The genome-wide significance threshold was adjusted for multiple testing to $\alpha = (0.05/18{,}850)/16 = 1.66\times10^{-7}$.

Genome-wide genetic correlations

To assess the overlap in genetic architecture between FC/SC within RSNs while taking the influence of global FC/SC into account, we designed a genetic correlation (r_g) analysis pipeline. This pipeline consisted of three steps. 1) In the first step, genome-wide r_g between 42 combinations of RSNs were estimated using LDSC ($\alpha = (0.05/42 =) 1.19 \times 10^{-3}$). The summary statistics of SNP-based GWAS were used as input for LDSC. We excluded FC-VN, because both the lambda (<1.02) and ratio (>0.20) values were out of bound for LDSC. 2) For all RSNs included in a significant bivariate r_g, additional r_g with global FC/SC were calculated in LDSC. 3) If one or both RSNs from the significant bivariate r_g showed additional significant r_g with global FC/SC, we recalculated of the genome-wide r_g between the two RSNs with global FC/SC taken into account. Since such residual genome-wide r_g analyses are not implemented in LDSC, we applied Genomic Structural Equation Modelling (genomic SEM)[89]. Genomic SEM is a method that enables to model the multivariate genetic architecture and covariance structure of complex traits using GWAS summary statistics and allows for sample overlap. We modelled residual covariance between RSN as the covariance between the residual variance of the two RSNs involved after taking the global factor into account (Figure 6). A confirmatory factor analysis was then ran using Diagonally Weighted Least Square estimation.

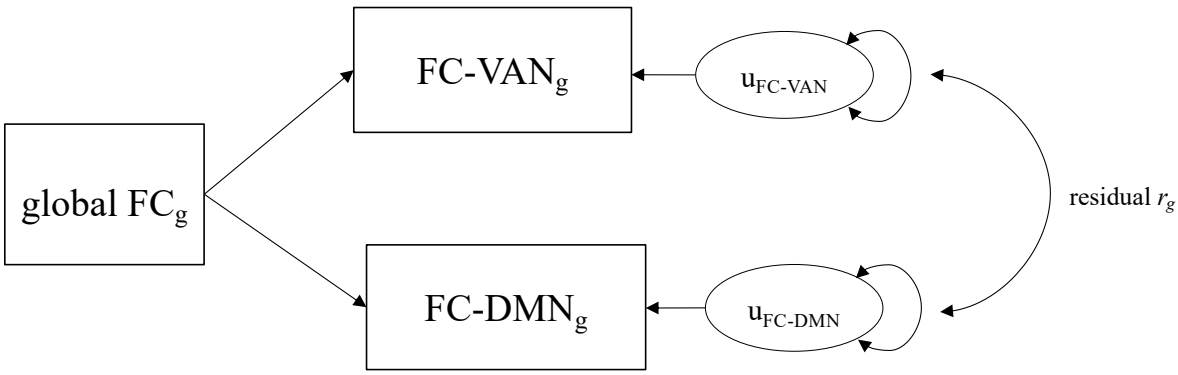


Figure 6. Path diagram of genomic SEM model. The summary statistics of two RSNs that have shown to significantly correlate with global connectivity will be used as input together with summary statistics of the global connectivity GWAS. In this way, r_g between the two RSNs can be estimated while taking global connectivity into account.

Local genetic correlations

The genome-wide r_g's described above are an average correlation of genetic effects across the genome, implicating that contrasting local r_g's are possibly cancelling each other out. Running r_g analysis on a locus level has the potential to uncover loci that show genetic similarity between traits. For this purpose, we adopted a three-step local r_g analysis pipeline similar to the genome-wide r_g analysis approach described above. All three steps were performed in LAVA[23], a local r_g analysis tool, using SNP-based GWAS summary statistics as input. We followed the suggested sample overlap procedure (as described on https://github.com/josefin-werme/LAVA) to enable LAVA to model shared variance due to sample overlap as residual covariance and consequently remove upward bias in local r_g estimates[23]. Since our GWASs included European samples, the 1,000 Genomes Phase 3 European data served as genotype reference and formed the basis of the locus definition file. For every locus, the first step of our pipeline consisted of estimating local bivariate r_g between 49 combinations of RSNs. However, RSNs that were devoid of heritable signal ($p > 1\times10^{-4}$) in the locus were excluded from local bivariate r_g analysis to ensure interpretability and reliability. A total of 774 bivariate tests were performed across 337 loci, leading to an adjusted significance threshold of $\alpha = (0.05/774 =) 6.46\times10^{-5}$. In the second step, RSNs that showed significant local r_g were additionally tested for r_g in that locus with global FC/SC. Note that if this was not possible, because global FC/SC showed no significant heritability in that locus, the local bivariate r_g between RSNs could not be biased by global FC/SC. If one or both RSNs did show additional significant r_g with global FC/SC, we ran a partial local r_g between the RSNs conditioned on the SC-global and/or FC-global phenotype in step three. If the partial local r_g between the RSNs no longer remained significant, we concluded that the initial r_g was driven by global FC/SC and did not reflect genetic overlap specific for these RSNs.

Acknowledgements

D.P. was funded by The Netherlands Organization for Scientific Research (NWO VICI 453-14-005), NWO Gravitation: BRAINSCAPES: A Roadmap from Neurogenetics to Neurobiology (Grant No. 024.004.012), and a European Research Council advanced grant (Grant No, ERC-2018-AdG

GWAS2FUNC 834057). The work of S.L. was supported by ZonMw Open Competition, project REMOVE 09120011910032. C.A.d.L. is funded by Hoffman-La Roche. The work of M.H. was supported by a VIDI (452-16-015) grant from the Netherlands Organization for Scientific Research (NWO) and an ERC Consolidator of the European Research Council (101001062). J.E.S. was supported by a VENI (201G-064) grant from the NOW. The research has been conducted using the UK Biobank Resource (application no. 16406). Analyses were carried out on the Genetic Cluster Computer hosted by the Dutch National computing and Networking Services SURFsara.

Author contributions

E.P.T., M.P.v.d.H. and D.P. conceived of the study, and J.W., M.N. and C.A.d.L. contributed additionally to its design. J.E.S. performed preprocessing of genetic data, while E.P.T. and S.C.d.L. performed preprocessing of neuroimaging data together. E.P.T. performed the analyses with contributions from J.W. and Y.W. E.P.T., J.W., S.C.d.L., M.N., C.A.d.L., M.P.v.d.H. and D.P. contributed to the interpretation of the data. E.P.T. drafted the work. E.P.T., J.W., M.N., D.P. and M.P.v.d.H. substantively revised it and S.C.d.L., J.E.S., Y.W. and C.A.d.L. revised it.

Competing interests

The authors report no competing interests.

Data availability

Genome-wide summary statistics will be made publicly available via https://ctg.cncr.nl/software/summary_statistics/ upon publication.

References

1. Buckholtz, J. W. & Meyer-Lindenberg, A. Psychopathology and the Human Connectome: Toward a Transdiagnostic Model of Risk For Mental Illness. *Neuron* **74**, 990–1004 (2012).
2. Yeo, B. T. T. *et al.* The organization of the human cerebral cortex estimated by intrinsic

functional connectivity. *J. Neurophysiol.* **106**, 2322–45 (2011).

3. Bressler, S. L. & Menon, V. Large-scale brain networks in cognition: emerging methods and principles. *Trends Cogn. Sci.* **14**, 277–290 (2010).

4. van den Heuvel, M. P. & Sporns, O. A cross-disorder connectome landscape of brain dysconnectivity. *Nat. Rev. Neurosci.* (2019). doi:10.1038/s41583-019-0177-6

5. Miranda-Dominguez, O. *et al.* Heritability of the human connectome: A connectotyping study. *Netw. Neurosci.* **2**, 175–199 (2017).

6. 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. *Proc. Natl. Acad. Sci.* **114**, (2017).

7. Anderson, K. M. *et al.* Heritability of individualized cortical network topography. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).

8. Meda, S. A. *et al.* Multivariate analysis reveals genetic associations of the resting default mode network in psychotic bipolar disorder and schizophrenia. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 2066–2075 (2014).

9. Barber, A. D., Hegarty, C. E., Lindquist, M. & Karlsgodt, K. H. Heritability of Functional Connectivity in Resting State: Assessment of the Dynamic Mean, Dynamic Variance, and Static Connectivity across Networks. *Cereb. cortex* **31**, 2834–2844 (2021).

10. Adhikari, B. M., Jahanshad, N., Reynolds, R. C., Cox, R. W. & Nichols, T. E. Heritability estimates on resting state fMRI data using ENIGMA analysis pipeline. *Biocomputing* 307–318 (2018).

11. Foo, H. *et al.* Novel genetic variants associated with brain functional networks in 18,445 adults from the UK Biobank. *Sci. Rep.* **11**, 1–12 (2021).

12. Zhao, B. *et al.* Common variants contribute to intrinsic human brain functional networks. *Nat. Genet.* (2022). doi:10.1038/s41588-022-01039-6

13. Roelfs, D. *et al.* Genetic overlap between multivariate measures of human functional brain connectivity and psychiatric disorders. *medRxiv* 1–24 (2022).

14. Van Den Heuvel, M. P., Mandl, R. C. W., Kahn, R. S. & Hulshoff Pol, H. E. Functionally

598 linked resting-state networks reflect the underlying structural connectivity architecture of the
599 human brain. *Hum. Brain Mapp.* **30**, 3127–3141 (2009).
600 15. Batista-García-Ramó, K. & Fernández-Verdecia, C. I. What we know about the brain
601 structure-function relationship. *Behav. Sci. (Basel).* **8**, (2018).
602 16. Mollink, J. *et al.* The spatial correspondence and genetic influence of interhemispheric
603 connectivity with white matter microstructure. *Nat. Neurosci.* **22**, (2019).
604 17. Gu, Z., Jamison, K. W., Sabuncu, M. R. & Kuceyeski, A. Heritability and interindividual
605 variability of regional structure-function coupling. *Nat. Commun.* **12**, 1–12 (2021).
606 18. Zhao, B. *et al.* Common genetic variation influencing human white matter microstructure.
607 *Science (80-.).* **372**, (2021).
608 19. Smith, S. M. *et al.* An expanded set of genome-wide association studies of brain imaging
609 phenotypes in UK Biobank. *Nat. Neurosci.* **24**, 737–745 (2021).
610 20. Sha, Z., Schijven, D., Fisher, S. E. & Francks, C. Genetic architecture of the white matter
611 connectome of the human brain. *bioRxiv* (2022).
612 21. Lange, S. C. de *et al.* Shared vulnerability for connectome alterations across psychiatric and
613 neurological brain disorders. *Nat. Hum. Behav.* (2019). doi:10.1101/360586
614 22. Uffelmann, E. & Posthuma, D. Emerging methods and resources for biological interrogation
615 of neuropsychiatric polygenic-signal. *Biol. Psychiatry* (2020).
616 doi:10.1016/j.biopsych.2020.05.022
617 23. Werme, J., Van Der Sluis, S., Posthuma, D. & de Leeuw, C. A. An integrated framework for
618 local genetic correlation analysis. *Nat. Genet* **54**, 274–282 (2022).
619 24. Wei, Y. *et al.* Genetic mapping and evolutionary analysis of human-expanded cognitive
620 networks. *Nat. Commun.* **10**, 1–11 (2019).
621 25. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and
622 annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
623 26. Mufford, M. S. *et al.* The Genetic Architecture of Amygdala Nuclei. *medRxiv*
624 2021.06.30.21258615 (2021).
625 27. Cooper, E. M., Boeke, J. D. & Cohen, R. E. Specificity of the BRISC deubiquitinating enzyme

626 is not due to selective binding to Lys63-linked polyubiquitin. *J. Biol. Chem.* **285**, 10344–
627 10352 (2010).
628 28. DiAntonio, A. & Hicke, L. Ubiquitin-dependent regulation of the synapse. *Annu. Rev.*
629 *Neurosci.* **27**, 223–246 (2004).
630 29. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human
631 genetic variants. *Nat. Genet.* **46**, 310–315 (2014).
632 30. Zhou, Y., Gunput, R. A. F. & Pasterkamp, R. J. Semaphorin signaling: progress made and
633 promises ahead. *Trends Biochem. Sci.* **33**, 161–170 (2008).
634 31. Jakobsson, M. E., Małecki, J. & Falnes, P. Ø. Regulation of eukaryotic elongation factor 1
635 alpha (eEF1A) by dynamic lysine methylation. *RNA Biol.* **15**, 314–319 (2018).
636 32. Kizil, C. *et al.* Simplet/Fam53b is required for Wnt signal transduction by regulating β-catenin
637 nuclear localization. *Development* **141**, 3529–3539 (2014).
638 33. Rao, J. *et al.* Advillin acts upstream of phospholipase C ϵ1 in steroid-resistant nephrotic
639 syndrome. *J. Clin. Invest.* **127**, 4257–4269 (2017).
640 34. Backman, J. T., Filppula, A. M., Niemi, M. & Neuvonen, P. J. Role of Cytochrome P450 2C8
641 in drug metabolism and interactions. *Pharmacol. Rev.* **68**, 168–241 (2016).
642 35. Wang, Q. *et al.* A Bayesian framework that integrates multi-omics data and gene networks
643 predicts risk genes from schizophrenia GWAS data. *Nat. Neurosci.* **22**, 691–699 (2019).
644 36. Kaiser, R. H., Andrews-Hanna, J. R., Wager, T. D. & Pizzagalli, D. A. Large-scale network
645 dysfunction in major depressive disorder: A meta-analysis of resting-state functional
646 connectivity. *JAMA Psychiatry* **72**, 603–611 (2015).
647 37. Watanabe, K. *et al.* A global overview of pleiotropy and genetic architecture in complex traits.
648 *Nat. Genet.* (2019). doi:10.1101/500090
649 38. Blake, J. A. & Ziman, M. R. Pax genes: Regulators of lineage specification and progenitor cell
650 maintenance. *Dev.* **141**, 737–751 (2014).
651 39. Di Magliano, M. P., Di Lauro, R. & Zannini, M. Pax8 has a key role in thyroid cell
652 differentiation. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 13144–13149 (2000).
653 40. Kumar, M. *et al.* Alteration in intrinsic and extrinsic functional connectivity of resting state

654 networks associated with subclinical hypothyroidism. *J. Neuroendocrinol.* **30**, 1–11 (2018).

655 41. Singh, S. *et al.* Alterations of Functional Connectivity Among Resting-State Networks in
656 Hypothyroidism. *J. Neuroendocrinol.* **27**, 609–615 (2015).

657 42. Emrani, S., Arain, H. A., DeMarshall, C. & Nuriel, T. APOE4 is associated with cognitive and
658 pathological heterogeneity in patients with Alzheimer's disease: a systematic review.
659 *Alzheimer's Res. Ther.* **12**, 1–19 (2020).

660 43. Wightman, D. P. *et al.* A genome-wide association study with 1,126,563 individuals identifies
661 new risk loci for Alzheimer's disease. *Nat. Genet.* **53**, (2021).

662 44. Bis, J. C. *et al.* Common variants at 12q14 and 12q24 are associated with hippocampal
663 volume. *Nat. Genet.* **44**, (2012).

664 45. Menon, V. & Uddin, L. Q. Saliency, switching, attention and control: a network model of
665 insula function. *Brain Struct. Funct.* **214**, 655–667 (2010).

666 46. Fox, M. D. & Raichle, M. E. Spontaneous fluctuations in brain activity observed with
667 functional magnetic resonance imaging. *Nat. Rev. Neurosci.* **8**, 700–711 (2007).

668 47. Seeley, W. W., Crawford, R. K., Zhou, J., Miller, B. L. & Greicius, M. D. Neurodegenerative
669 Diseases Target Large-Scale Human Brain Networks. *Neuron* **62**, 42–52 (2009).

670 48. Chiesa, P. A. *et al.* Revolution of resting-state functional neuroimaging genetics in
671 Alzheimer's Disease. *Trends Neurosci.* **40**, (2017).

672 49. Rolls, E. T. The cingulate cortex and limbic systems for emotion, action, and memory. *Brain*
673 *Struct. Funct.* **224**, 3001–3018 (2019).

674 50. Liu, C., Pu, W., Wu, G., Zhao, J. & Xue, Z. Abnormal resting-state cerebral-limbic functional
675 connectivity in bipolar depression and unipolar depression. *BMC Neurosci.* **20**, 1–8 (2019).

676 51. Shi, Y. *et al.* Genetic variation in the calcium/calmodulin-dependent protein kinase (CaMK)
677 pathway is associated with antidepressant response in females. *J. Affect. Disord.* **136**, 558–566
678 (2012).

679 52. Kanders, S. H. *et al.* A pharmacogenetic risk score for the evaluation of major depression
680 severity under treatment with antidepressants. *Drug Dev. Res.* **81**, 102–113 (2020).

681 53. Kim, H. Dissociating the roles of the default-mode, dorsal, and ventral networks in episodic

682 memory retrieval. *Neuroimage* **50**, 1648–1657 (2010).

683 54. Raichle, M. E. The Brain's Default Mode Network. *Annu. Rev. Neurosci.* **38**, 433–447 (2015).

684 55. Domakonda, M. J., He, X., Lee, S., Cyr, M. & Marsh, R. Increased Functional Connectivity

685 Between Ventral Attention and Default Mode Networks in Adolescents With Bulimia

686 Nervosa. *J. Am. Acad. Child Adolesc. Psychiatry* **58**, 232–241 (2019).

687 56. Moqrich, A. *et al.* Cloning of human striatin cDNA (STRN), gene mapping to 2p22-p21, and

688 preferential expression in brain. *Genomics* **51**, 136–139 (1998).

689 57. Eastwood, S. L., Law, A. J., Everall, I. P. & Harrison, P. J. The axonal chemorepellant

690 semaphorin 3A is increased in the cerebellum in schizophrenia and may contribute to its

691 synaptic pathology. *Mol. Psychiatry* **8**, 148–155 (2003).

692 58. Damaraju, E. *et al.* Dynamic functional connectivity analysis reveals transient states of

693 dysconnectivity in schizophrenia. *NeuroImage Clin.* **5**, 298–308 (2014).

694 59. Ford, J. M. *et al.* Visual hallucinations are associated with hyperconnectivity between the

695 amygdala and visual cortex in people with a diagnosis of schizophrenia. *Schizophr. Bull.* **41**,

696 223–232 (2015).

697 60. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat.*

698 *Genet.* **47**, 1236–1241 (2015).

699 61. Suárez, L. E., Markello, R. D., Betzel, R. F. & Misic, B. Linking Structure and Function in

700 Macroscale Brain Networks. *Trends Cogn. Sci.* (2020). doi:10.1016/j.tics.2020.01.008

701 62. Visscher, P. M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am.*

702 *J. Hum. Genet.* **101**, 5–22 (2017).

703 63. Sudlow, C. *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a

704 Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med.* **12**, 1–10 (2015).

705 64. Jansen, P. R. *et al.* Genome-wide meta-analysis of brain volume identifies genomic loci and

706 genes shared with intelligence. *Nat. Commun.* **11**, (2020).

707 65. Smith, S. M., Alfaro-almagro, F. & Miller, K. L. UK Biobank Brain Imaging Documentation.

708 *biobank.ctsu.ox.ac.uk/crystal/docs/brain_mri.pdf* (2020).

709 66. Fischl, B. FreeSurfer. *Neuroimage* **62**, 774–781 (2012).

67. Lange, S. C. de & van den Heuvel, M. P. Structural and functional connectivity reconstruction with CATO - A Connectivity Analysis TOolbox. *bioRxiv* **43**, 745–753 (2021).

68. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. FSL. *Neuroimage* **62**, 782–790 (2012).

69. Cammoun, L. *et al.* Mapping the human connectome at multiple scales with diffusion spectrum MRI. *J. Neurosci. Methods* **203**, 386–397 (2012).

70. Mori, S., Crain, B. J., Chacko, V. P. & Zijl, P. C. M. van. Three-Dimensional Tracking of Axonal Projections in the Brain by Magnetic Resonance Imaging. *Ann. Neurol.* **45**, 265–269 (1999).

71. Beaulieu, C. The basis of anisotropic water diffusion in the nervous system - A technical review. *NMR Biomed.* **15**, 435–455 (2002).

72. 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).

73. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

74. Day, F. R., Loh, P. R., Scott, R. A., Ong, K. K. & Perry, J. R. B. A Robust Example of Collider Bias in a Genetic Association Study. *Am. J. Hum. Genet.* **98**, 392–393 (2016).

75. Abraham, G., Qiu, Y. & Inouye, M. FlashPCA2: principal component analysis of Biobank-scale genotype datasets. *Bioinformatics* **33**, 2776–2778 (2017).

76. Alfaro-Almagro, F. *et al.* Confound modelling in UK Biobank brain imaging. *Neuroimage* 117002 (2020). doi:10.1016/j.neuroimage.2020.117002

77. Visscher, P. M., Hill, W. G. & Wray, N. R. Heritability in the genomics era - Concepts and misconceptions. *Nat. Rev. Genet.* **9**, 255–266 (2008).

78. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, 1–7 (2010).

79. Boyle, A. P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22**, 1790–1797 (2012).

80. Ernst, J. & Kellis, M. ChromHMM: Automating chromatin-state discovery and

738 characterization. *Nat. Methods* **9**, 215–216 (2012).

739 81. Ramasamy, A. *et al.* Genetic variability in the regulation of gene expression in ten regions of
740 the human brain. *Nat. Neurosci.* **17**, 1418–1428 (2014).

741 82. The GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across
742 human tissues. *Science (80-.).* **369**, 1318–1330 (2020).

743 83. Ng, B. *et al.* An xQTL map integrates the genetic architecture of the human brain's
744 transcriptome and epigenome. *Nat. Neurosci.* **20**, 1418–1426 (2017).

745 84. Wang, D. *et al.* Comprehensive functional genomic resource and integrative model for the
746 human brain. *Science (80-.).* **362**, (2018).

747 85. Giusti-Rodriguez, P. *et al.* A Chromatin Catalog for the Interpretation of Genetic Associations
748 of Psychiatric Disorders. *Eur. Neuropsychopharmacol.* **29**, (2019).

749 86. Schmitt, A. D. *et al.* A Compendium of Chromatin Contact Maps Reveals Spatially Active
750 Regions in the Human Genome. *Cell Rep.* **17**, 2042–2059 (2016).

751 87. Roadmap Epigenomics Consortium *et al.* Integrative analysis of 111 reference human
752 epigenomes. *Nature* **518**, 317–329 (2015).

753 88. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set
754 Analysis of GWAS Data. *PLoS Comput. Biol.* **11**, 1–19 (2015).

755 89. Grotzinger, A. D. *et al.* Genomic structural equation modelling provides insights into the
756 multivariate genetic architecture of complex traits. *Nat. Hum. Behav.* **3**, 513–525 (2019).

757