Supplementary Materials

Functional connectome-wide associations of schizophrenia polygenic risk

Hengyi Cao, Hang Zhou, Tyrone D. Cannon

Yale University

Supplementary Methods

Subjects

The CNP sample was recruited from the greater Los Angeles area. All participants in the studied sample completed a battery of seven paradigms employed in the consortium (i.e. resting state, risk taking, working memory, episodic memory encoding, episodic memory retrieval, stop signal, task switching). Diagnoses for the three disorders in the sample were based on the Structured Clinical Interview for DSM-IV [1] and the Adult ADHD Interview [2]. Any patients who met criteria for more than one diagnosis were excluded. For patients with schizophrenia and bipolar disorders, the positive and negative symptoms were assessed by the Scale for the Assessment of Positive Symptoms (SAPS [3]) and the Scale for the Assessment of Negative Symptoms (SANS [3]), respectively. Verbal IQ was assessed by the vocabulary subtest of the Wechsler Adult Intelligence Scale (WAIS [4]), and performance IQ was quantified by the matrix reasoning subtest of WAIS. The overall IQ scores were calculated as the sum of the scaled verbal and performance IQ scores. Healthy subjects were excluded if they had a life-time diagnosis of a major Axis-I disorder, substance abuse or significant medical illness. See Poldrack et al. [5] for a detailed sample description of this public dataset.

Data acquisition

The imputed HCP genetic data were downloaded from https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001364.v1.p1. Briefly, DNAs were extracted from the blood samples for each subject, and genotyping was based on the Illumina Infinium Multi-Ethnic Genotyping Array (MEGA). The genetic imputation was performed using IMPUTE2 [6] based on the 1000 Genomes released v.5 phase 3 reference panel (ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20130502). We performed PRS calculations based on these imputed data.

The preprocessed HCP imaging data were downloaded from https://db.humanconnectome.org/. Briefly, data were collected from a customized 3T Siemens Connectom scanner equipped with a 32-channel head coil. The resting-state paradigm was scanned with four sessions spanned across two consecutive days, and the other paradigms were scanned with two sessions on a single day. To balance the amount of data across paradigms and to avoid overrepresentation of resting state in the following cross-paradigm computations, we only used the resting-state data from the two sessions acquired on the first day. The following parameters were used for fMRI data acquisition: TR = 720 ms, TE = 33.1 ms, FA = 52°, FOV = 208 × 180 mm^2, 2 mm slice thickness, 72 slices, multi-band factor = 8. Of note, unlike the commonly applied phase encoding directions along the "anterior-posterior" axis, the HCP fMRI data were acquired using left-right and right-left phase encoding directions to accelerate scan time and to minimize image distortion [7]. The CNP data were downloaded from https://openneuro.org/datasets/ds000030. Data were collected from 3T Siemens Trio scanners located at UCLA using the same GRE-EPI sequence: TR = 2s, TE = 30 ms, 90 degree flip angle, 34 4-mm slices, 192 mm FOV.

Imaging Processing

The HCP imaging data were preprocessed based on the standard HCP pipeline [8], including a total

of five major steps (PreFreeSurfer, FreeSurfer, PostFreeSurfer, fMRIVolume, fMRISurface). In brief, images were corrected for gradient nonlinearity induced distortion, head motion, and phase-encoding related distortion, registered to individuals' T1 weighted images, and normalized to the MNI space. The CNP data were preprocessed using the standard pipeline implemented in the Statistical Parametric Mapping software (SPM12, https://www.fil.ion.ucl.ac.uk/spm/software/spm12/), including slice timing correction, head motion realignment, functional-structural image coregistration, MNI space normalization, and spatial smoothing.

The preprocessed data were further scrutinized for excessive head motion. All subjects in the HCP sample had an average frame-wise displacement (FD) below 0.4mm. Therefore, no extra subjects were excluded. The mean time series for each of the 270 nodes in the expanded Power brain atlas [9-11] were extracted from the preprocessed images. The extracted time series were further corrected for task-related coactivations (for task data), white matter and cerebrospinal fluid signals, 24 head motion parameters (i.e. the 6 rigid-body parameters generated from the realignment step, their first derivatives, and the squares of these 12 parameters), and FD. The corrected time series were then high-pass filtered at 0.008 Hz to account for scan noise. Subsequently, a 270 × 270 pairwise connectivity matrix was generated using Pearson correlations for each subject during each paradigm and scan session. Following the previous suggestions for the HCP data [7], we averaged the derived connectivity matrices across the two scan sessions for each paradigm to boost signals, generating eight paradigm-dependent connectivity matrices for each subject in the HCP data.

All paradigm-specific connectivity matrices for the same subjects were then entered into a principal component analysis (PCA) using singular value decomposition. Essentially, the PCA method decomposes the original connectivity matrices into a set of principal components (PCs) that are orthogonal to each other. Each generated component is a linear combination of original matrices, and the components are organized in the way such that the first PC accounts for the largest variance in the original data, and the second PC accounts for the second largest variance in the original data, and so on. In our data, the PCA yielded a total of eight PCs, where the first PC scores represent the shared connectivity patterns that explain the most variance across all paradigms. We extracted these first PC scores for each subject for further statistical analysis.

The network-based statistic (NBS) analysis was used to associate whole-brain connectivity with PRSs. For details of this established method, please see previous publications [12, 13]. In brief, NBS is an approach to effectively control for cluster-level family-wise error (FWE) in the connectomic analysis, which offers a larger power than mass-univariate tests on independent connections. Specifically, this method used a non-parametric permutation-based approach to control for type I error for the identified connectivity clusters. First, the original cluster was identified using an initial threshold of $P < 0.001$. The data were then permuted for 10,000 times, where during each permutation the same threshold ($P < 0.001$) was applied and the size (T values) of the identified cluster was recalculated. This generated a null distribution for the identified cluster. The corrected P value was then determined by the proportion of the derived cluster sizes in the permutation distribution that were larger than the observed effect. A P value below 0.05 indicated that the

family-wise false positive rate of the identified cluster was smaller than 5%.

Supplementary Results

NBS analysis on each of the eight paradigms in the HCP data

To assess whether single-paradigm fMRI data would be sufficient to detect the observed effect, we further performed the NBS analysis on each of the eight paradigms in the HCP data. No significant effects (P_{FWE} > 0.05) were found for any of the eight paradigms alone, suggesting that single paradigm may not have sufficient power to detect the observed PRS-connectome relationships.

Supplementary analysis using a different temporal filter

In this study, we used a high-pass filter (cutoff 0.008Hz) for data processing due to the following reasons: 1) While band-pass filtering is a common practice for resting-state data, high-pass filtering is normally used for task-based data in order to maintain high-frequency task-related effects in signals. These effects primarily include the trial-to-trial variability, a critical resource associated with neurophysiology and human behaviors [14, 15]; and 2) Even for resting-state data, the use of band-pass filter to remove high-frequency signals is under scrutiny in recent years. Several studies have shown that high-frequency signals may contain neural information as well and may meaningfully contribute to functional connectivity [16-19]. As task-based data were the major focus in this study, we used high-pass filtering to maximally maintain useful neural signals. Here, in order to test the robustness of our results across different frequency ranges, we further reprocessed the data using band-pass filtering (0.008-0.1Hz). We found that the identified network in data derived from band-pass filtering was also significantly correlated with PRSs (R = -0.35, P = 2e-16), with the effect size very similar to that observed using high-pass filtering. This analysis suggests that the observed connectomic associations are not driven by a specific filtering approach.

Supplementary analysis using a reverse strategy

To further investigate whether connectivity in other brain networks would relate to both cognitive ability and polygenic risk, we used a reverse strategy as the one we used in the paper. That is, we first used NBS analysis to identify connectomic associations with IQ scores in the CNP sample, and then examined whether the identified networks would be correlated with PRSs in the HCP sample. The NBS analysis used the same procedure and parameters as reported in the paper, controlling for age, sex, head motion, and diagnostic group. This analysis revealed two large-scale networks in the CNP sample, with one positively correlated with IQ and the other negatively correlated with IQ (P_{FWE} < 0.05, see Fig S3). However, the connectivity of these two networks were not correlated with PRSs in the HCP sample (P > 0.42). This analysis suggested that while cognitive deficits in patients are associated with more widely distributed networks in the brain, only the network reported in the manuscript is related to both polygenic risk and cognitive functioning in patients with schizophrenia.

Supplementary Tables

Table S1. Demographic and clinical data for the CNP sample.

	SZ (n = 44)	BD (n = 43)	ADHD (n = 34)	HC (n = 77)	*P* value
Age (years)	35.80 ± 8.94	35.21 ± 8.87	31.09 ± 9.85	30.70 ± 8.54	0.004 (overall) 0.02 (SZ vs HC)
Sex (M/F)	34/10	25/18	18/16	43/34	0.08 (overall)
IQ – Performance	10.05 ± 2.87	13.16 ± 3.05	13.53 ± 2.63	13.39 ± 2.86	<0.001 (overall) <0.001 (SZ vs HC)
IQ – Verbal	7.57 ± 2.22	10.28 ± 2.67	10.62 ± 2.52	10.65 ± 2.25	<0.001 (overall) <0.001 (SZ vs HC)
IQ – Overall	17.61 ± 4.14	23.44 ± 4.86	24.15 ± 4.29	24.04 ± 4.47	<0.001 (overall) <0.001 (SZ vs HC)
SAPS – Hallucinations	8.27 ± 7.58	0.15 ± 0.80	-	-	<0.001 (SZ vs BD)
SAPS – Delusions	12.45 ± 10.26	2.10 ± 3.32	-	-	<0.001 (SZ vs BD)
SAPS – Bizarre behavior	2.82 ± 3.87	1.55 ± 2.34	-	-	0.08 (SZ vs BD)
SAPS – Thought disorder	7.82 ± 8.26	4.75 ± 5.78	-	-	0.05 (SZ vs BD)
SANS – Affective blunting	5.73 ± 7.04	2.43 ± 4.98	-	-	0.02 (SZ vs BD)
SANS – Alogia	3.16 ± 3.67	0.53 ± 1.30	-	-	<0.001 (SZ vs BD)
SANS – Avolition	10.48 ± 6.04	7.98 ± 2.25	-	-	0.05 (SZ vs BD)
SANS – Anhedonia	10.52 ± 6.29	7.48 ± 5.57	-	-	0.02 (SZ vs BD)
SANS – Attention	5.76 ± 3.55	2.85 ± 2.77	-	-	<0.001 (SZ vs BD)

Table S2. Nodes and edges in the identified PRS-associated network. The vast majority of nodes were mapped to the visual system (in red), default-mode system (in blue), and frontoparietal system (in yellow).

	Edges	
Edge No.	**Node 1 (belonged system)**	**Node 2 (belonged system)**
1	Occipital_Inf_L (visual)	Temporal_Inf_L (default-mode)
2	Temporal_Inf_L (default-mode)	Supp_Motor_Area_L (cingulo-opercular)
3	Temporal_Mid_R (default-mode)	Frontal_Med_Orb_R (default-mode)
4	Temporal_Sup_R (auditory)	Precuneus_L (default-mode)
5	SupraMarginal_L (auditory)	Precuneus_L (default-mode)
6	Frontal_Med_Orb_R (default-mode)	Frontal_Mid_L (default-mode)
7	Frontal_Med_Orb_R (default-mode)	Frontal_Sup_Medial_L (default-mode)
8	Precentral_L (sensorimotor)	Temporal_Mid_R (default-mode)
9	Frontal_Sup_Medial_L (default-mode)	Temporal_Mid_R (default-mode)
10	Paracentral_Lobule_L (sensorimotor)	Temporal_Pole_Mid_R (default-mode)
11	Frontal_Sup_Medial_L (default-mode)	Temporal_Pole_Mid_R (default-mode)
12	Temporal_Mid_R (default-mode)	Lingual_L (visual)
13	Temporal_Mid_R (default-mode)	Lingual_R (visual)
14	Temporal_Mid_R (default-mode)	Calcarine_R (visual)
15	Temporal_Mid_R (default-mode)	Calcarine_R (visaul)
16	Temporal_Mid_R (default-mode)	Calcarine_L (visual)
17	Precuneus_L (default-mode)	Calcarine_L (visual)
18	Temporal_Mid_R (default-mode)	Calcarine_L (visual)
19	Lingual_R (visual)	Occipital_Mid_L (visual)
20	Temporal_Mid_R (default-mode)	Fusiform_R (visual)
21	Precuneus_L (default-mode)	Calcarine_L (visual)
22	Temporal_Mid_R (default-mode)	Calcarine_L (visual)
23	Temporal_Inf_L (default-mode)	Occipital_Inf_L (visual)
24	Lingual_R (visual)	Occipital_Sup_R (visual)
25	Temporal_Mid_R (default-mode)	Fusiform_R (visual)
26	Frontal_Sup_Orb_R (frontoparietal)	Occipital_Mid_L (visual)
27	Temporal_Mid_R (default-mode)	Occipital_Mid_L (visual)
28	Temporal_Mid_R (default-mode)	Occipital_Mid_L (visual)
29	Lingual_R (visual)	Occipital_Mid_L (visual)
30	Temporal_Inf_L (default-mode)	Occipital_Mid_R (visual)
31	Lingual_R (visual)	Occipital_Mid_R (visual)
32	Temporal_Inf_L (default-mode)	Calcarine_R (visual)
33	Temporal_Mid_R (default-mode)	Calcarine_R (visual)
34	Temporal_Mid_R (default-mode)	Calcarine_R (visual)
35	Temporal_Mid_R (default-mode)	Occipital_Inf_L (visual)
36	Temporal_Mid_R (default-mode)	Occipital_Mid_R (visual)
37	Temporal_Inf_R (default-mode)	Frontal_Inf_Tri_R (frontoparietal)

38	Temporal_Mid_R (default-mode)	Frontal_Inf_Tri_R (frontoparietal)
39	Lingual (visual)	Frontal_Inf_Tri_R (frontoparietal)
40	Frontal_Sup_R (default-mode)	Frontal_Inf_Tri_L (frontoparietal)
41	Temporal_Mid_R (default-mode)	Frontal_Inf_Tri_L (frontoparietal)
42	Precuneus_L (default-mode)	Parietal_Inf_L (frontoparietal)
43	Temporal_Mid_R (default-mode)	Temporal_Inf_R (frontoparietal)
44	Temporal_Mid_R (default-mode)	Temporal_Inf_R (frontoparietal)
45	Temporal_Mid_R (default-mode)	Precentral_R (frontoparietal)
46	Temporal_Mid_R (default-mode)	Precentral_L (frontoparietal)
47	Precuneus_L (default-mode)	Precentral_L (frontoparietal)
48	Temporal_Mid_R (default-mode)	Precentral_L (frontoparietal)
49	Temporal_Mid_R (default-mode)	Parietal_Inf_L (frontoparietal)
50	Lingual_R (visual)	Parietal_Inf_R (frontoparietal)
51	Temporal_Mid_R (default-mode)	Frontal_Sup_Medial_L (frontoparietal)
52	Frontal_Med_Orb_R (default-mode)	Frontal_Sup_Medial_L (frontoparietal)
53	Precuneus_L (default-mode)	Frontal_Sup_Medial_L (frontoparietal)
54	Temporal_Mid_R (default-mode)	Frontal_Sup_Medial_L (frontoparietal)
55	Temporal_Pole_Mid_R (default-mode)	Frontal_Sup_Medial_L (frontoparietal)
56	Precuneus_L (default-mode)	Insula_L (salience)
57	Precuneus_L (default-mode)	Supp_Motor_Area_L (salience)
58	Temporal_Mid_R (default-mode)	Supp_Motor_Area_L (salience)
59	Precuneus_L (default-mode)	Frontal_Mid_L (salience)
60	Temporal_Mid_R (default-mode)	Cingulum_Mid_R (salience)
61	Precuneus_L (default-mode)	Frontal_Mid_R (salience)
62	Precuneus_L (default-mode)	Cingulum_Mid_R (salience)
63	Temporal_Mid_R (default-mode)	Putamen_R (subcortex)
64	Temporal_Mid_R (default-mode)	Temporal_Mid_L (attention)
65	Temporal_Mid_R (default-mode)	Temporal_Sup_R (attention)
66	Lingual_R (visual)	Occipital_Sup_R (attention)
67	Temporal_Inf_L (default-mode)	Occipital_Mid_L (attention)
68	Temporal_Mid_R (default-mode)	Temporal_Inf_L (attention)
69	Frontal_Inf_Tri_R (frontoparietal)	Accumben_R (subcortex)

Supplementary Figures

Fig S1. Percent of variance explained by the first principal components (PCs) in the cross-paradigm analyses. For all groups in both samples, the first PCs explained 64% - 70% of all variance across paradigms, supporting a state-independent "trait" architecture of functional connectome. No significant between-group differences were shown for the percentages, suggesting that the observed group differences are not driven by differences in the amount of extracted variance.

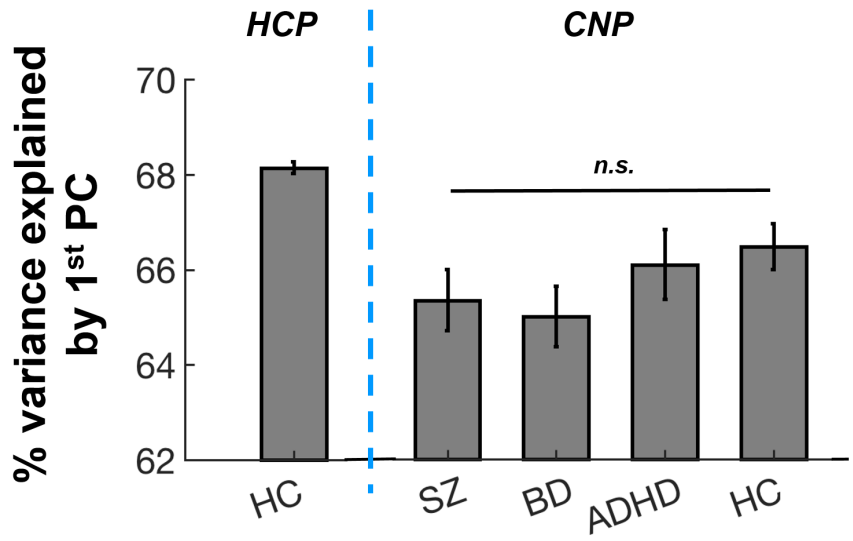


Fig S2. Robustness of the observed connectomic finding across multiple GWAS *P*-value thresholds. The associations remained significant for all computed thresholds (from 5E-7 to 5E-2).

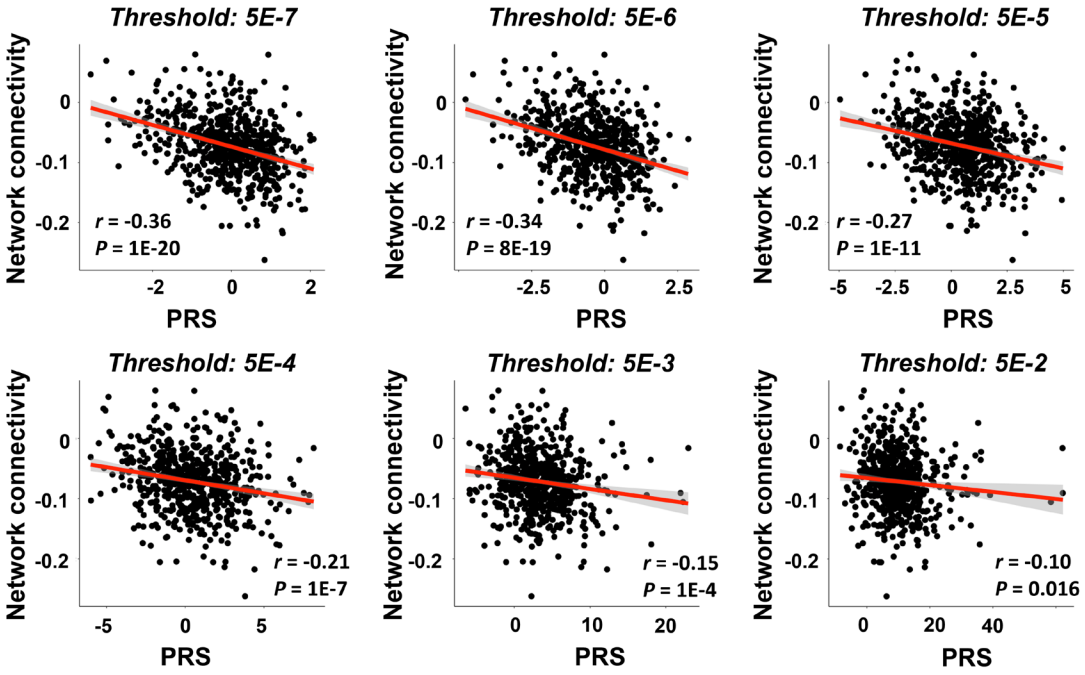
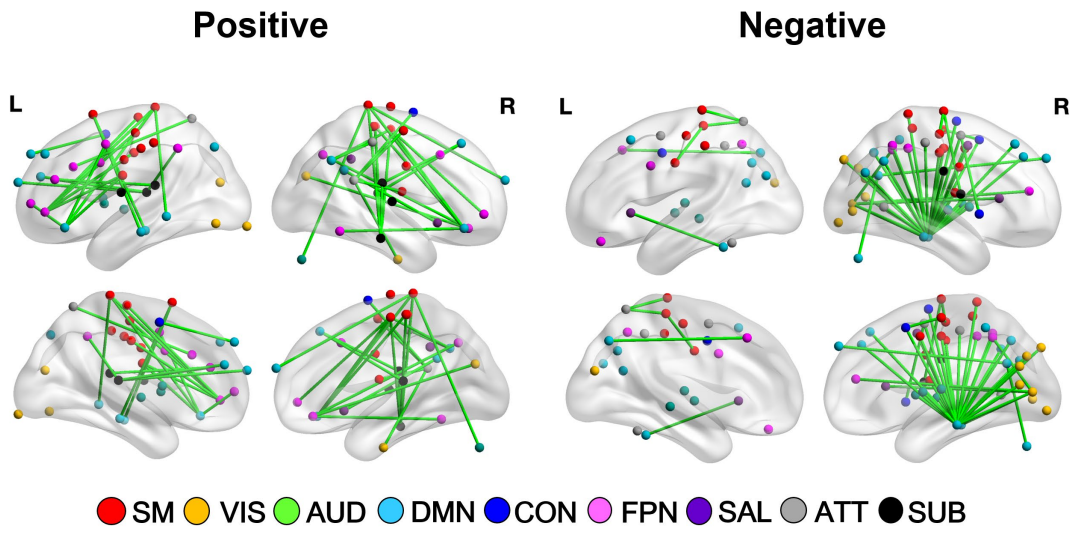


Fig S3. Networks associated with IQ scores in the CNP sample as revealed by the NBS analysis (left panel: positive associations; right panel: negative associations). Neither of these networks were related to PRSs in the HCP sample.



Supplementary References

1. First MB, Spitzer RL, M.; G, Williams JBW. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P). *New York: Biometrics Research, New York State Psychiatric Institute* 2002.

2. Kaufman J, Birmaher B, Brent DA, Ryan ND, Rao U. K-Sads-Pl. *Journal of the American Academy of Child and Adolescent Psychiatry* 2000; **39**(10)**:** 1208.

3. Andreasen NC. Methods for assessing positive and negative symptoms. *Schizophrenia: Positive and negative symptoms and syndromes.* Karger: Basel, Switzerland, 1990, pp 73-88.

4. Wechsler D. Wechsler Adult Intelligence Scale - Third Edition (WAIS-III). *The Psychological Corporation* 1997.

5. Poldrack RA, Congdon E, Triplett W, Gorgolewski KJ, Karlsgodt KH, Mumford JA *et al.* A phenome-wide examination of neural and cognitive function. *Sci Data* 2016; **3:** 160110.

6. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 2009; **5**(6)**:** e1000529.

7. Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G *et al.* Resting-state fMRI in the Human Connectome Project. *Neuroimage* 2013; **80:** 144-168.

8. Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* 2013; **80:** 105-124.

9. Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA *et al.* Functional network organization of the human brain. *Neuron* 2011; **72**(4)**:** 665-678.

10. Cao H, Chen OY, Chung Y, Forsyth JK, McEwen SC, Gee DG *et al.* Cerebello-thalamo-cortical hyperconnectivity as a state-independent functional neural signature for psychosis prediction and characterization. *Nat Commun* 2018; **9**(1)**:** 3836.

11. Cao H, McEwen SC, Forsyth JK, Gee DG, Bearden CE, Addington J *et al.* Toward Leveraging Human Connectomic Data in Large Consortia: Generalizability of fMRI-Based Brain Graphs Across Sites, Sessions, and Paradigms. *Cereb Cortex* 2018.

12. Zalesky A, Fornito A, Bullmore ET. Network-based statistic: identifying differences in brain networks. *Neuroimage* 2010; **53**(4)**:** 1197-1207.

13. Zalesky A, Cocchi L, Fornito A, Murray MM, Bullmore E. Connectivity differences in brain networks. *Neuroimage* 2012; **60**(2)**:** 1055-1062.

14. Fox MD, Snyder AZ, Zacks JM, Raichle ME. Coherent spontaneous activity accounts for trial-to-trial

variability in human evoked brain responses. *Nat Neurosci* 2006; **9**(1)**:** 23-25.

15. Pessoa L, Gutierrez E, Bandettini P, Ungerleider L. Neural correlates of visual working memory: fMRI amplitude predicts task performance. *Neuron* 2002; **35**(5)**:** 975-987.

16. Boubela RN, Kalcher K, Huf W, Kronnerwetter C, Filzmoser P, Moser E. Beyond noise: using temporal ICA to extract meaningful information from high-frequency fMRI signal fluctuations during rest. *Frontiers in Human Neuroscience* 2013; **7**.

17. Yuan BK, Wang J, Zang YF, Liu DQ. Amplitude differences in high-frequency fMRI signals between eyes open and eyes closed resting states. *Frontiers in Human Neuroscience* 2014; **8**.

18. Trapp C, Vakamudi K, Posse S. On the detection of high frequency correlations in resting state fMRI. *Neuroimage* 2018; **164:** 202-213.

19. Niazy RK, Xie JY, Miller K, Beckmann CF, Smith SM. Spectral characteristics of resting state networks. *Prog Brain Res* 2011; **193:** 259-276.