SUPPLEMENTARY INFORMATION

Genetic Heterogeneity Shapes Brain Connectivity in Psychiatry

Supplementary Materials and Methods

Cohorts

SFARI SimonsVIP-dataset

Copy number variants (CNV) carriers were clinically ascertained. Imaging data of 16p11.2 CNV carriers and typically developing controls were acquired by the Simons variation in individuals project (VIP) consortium (1) across 2 sites. We excluded 50 individuals from the analysis due to the insufficient quality of the imaging data. The final 16p11.2 sample includes 146 individuals: 16p11.2 deletion (n=25) and duplication (n=26) carriers, 1q21.1 deletion (n=7) and duplication (n=4) carriers, and extrafamilial controls (n=84). Over 90% of the deletion carriers and 69% of the duplication carriers met the criteria for at least one clinical psychiatric diagnosis. Control subjects were recruited from the general population (extra-familial subjects) and had no major DSM-V diagnosis. The 16p11.2 duplication group includes 2 individuals with a triplication.

UCLA 22q11.2-dataset

CNV carriers were clinically ascertained. Imaging data of 22q11.2 CNV carriers and typically developing (TD) controls were acquired at the University of California, Los Angeles (UCLA). Patients were ascertained from the UCLA or Children's Hospital, Los Angeles Pediatric Genetics, Allergy/Immunology and/or Craniofacial Clinics. Demographically comparable TD comparison subjects were recruited from the same communities as patients via web-based advertisements and by posting flyers and brochures at local schools, pediatric clinics, and other community sites. Exclusion criteria for all study participants included significant neurological or medical conditions (unrelated to 22q11.2 mutation) that might affect brain structure, history of head injury with loss of consciousness, insufficient fluency in English, and/or substance or alcohol abuse or dependence within the past 6 months. The UCLA Institutional Review Board approved all study procedures and informed consent documents. Scanning was conducted on an identical 3 T Siemens Trio MRI scanner with a 12-channel

head coil at the University of California at Los Angeles Brain Mapping Center or at the Center for Cognitive Neuroscience (2). We excluded 25 individuals from the analysis due to the insufficient quality of the imaging data. The final 22q11.2 sample includes 22q11.2 BPA-BPD deletion (n=43) and BPA-BPD duplication (n=10) carriers and extrafamilial controls (n=43).

Montreal rare genomic disorder family dataset (MRG)

Cognitive and behavioral measures were collected in families with at least one child who carries a CNVs classified as pathogenic or VUS (variation of uncertain significance). CNV carriers and their first-degree relatives were ascertained neurodevelopmental disorder clinic at the Sainte Justine Montreal hospital, Quebec-Canada (MP-21-2016-946, 4165). The characteristics of this cohort reflect the criteria for Chromosomal microarray testing in the neurodevelopmental disorder clinic which include: intellectual disabilities, learning disabilities, autism spectrum disorder (ASD) as well as children with several comorbidities including attention deficit hyperactivity disorder (ADHD), speech and language disorders and developmental coordination disorders. The same assessments performed in first-degree relatives (carriers and non-carriers) allow adjusting for the effect of the additional genetic and environmental background. All the MRI scans have been performed at the Montreal Neurological Institute using the same scanner (Prisma 3T). The acquisition time for the resting-state sequence was 7 minutes. The neuroimaging protocol was designed by John D. Lewis and is available online: http://www.bic.mni.mcgill.ca/users/jlewis/BrainCanada/MCIN/

Data from this cohort include 16p11.2 deletion (n=7) and duplication (n=3) carriers, 1q21.1 deletion (n=5) and duplication (n=1) carriers, 15q11.2 (n=1) duplication carrier, 22q11.2 (n=1) duplication carrier, and intrafamilial controls (n=39) after exclusion of 16 individuals due to quality control criteria.

Cardiff DEFINE-dataset

The Cardiff CNV cohort was supported by the Wellcome Trust Strategic Award “DEFINE” and the National Centre for Mental Health with funds from Health and Care Research Wales. CNV carriers were clinically ascertained. MRI data were acquired on a 3 Tesla General Electric HDx MRI system (GE Medical Systems, Milwaukee, WI) using an eight-channel receive-only head RF coil (as described

in (3)). The acquisition time for functional resting-state data was 7 minutes. Following parameters have been used: Repetition Time = 2000 ms, Echo time = 35 ms, Slice Thickness = 3.4 mm (eyes open, fixation cross). The full protocol has been developed for the 100 brains project. 16p11.2 deletion (n=1) carrier, 1q21.1 deletion (n=3) and duplication (n=1) carriers, 15q11.2 (n=1) deletion carrier, 22q11.2 (n=1) duplication carrier, and extrafamilial controls (n=8) were included after the exclusion of 4 individuals due to quality control criteria.

Idiopathic autism dataset - ABIDE1

The ABIDE dataset (4,5) is an aggregate sample of different studies including imaging and behavioral data for individuals with an autism diagnosis and typically developing peers matched for age. Due to the small number of females in the ABIDE dataset, we excluded female individuals. To better account for biases in connectivity estimation due to differences in recording sites, subject age, and scanner motion, we created age and motion-matched subsamples for each recording site in ABIDE of individuals that passed our quality control criteria. We then excluded recording sites with fewer than 20 individuals (10 participants with autism, 10 controls). Our final ABIDE sample thus includes 943 male individuals, 472 individuals with autism, and 471 healthy controls, from 28 recording sites.

Idiopathic schizophrenia dataset

We used fMRI data retrospectively aggregated from 10 distinct sites and studies. Brain imaging multi-state data were obtained through either the SchizConnect and OpenfMRI data sharing platforms (http://schizconnect.org (6); https://openfmri.org (7)) or local scanning at the University of Montréal. All patients were diagnosed with schizophrenia (SZ) according to DSM-IV or DSM-V criteria, as a function of the time of the study. Sites samples were obtained after subjects were selected in order to ensure even proportions of SZ patients and controls within each site (from N = 9 to N = 42 per group) and to reduce between-group differences with regards to gender ratio (74% vs. 75% males in patients and controls, respectively), age distribution (34 vs. 32 years old on average) and motion levels (averaged frame displacement: 0.16 vs. 0.14 mm). Such matching of SZ and control subjects was achieved based on propensity scores. In total, we retained 242 SZ patients and 242 healthy controls in statistical

analyses. Depending on the study, positive and negative symptoms were assessed with either the Positive and negative syndrome scale (PANSS,(8)) or the Scales for the assessment of positive/negative symptoms (SAPS/SANS,(9)). In order to allow for group analyses, SAPS/SANS scores were converted into PANSS scores using published regression-based equations(10).

Idiopathic ADHD dataset - ADHD200

We used data provided by the ADHD-200 consortium and the neuro bureau ADHD-200 Preprocessed repository (8 cohorts http://fcon_1000.projects.nitrc.org/indi/adhd200/ (11)). Data from seven sites were retained after the exclusion of 447 individuals. We included in our study a total of 518 subjects, 187 patients diagnosed with ADHD and 331 healthy controls.

UCLA Consortium for Neuropsychiatric Phenomics - CNP

We downloaded T1-weighted Anatomical MPRAGE and resting-state fMRI BOLD data on the OpenfMRI platform (ds000030, https://www.openfmri.org/dataset/ds000030/) (12). Subjects were scanned across 2 sites (UCLA) - with a similar 3T Siemens Trio machine. We excluded 13 individuals after visual quality control. Data included in our study (n=237) encompassed healthy individuals (n=113) and individuals diagnosed with SZ (n=41), BIP (n=44), or ADHD (n=39). These diagnoses followed the Structured Clinical Interview for DSM-IV.

UK-BioBANK dataset

The UK biobank genomics data (13) were available for 31,225 individuals with MRI. After standard quality control procedure, we were able to add to our sample: 2q13 deletion (n=183) and duplication (n=88) carriers, 16p11.2 deletion (n=4) and duplication (n=6) carriers, 1q21.1 deletion (n=10) and duplication (n=13) carriers, 1q21.1 TAR duplication carriers (n=29), 15q11.2 deletion (n=103), and duplication (n=136) carriers, 15q13.3 duplication (n=190) carriers and non-carriers (n=30,185) from UK-Biobank.

CNVs calling procedure in UK-BioBANK dataset

The individual DNA (blood samples) were genotyped using the Affymetrix Axiom Array and UK BiLEVE Axiom array with common 733,256 markers mapped on the human genome version hg19. The UK Biobank provided normalized signal data of Log R Ratio and B Allele Frequency for each marker, which was formatted into standard input data suitable for the most common CNV calling software. We developed a CNV calling workflow with the ability to compute per individual CNV calling in parallel. The pipeline is compatible with Unix architectures and optimized for low-resource computers. We implemented CNV calling procedures for PennCNV(14) and QuantiSNP (15). Our protocol was executed on Compute Canada servers. The pipeline is available through GitHub (https://github.com/labjacquemont/MIND-GENESPARALLELCNV) and a QuantiSNP docker image version were also made available (16). We used default parameters for both algorithms to harmonize the CNV calling procedure (SNP coverage ≥ 3, likelihood score ≥ 15, and CNV length ≥1000 nt). Results were merged using CNVision script (n_{cnv}= 97,252) (17). We annotated CNV within the 9 loci (1q21.1, TAR, 2q13, 13q12.12, 15q11.2, 15q13.3, 16p11.2, 16p13.11, 22q11.2). Each detected CNV was visually checked with SnipPeep (http://snippeep.sourceforge.net/). We pooled these data with previous observations reported by Kendall et al 2017(18).

Selecting CNVs, conditions and traits.

We analyzed all of the available rsfMRI data for developmental psychiatric CNVs with at least n=20 carriers to allow for the detection of large effect sizes (Cohen's d > 0.8) previously reported for CNVs. As a result, selected CNVs are those most frequently identified in the clinic: 22q11.2, 1q21.1, 15q11.2, 16p11.2. We also selected all negative control CNVs (defined as CNVs without any reported association with psychiatric conditions or neurodevelopmental disorders) available for analysis with at least n=20 carriers. Some of these CNVs have mild effect sizes on cognitive ability (19) and may therefore mildly increase risk for psychiatric conditions.

Cognitive ability and neuroticism were selected because 1) all CNVs that increase risk for autism and/or SZ decrease cognitive ability (20,21) and 2) both traits show the highest genetic correlation, amongst commonly measured traits, with autism (22) as well as with schizophrenia (23,24).

Inflammatory bowel disease (IBD) was selected as a non-psychiatric control condition with a sample size similar to those available for the psychiatric conditions included in the study.

SNP Sample Size Determination for polygenic scores

Polygenic prediction via Bayesian regression and continuous shrinkage priors (PRS-CS) (25) requires a single SNP sample size as an input. Given that most of our discovery GWAS were actually meta-GWAS that were comprised of individual studies that varied in terms of their sample size and the SNPs they included, the sample size often varied considerably between SNPs. To account for this reality, we examined the distribution of SNP sample sizes in R and excluded SNPs that had sample sizes that were less than half of the maximum SNP sample size. Of the remaining SNPs, the median SNP sample size was used as the PRS-CS sample size input.

To ensure convergence of the underlying Gibbs sampler algorithm, we ran 25,000 Markov chain Monte Carlo iterations and designated the first 10,000 MCMC iterations as burn-in. The PRS-CS global shrinkage parameter was set to 0.01 when the discovery GWAS had a sample size that was less than 200,000; otherwise, it was learned from the data using a fully Bayesian approach. Default settings were used for all other PRS-CS parameters. The EUR posterior effects were fed into PLINK 1.9 (26) to produce raw PRS separately for the EUR and white British UKBB cohorts, and R [(27)] was used to standardize the PRS for each cohort to mean = 0 and SD = 1.

Standardized PRS were then adjusted by regressing out the first ten within-ancestry PCs.

eTable 1. Discovery GWAS used to compute polygenic scores with PRS-CS

Discovery GWAS used to compute polygenic scores with PRS-CS

Trait	Discovery GWAS	SNP Sample Size[a] (for PRS-CS)	SNP Count[b] (for PRS-CS)
BIP	Stahl et al. (2019)(28)	51,710	1,105,066
SZ	Ruderfer et al. (2018)(29)	87,491	1,079,332
Autism	Grove et al. (2019)(30)	46,350	1,094,054
IQ	Savagem J.E. et al (2018)(23)	266,670	1,071,016
Cross-D	Lee, P.H. et al. (2019)(31)	438,997	1,028,452
LDL	Willer et al. (2011)(32)	89,856	977,929
CKD	Wuttke et al. (2019)(33)	438,997	1,111,519

SZ, schizophrenia; LDL, low-density lipoprotein cholesterol; CKD, chronic kidney disease, BIP, Bipolar disorder, Cross-D: Cross-disorder. Of note, the discovery GWAS used for IQ, and CKD included individuals from the UKBB.

[a]PRS-CS requires a single SNP sample size.

[b]The "SNP count" is the number of SNPs in common between the discovery GWAS, the PRS-CS LD panel, and the genomic dataset; only these SNPs were used to calculate posterior effects.

Legend: PRS were computed using discovery GWAS for 3 psychiatric conditions (BIP, SZ, Autism), cross-disorder, IQ, and 2 non-brain control traits (Low-Density Lipoprotein; Chronic Kidney Disease). We used the most recent GWAS to compute all PRS. There were instances we intentionally used an older GWAS to avoid using a GWAS that included UKBB.

Diagnostic procedure

For autism, we used the ABIDE 1 and 2 datasets, which has previously been described in full detail (4,5): "*Sites reached ASD diagnoses by either (1) combining clinical judgment and 'gold standard' diagnostic instruments—Autism Diagnostic Observation Schedule (ADOS) and/or Autism Diagnostic Interview–Revised, clinical judgment only or 'gold standard' diagnostic instruments only. Among the 17 sites using the ADOS and/or Autism Diagnostic Interview-Revised, 16 obtained research-reliable administrations and scorings. Site-specific details are available at http://fcon_1000.projects.nitrc.org/indi/abide. Given participant ages (6 years) and IQ, most were evaluated with ADOS Modules 3 or 4. Average ADOS total scores were similar across sites, suggesting consistency in ASD severity. Calibrated severity scores (computed using the new ADOS algorithm for Modules 1, 2 and 3) were available for nine sites and confirmed this pattern.*"

The schizophrenia dataset is an aggregate from 11 distinct sites (12,34). Brain imaging multi-state data were obtained through either the SchizConnect (6) and OpenfMRI (7) data sharing platforms (http://schizconnect.org ; https://openfmri.org) or local scanning at the University of Montréal. All patients were diagnosed with schizophrenia (SZ) according to DSM-IV or DSM-V criteria, as a function of the time of the study. Depending on the study, positive and negative symptoms were assessed with either the Positive and negative syndrome scale (PANSS) or the Scales for the assessment of positive/negative symptoms (SAPS/SANS).

For ADHD, we used the data from 5 sites (ADHD-200): We provided a summary of the diagnostic assessment but more details are available at:

http://fcon_1000.projects.nitrc.org/indi/adhd200/index.html

"At the Kennedy Krieger Institute, *psychiatric diagnoses were based on evaluations with the Diagnostic Interview for Children and Adolescents, Fourth Edition (DICA-IV, 1997), a structured parent interview based on DSM-IV criteria; the Conners' Parent Rating Scale-Revised, Long Form (CPRS-R), and the DuPaul ADHD Rating Scale-IV(Reid, 1998). Intelligence was evaluated with the*

Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV) and academic achievement was assessed with the Wechsler Individual Achievement Test-II. Children assigned to the ADHD group met criteria for ADHD on the DICA-IV and either had a T-score of 65 or greater on the CPRS-R Long Form (DSM-IV Inattentive) and/or M (DSM-IV Hyperactive/Impulsive) or met criteria on the DuPaul ADHD Rating Scale IV (six out of nine items scored 2 or 3 from Inattention items and/or six out of nine scored 2 or 3 from the Hyperactivity/Impulsivity items).

*For the **NeuroIMAGE study** (Radboud University Nijmegen Medical Centre, Vrije Universiteit Amsterdam, UMC Groningen), Psychiatric diagnoses were based on evaluations with the Schedule of Affective Disorders and Schizophrenia for Children—Present and Lifetime Version (KSADS-PL) administered to parents and children and the Conners' Parent Rating Scale-Revised, Long version (CPRS-LV). Intelligence was evaluated with the Wechsler Abbreviated Scale of Intelligence (WASI). Inclusion in the ADHD group required a diagnosis of ADHD based on parent and child responses to the KSADS-PL. Inclusion criteria for TDC required the absence of any Axis-I psychiatric diagnoses per parent and child KSADS-PL interview, as well as T-scores below 60 for all the CPRS-R: LV ADHD summary scales.*

*For the **NYU Child Study Center,** psychiatric diagnoses were based on evaluations with the Schedule of Affective Disorders and Schizophrenia for Children—Present and Lifetime Version (KSADS-PL) administered to parents and children and the Conners' Parent Rating Scale-Revised, Long version (CPRS-LV). Intelligence was evaluated with the Wechsler Abbreviated Scale of Intelligence (WASI). Inclusion in the ADHD group required a diagnosis of ADHD based on parent and child responses to the KSADS-PL as well as on a T-score greater than or equal to 65 on at least one ADHD-related index of the CPRS-R: LV. Inclusion criteria for TDC required the absence of any Axis-I psychiatric diagnoses per parent and child KSADS-PL interview, as well as T-scores below 60 for all the CPRS-R: LV ADHD summary scales.*

*For the **Oregon Health Science University site,** psychiatric diagnoses were based on evaluations with the Kiddie Schedule for Affective Disorders and Schizophrenia (KSADS-I) administered to a parent; parent and teacher Connors' Rating Scale-3rd Edition; and a clinical review by a child psychiatrist and neuropsychologist who had to agree on the diagnosis. Intelligence was evaluated with a three-*

subtest short form (Block Design, Vocabulary, and Information) of the Wechsler Intelligence Scale for Children, Fourth Edition.

*For the **Peking University site**, all participants (ADHD and TDC) were evaluated with the Schedule of Affective Disorders and Schizophrenia for Children—Present and Lifetime Version (KSADS-PL) with one parent for the establishment of the diagnosis for study inclusion. The ADHD Rating Scale (ADHD-RS) IV was employed to provide dimensional measures of ADHD symptoms. Full-scale Wechsler Intelligence Scale for Chinese Children-Revised (WISCC-R) was performed."*

For the UCLA Consortium for Neuropsychiatric Phenomics (12):

"Diagnoses for ADHD, bipolar disorder and SZ followed the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition—Text Revision, and were based on the Structured Clinical Interview for DSM-IV (SCID-I) supplemented by the Adult ADHD Interview (a structured interview form derived from the Kiddie Schedule for Affective Disorders and Schizophrenia, Present and Lifetime Version (KSADS-PL)), in order to enable a more detailed characterization of lifetime history of ADHD in adults. Interviewers/raters were trained to criteria as described elsewhere, in brief minimum standards of acceptable symptom agreement were overall kappa of .75, a kappa specificity of .75 and sensitivity of .75, and .85 kappa for diagnostic accuracy. Diagnostic and Symptom elicitation skill was also assessed with the SCID Checklist of Interviewer Behaviors and the Symptom Checklist of Interviewer Behaviors. Ongoing quality assurance checks documented kappa above .75 for each rater annually during the course of the study."

Resting-state fMRI Preprocessing

All datasets were preprocessed using the same parameters with the same Neuroimaging Analysis Kit (NIAK) version 0.12.4, an Octave-based open-source processing and analysis pipeline (35). The first four volumes of each rs-fMRI time series were discarded to allow for magnetization to reach a steady state. Each data set was corrected for differences in slice acquisition time. Head motion parameters were estimated by spatially re-aligning individual timepoints with the median volume in the time series. This

reference median volume was then aligned with the individual anatomical T1 image, which in turn was co-registered onto the MNI152 template space using an initial affine transformation, followed by a nonlinear transformation. Finally, each individual time point was mapped to the MNI space (36) using the combined spatial transformations. Slow frequency drifts were modeled on the entire time series as discrete cosine basis functions with a 0.01 Hz high-pass cut-off. Timepoints with excessive in-scanner motion (greater than 0.5 mm framewise displacement) were then censored from the time series by removing the affected timepoint as well as the preceding and following two-time points (37). Nuisance covariates were regressed from the remaining time series: the previously estimated slow time drifts, the average signals in conservative masks of the white matter and lateral ventricles, and the first principal components (95% energy) of the estimated six rigid-body motion parameters and their squares. Data were then spatially smoothed with a 3D Gaussian kernel (FWHM = 6mm).

Quality Control - preprocessed rs-fMRI data

Preprocessed data were visually controlled for the quality of co-registration, head motion, and related artifacts by three raters. Not all datasets were examined by the same raters, yet all raters followed the same standardized quality-control procedure (38). Individuals were excluded from the analysis if co-registration errors could not be fixed. Individuals were also excluded from the analysis if the average framewise displacement after motion censoring exceeded 0.5 mm or if fewer than 40-time frames remained.

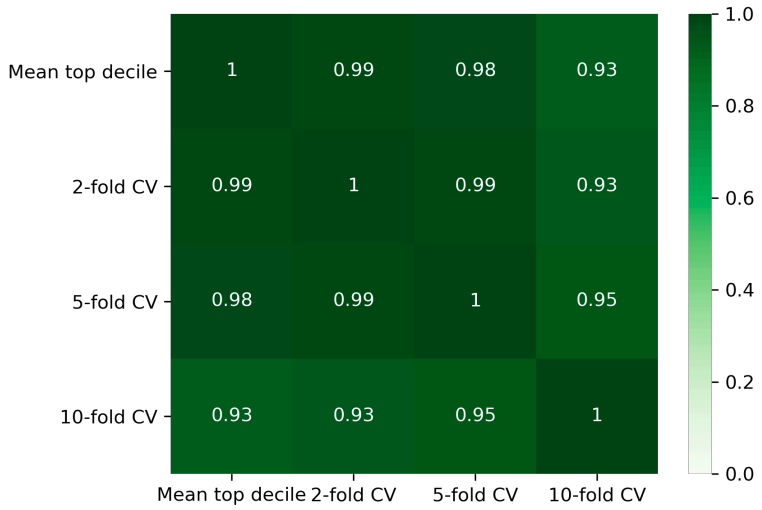
eTable2: Additional information on motion and scanning sites

CNV (hg19)	n (genes) / *Gene*	**Status**	**Motion**	Frames	**Sites**	Cohorts
15q11.2 15: 22.81-23.09	4 *CYFIP1*	DEL	0.19 (0.06)	389 (118)	3	UKBB
		DUP	0.19 (0.05)	393 (118)		
15q13.3 15: 31.08-32.46	5 *CHRNA7*	DUP	0.19 (0.05)	375 (127)		
2q13 2: 110.86-110.98	3 *NPHP1*	DEL	0.19 (0.05)	370 (123)		
		DUP	0.19 (0.05)	399 (98)		
16p13.11 16: 15.51-16.29	6 *MYH11*	DUP	0.18(0.04)	298 (101)		
13q12.12 13: 23.56-24.88	5 *SPATA13*	DEL	0.20(0.07)	391 (111)		
		DUP	0.20(0.06)	361 (125)		
TAR 1: 145.39-147.39	15 *RBM8*	DUP	0.17(0.05)	410 (95)		
1q21.1 1: 146.53-147.39	7 *CHD1L*	DEL	0.18 (0.07)	271 (172)	6	UKBB-MRG-Cardiff-SFARI
		DUP	0.21 (0.08)	276 (188))	7	
22q11.2 22: 19.04-21.47	49 *TBX1*	DEL	0.18 (0.07)	120 (36)	1	UCLA
		DUP	0.19 (0.09)	225 (163)	5	UCLA-UKBB Cardiff-MRG
16p11.2 16: 29.65-30.20	27 *KCTD13*	DEL	0.22 (0.09)	108 (99)	5	SFARI - MRG -UKBB
		DUP	0.21 (0.09)	148 (140)	4	
Idiopathic Psychiatric Conditions		**SZ**	0.17 (0.06)	147 (50)	12	Montreal-SZ CNP
		BIP	0.17(0.07)	127 (25)	2	CNP
		ASD	0.16 (0.05)	157 (118)	28	ABIDE1 ABIDE2
		ADHD	0.15 (0.04)	135 (23)	8	ADHD-200 CNP
Non psychiatric condition		IBD	0.19 (0.06)	371 (123)	3	UKBB
Polygenic scores	BIP, Cross-Disorder, LDL, CKD		0.18 (0.05)	388 (116)	3	UKBB
	ASD					
	SZ					
	IQ					
Controls	UKBB		0.19 (0.05)	387 (116)	3	UKBB
	SFARI		0.18 (0.07)	81.4 (19)	2	SFARI
	MRG		0.21 (0.07)	381 (177)	1	MRG
	Cardiff		0.12 (0.06)	151.4 (35)	1	Cardiff
	UCLA		0.14 (0.04)	130 (24)	1	UCLA
	Psychiatric cohorts		0.15 (0.05)	163 (100)	44	-

Legend: CNV carriers, individuals with idiopathic psychiatric conditions, and controls after MRI quality control. Chr: chromosome number, coordinates are presented in Megabases (Mb, Hg19). Motion: framewise displacement (in mm). Quantitative variables are expressed as the mean ± standard

deviation. All site scanned controls and sensitivity analyses were performed to investigate the potential bias introduced by differences in scanning site, age, and sex.

eFigure 1 Original effect sizes observations versus effect size obtained by cross-validation



Legend: Correlation between original effect sizes observations (Mean of the top decile of connections) and those computed using a 2-, 5-, and 10-fold cross-validation. Values in the boxes are Pearson r.

eTable 3. Psychiatric diagnoses and FSIQ

CNV	Status	n clin	FSIQ	ASD	ADHD	SZ	BIP
1q21.1	DEL	15	97 (17)	2	2	0	0
	DUP	6	88 (29)	2	2	0	0
22q11.2	DEL	43	77 (14)	23	19	3	0
	DUP	12	96 (20)	4	5	0	0
16p11.2	DEL	28	87 (15)	7	7	0	0
	DUP	29	90 (21)	3	3	0	0
Idiopathic Psychiatric Conditions	**SZ**	283	-	-	-	242	-
	BIP	44	-	-	-	-	44
	ASD	225	104 (17)	225	-	-	-
	ADHD	226	107 (14)	-	289	-	-

Legend: Diagnoses information and full-scale intelligence quotient (FSIQ, mean (standard deviation))

for the clinically ascertained CNV carriers, and psychiatric conditions. n clin: number of participants clinically ascertained; DEL: deletion; DUP: duplication; SZ: schizophrenia, ASD: Autism Spectrum Disorder; ADHD: Attention-Deficit/Hyperactivity-Disorder, BIP: Bipolar disorder.

eTable 4. Psychiatric diagnoses in UK-Biobank

CNV	Status	F17.1 Harmful use	F20-F29 SZ	F30.0 Hypomania	F32 Depressive episode	F41.9 Anxiety disorder	F42 OCD	F03 Dementia	F84.5 Asperger's syndrome	F90.0 ADHD	F99 Mental disorder unspecified
15q11.2	DEL	2	0	0	2	0	0	0	0	0	0
	DUP	0	0	0	2	3	0	1	0	0	0
15q13.3	DUP	0	0	1	1	0	0	0	0	0	0
2q13	DEL	1	0	0	0	3	0	0	0	0	0
	DUP	1	0	0	0	1	0	0	0	0	0
Controls	CON	69	6	1	97	44	2	0	1	0	0
1q21.1	DEL	0	0	0	0	0	0	0	0	0	0
	DUP	0	0	0	1	0	0	0	0	1	1
	DUP-TAR	0	0	0	1	1	0	0	0	0	0
22q11.2	DUP	0	0	0	1	0	0	0	0	0	0
16p11.2	DEL	0	0	0	1	0	0	0	0	0	0
	DUP	0	0	0	0	0	0	0	0	0	0

Legend: Diagnosis information for the CNV carriers identified in the UK Biobank. More information about diagnostic codes is available on UK Biobank website.

Supplementary Results

eTable 5. Connectome wide association studies without global signal adjustment

Metrics		No global signal adjustment					
		Connections		Beta values		mean	*pvalue* shift
		pos	*neg*	*min*	*max*		
CNV Deletion	1q21.1	0	1	-1.08	0.61	-0.02	ns
	2q13	0	0	-0.11	0.22	0.04	ns
	13q12.12	0	0	-0.52	0.53	0.03	ns
	15q11.2	0	0	-0.26	0.37	0.04	ns
	16p11.2	183	9	-0.84	1.82	0.18	0.04
	22q11.2	0	38	-1.6	0.80	-0.25	0.04
CNV Duplication	TAR	0	0	-0.54	0.42	-0.08	ns
	1q21.1	102	0	-0.39	1.12	0.30	0.008
	2q13	0	0	-0.37	0.24	-0.03	ns
	13q12.12	0	0	-0.31	0.79	0.31	0.009
	15q11.2	0	0	-0.28	0.18	-0.06	ns
	15q13.3	0	0	-0.2	0.18	0.007	ns
	16p11.2	0	3	-1.09	0.48	-0.07	ns
	16p13.11	0	0	-0.55	0.33	-0.16	0.03
	22q11.2	0	0	-0.55	0.93	0.24	0.03
Polygenic score	Cross Dis	1	0	-0.02	0.03	-0.002	ns
	ASD	0	0	-0.02	0.02	0	ns
	Schizophrenia	1	25	-0.03	0.02	-0.005	ns
	BIP	9	0	-0.02	0.03	0.004	ns
	IQ	63	2	-0.02	0.03	0.004	ns
	LDL	0	0	-0.02	0.01	-0.003	ns
	CKD	0	0	-0.02	0.02	0	ns
Psychiatric conditions	ASD	5	23	-0.31	0.27	-0.03	ns
	Schizophrenia	20	1316	-0.64	0.31	-0.22	<2e-4
	Bipolar	0	58	-0.77	0.49	-0.17	0.08
	ADHD	0	0	-0.19	0.25	0.03	0.21
	IBD	0	0	-0.19	0.14	-0.03	ns
Cognitive scores	Fluid Intel	932	39	-0.03	0.05	0.01	0.0004
	Neuroticism	23	782	-0.04	0.03	-0.01	0.002

Legend: The number of significantly altered connections (FDR corrected) for each connectome-wide association study (n=29) before Global signal adjustment (GSA).
min-max: minimum-maximum of z-scored beta values; top decile: top decile of beta values; Connection pos: number of positive connections surviving FDR; Connection neg: number of negative connections surviving. Abbreviations: DEL: deletion; DUP: duplication; SZ: schizophrenia, ASD: Autism Spectrum Disorder; ADHD: Attention-Deficit / Hyperactivity-Disorder, BIP: Bipolar disorder, CrossD: Cross-disorder, LDL: Low-Density Lipoprotein, CKD: Chronic Kidney Disease; IQ: intelligence quotient.

All the *beta*-values, *p*-values, and *q*-values are available on Github for the 29 FC-profiles:
https://github.com/claramoreau9/NeuropsychiatricCNVs_Connectivity/tree/master/results_tables
Interactive representations on brain maps (39) are also available, with and without GSA:
https://claramoreau9.github.io/Braimaps_Github.html

CNVs sensitivity analyses

16p11.2 CNVs

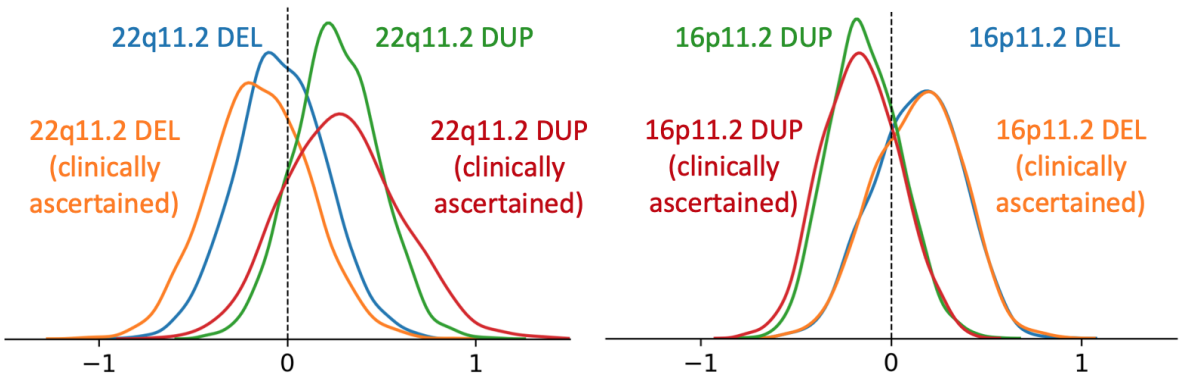
FC profiles obtained before GSA were correlated with our previous study (r=0.70, for 16p11.2 deletion, r=0.83, CI=[0.82; 0.84] for 16p11.2 duplication) (40). We performed a sensitivity analysis only including clinically ascertained 16p11.2 carriers (MRG, Define Cardiff, Lausanne, and SVIP cohorts, eFigure 2). FC-profiles were strongly correlated to our previously published findings (r=0.79, CI=[0.77; 0.8] for 16p11.2 deletion, r=0.87, CI=[0.86; 0.88] for 16p11.2 duplication) (40).

22q11.2 CNVs

The GSA FC profile was correlated with the previously published FC profile without GSA (40) r=0.86 (CI=[0.85; 0.87]).
We performed a sensitivity analysis for 22q11.2 duplication, showing that FC-profiles using the full dataset and the FC-profiles excluding the non-clinically ascertained participants (UK-Biobank scanning sites). This new FC profile computed using UCLA, MRG, and Define Cardiff data, was highly

correlated to previous findings (r=0.87, eFigure 2 (40)). We performed an additional sensitivity analysis matching controls for scanning site, age, sex, and motion (n=246) using the "matchControls" function in the {e1071} R package (41). FC-profiles before and after matching were strongly correlated (r=0.96, CI=[0.95;0.97]).

eFigure 2. Mean connectivity shifts in clinically and non-clinically ascertained 16p11.2 and 22q11.2 CNVs carriers

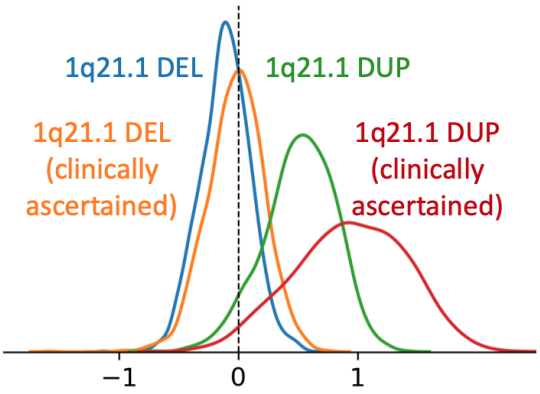


Legend: Density plots represent the FC profile distribution (2080 beta estimates from whole-brain contrast of cases versus controls) for the 22q11.2 CNVs (left) and 16p11.2 CNVs (right) groups, using all the available subjects (blue and green, for deletion and duplication respectively) or only the clinically ascertained CNVs carriers (orange and red).

1q21.1 CNVs

Sensitivity analyses showed that FC-profiles obtained using all sites and only those including clinically ascertained carriers were correlated (r=0.75 for 1q21.1 deletion, and r=0.78 for 1q21.1 duplication, eFigure 3).

eFigure 3. Mean connectivity shifts in clinically and non-clinically ascertained 1q21.1 CNVs carriers



Legend. Density plots represent the distribution of 2080 beta estimates for the connectome wide association study (whole brain contrast of cases versus controls) for 1q21.1 CNVs, using all the available subjects (blue and green, for deletion and duplication respectively) or only the clinically ascertained CNVs carriers (orange and red).

Sensitivity analyses of age-related effect on FC

Keeping in mind our limited power, we performed sensitivity analyses by stratifying the 16p11.2 deletion, PRS-SZ, and PRS-IQ into 2 age groups. For the 16p11.2 deletion, the corresponding FC-profiles of both sub-samples (above 12 years: n=17 and under 12 years old: n=14) were correlated with the FC-profile reported in the main text computed with the entire sample (r=0.81 and 0.85 respectively). For PRS-SZ, the FC-profiles of both subsamples (below 65 years old, n=15236 and above 65 years, n=14404) were correlated with the FC-profile reported in the manuscript and computed with the entire sample (r=0.84 and 0.9 respectively). The same observation was reported for PRS-IQ (r=0.82 and r=0.89 respectively). We also added age^2 as a confounding variable in all three CWAS (16p11.2, PRS-SZ and PRS-IQ). All three FC profiles were highly correlated with the initial CWAS profile reported in the manuscript (ranging from r=0.93 to 0.99).

Sensitivity analyses excluding female participants in schizophrenia dataset

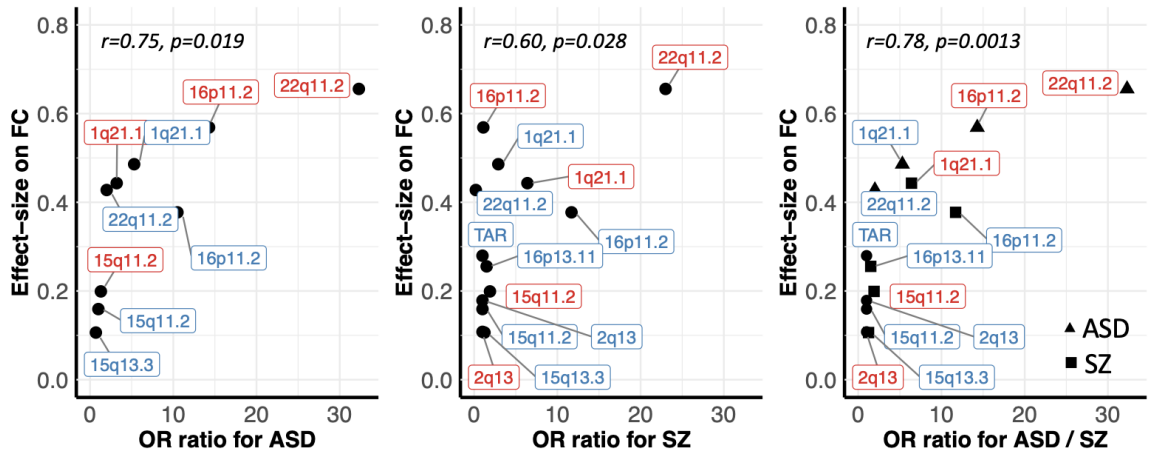
A sensitivity analysis was performed after excluding females in the schizophrenia cases control CWAS. In the all male sample, the CWAS of idiopathic SZ showed 309 connections surviving FDR. The beta map of this subsample was highly correlated (r= 0.95) with the initial CWAS performed on the full sample.

Sensitivity analyses excluding patients with autism and ADHD with medication

A sensitivity analysis was performed on 122 participants with autism from ABIDE1 and the respective controls excluding subjects with medication at the MRI scan time and subjects without information about pharmacological treatment. The beta map of this new connectome wide association analysis was correlated (r= 0.81) with the initial connectome wide association study performed on the full sample including subjects with medication. The same analysis was performed for ADHD using “Medication naive versus Not medication naive” information provided by the ADHD-200 consortium. We obtained

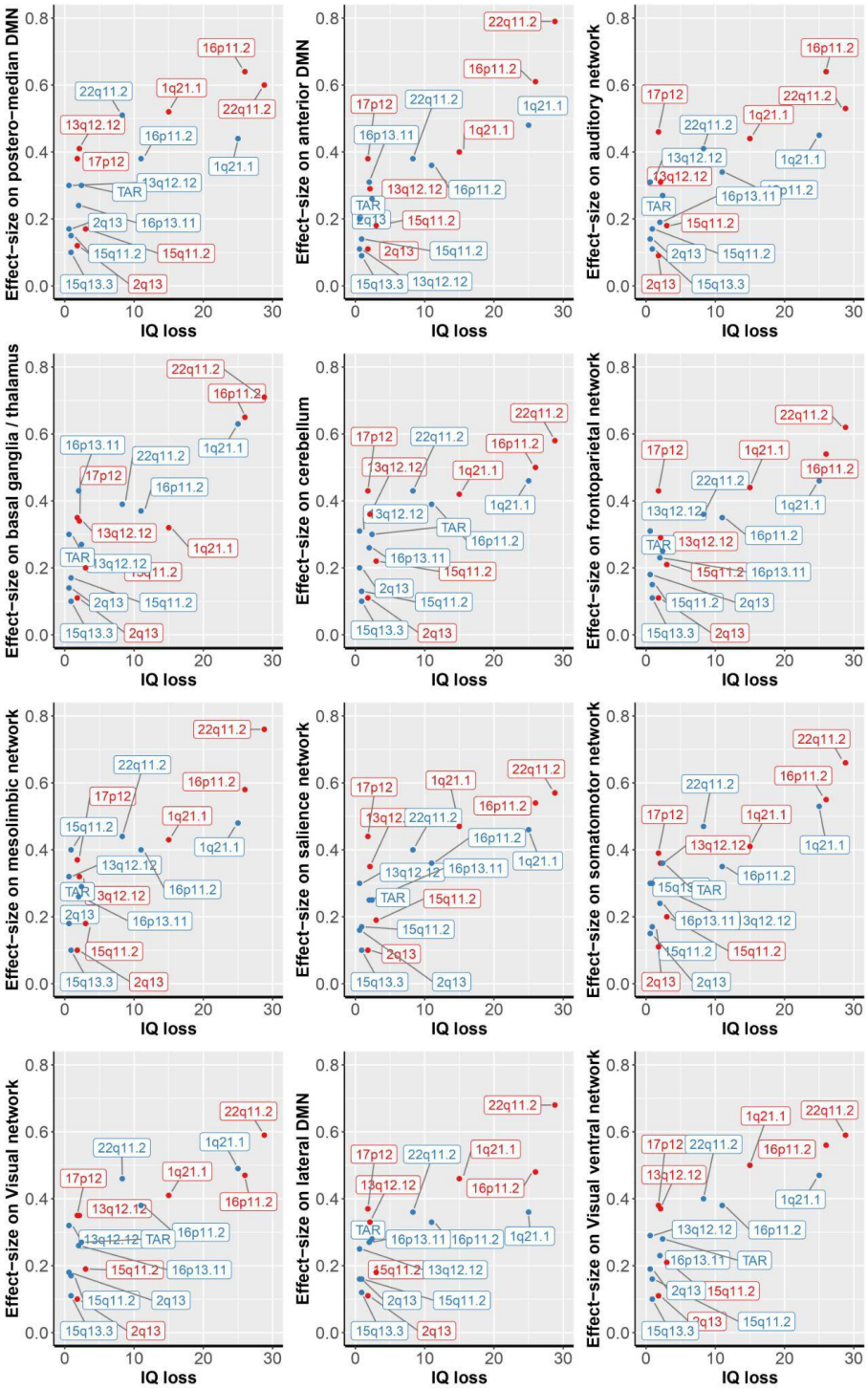
the same result (no connection survived FDR). We did not have individual medication information for the SZ cohorts.

eFigure 4. Correlation between sizes of CNVs on risk for autism and SZ and their effect sizes on FC



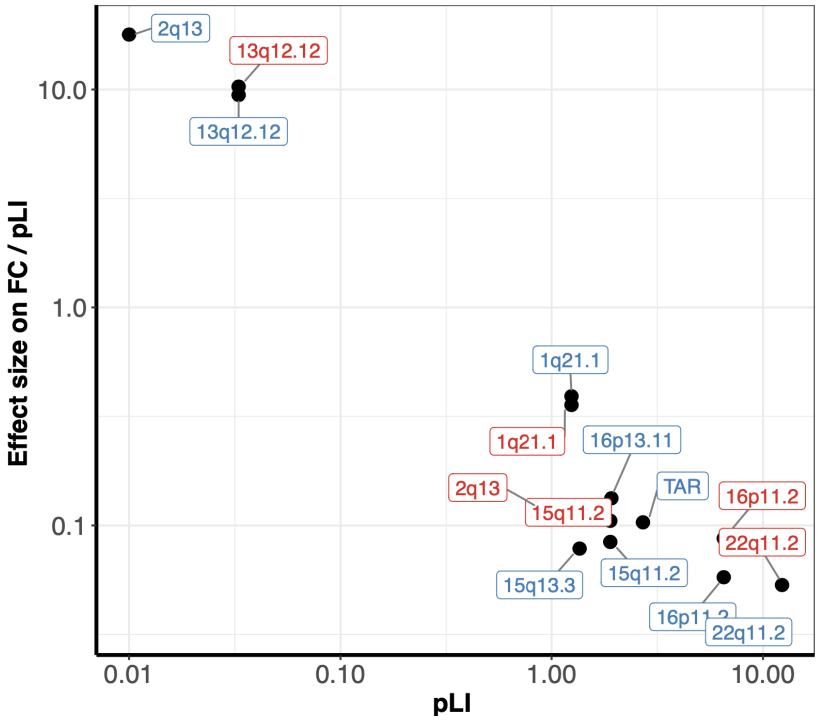
*Legend : (left) Correlation between previously published effect sizes of CNVs on autism risk and their effect sizes on FC. (middle) Correlation between previously published effect sizes of CNVs on SZ risk and their effect sizes on FC. (right) Correlation between previously published effect sizes of CNVs on autism or SZ risk (17,42–44) and their effect sizes on FC. X-axis: Odd ratios for either autism or SZ (*the highest risk conferred by each CNV for either *autism* (17,44) or SZ (42,43)). *Y-axis: effect sizes on FC (top decile of estimates).*

eFigure 5. Correlation between effect sizes of CNVs on IQ and on FC for 12 functional networks



Legend: Correlation between previously published effect sizes of CNVs on IQ (16) (X-axis) and their effect sizes on FC of 12 functional networks (using the mean of the top decile of network-wide estimates)(Y-axis).

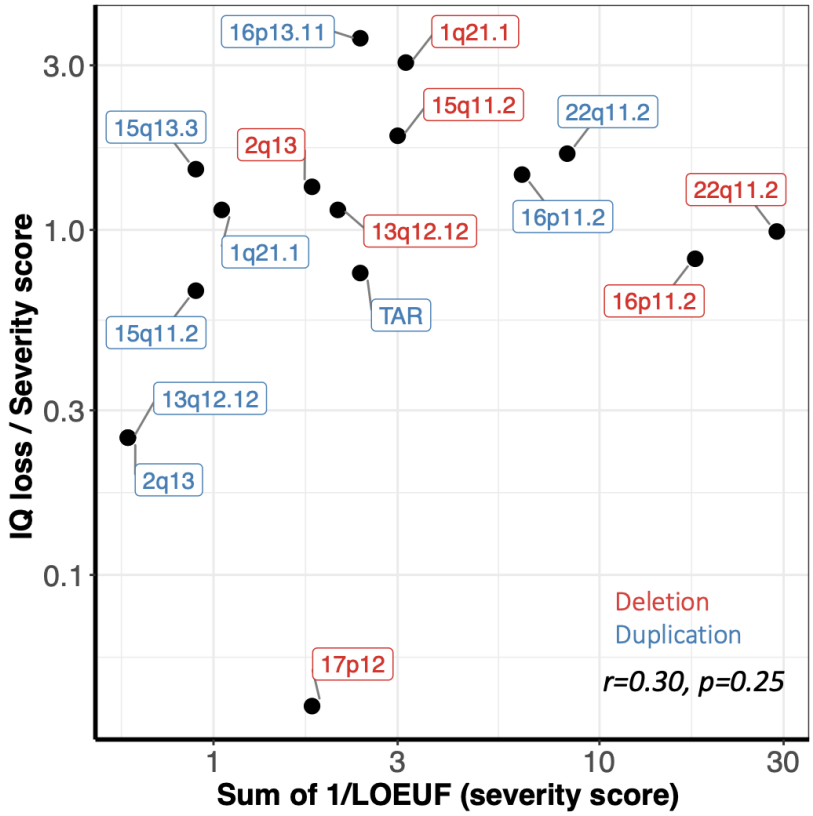
eFigure 6. Adjusted effect sizes on FC and multigenicity (severity score)



Y-axis: mean effect on FC of one point of pLI (CNV effect sizes adjusted by the severity score).

X-axis: Severity score measured by pLI for each CNV.

eFigure 7. Concordance analysis between the CNVs effect-size on IQ and the effect sizes predicted by the linear effect of the severity score



Legend: Relationship between CNV effect sizes on IQ normalized by the severity score (y-axis), and the severity of the CNV (x-axis, measured by the sum of 1/LOEUF).

Supplementary References

1. Simons Vip Consortium (2012): Simons Variation in Individuals Project (Simons VIP): a genetics-first approach to studying autism spectrum and related neurodevelopmental disorders. *Neuron* 73: 1063–1067.
2. Lin A, Ching CRK, Vajdi A, Sun D, Jonas RK, Jalbrzikowski M, *et al.* (2017): Mapping 22q11.2 Gene Dosage Effects on Brain Morphometry. *J Neurosci* 37: 6183–6199.
3. Drakesmith M, Parker GD, Smith J, Linden SC, Rees E, Williams N, *et al.* (2019): Genetic risk for schizophrenia and developmental delay is associated with shape and microstructure of midline white-matter structures. *Transl Psychiatry* 9: 102.
4. Di Martino A, Yan C-G, Li Q, Denio E, Castellanos FX, Alaerts K, *et al.* (2014): The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol Psychiatry* 19: 659–667.
5. Di Martino A, O'Connor D, Chen B, Alaerts K, Anderson JS, Assaf M, *et al.* (2017): Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. *Sci Data* 4: 170010.
6. Wang L, Alpert KI, Calhoun VD, Cobia DJ, Keator DB, King MD, *et al.* (2016): SchizConnect: Mediating neuroimaging databases on schizophrenia and related disorders for large-scale integration. *Neuroimage* 124: 1155–1167.
7. Poldrack RA, Barch DM, Mitchell JP, Wager TD, Wagner AD, Devlin JT, *et al.* (2013): Toward open sharing of task-based fMRI data: the OpenfMRI project. *Front Neuroinform* 7: 12.
8. Kay SR, Fiszbein A, Opler LA (1987): The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull* 13: 261–276.
9. Andreasen NC (1989): The Scale for the Assessment of Negative Symptoms (SANS): conceptual and theoretical foundations. *Br J Psychiatry Suppl* 49–58.
10. van Erp TGM, Preda A, Nguyen D, Faziola L, Turner J, Bustillo J, *et al.* (2014): Converting positive and negative symptom scores between PANSS and SAPS/SANS. *Schizophr Res* 152: 289–294.
11. ADHD-200 Consortium (2012): The ADHD-200 Consortium: A Model to Advance the Translational Potential of Neuroimaging in Clinical Neuroscience. *Front Syst Neurosci* 6: 62.
12. Poldrack RA, Congdon E, Triplett W, Gorgolewski KJ, Karlsgodt KH, Mumford JA, *et al.* (2016): A phenome-wide examination of neural and cognitive function. *Sci Data* 3: 160110.
13. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, *et al.* (2015): 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.

14. Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SFA, *et al.* (2007): PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res* 17: 1665–1674.
15. Colella S, Yau C, Taylor JM, Mirza G, Butler H, Clouston P, *et al.* (2007): QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids Res* 35: 2013–2025.
16. Huguet G, Schramm C, Douard E, Jiang L, Labbe A, Tihy F, *et al.* (2018): Measuring and Estimating the Effect Sizes of Copy Number Variants on General Intelligence in Community-Based Samples. *JAMA Psychiatry*. https://doi.org/10.1001/jamapsychiatry.2018.0039
17. Sanders SJ, He X, Willsey AJ, Ercan-Sencicek AG, Samocha KE, Cicek AE, *et al.* (2015): Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron* 87: 1215–1233.
18. Kendall KM, Rees E, Escott-Price V, Einon M, Thomas R, Hewitt J, *et al.* (2017): Cognitive Performance Among Carriers of Pathogenic Copy Number Variants: Analysis of 152,000 UK Biobank Subjects. *Biol Psychiatry* 82: 103–110.
19. Huguet G, Schramm C, Douard E, Tamer P, Main A, Monin P, *et al.* (2021): Genome-wide analysis of gene dosage in 24,092 individuals estimates that 10,000 genes modulate cognitive ability. *Mol Psychiatry*. https://doi.org/10.1038/s41380-020-00985-z
20. Stefansson H, Meyer-Lindenberg A, Steinberg S, Magnusdottir B, Morgen K, Arnarsdottir S, *et al.* (2014): CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature* 505: 361–366.
21. Douard E, Zeribi A, Schramm C, Tamer P, Loum MA, Nowak S, *et al.* (2021): Effect Sizes of Deletions and Duplications on Autism Risk Across the Genome. *Am J Psychiatry* 178: 87–98.
22. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, *et al.* (2019): Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet* 51: 431–444.
23. Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, de Leeuw CA, *et al.* (2018): Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat Genet* 50: 912–919.
24. Nagel M, Jansen PR, Stringer S, Watanabe K, de Leeuw CA, Bryois J, *et al.* (2018): Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and

pathways. *Nat Genet* 50: 920–927.

25. Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW (2019): Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun* 10: 1776.
26. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ (2015): Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4. https://doi.org/10.1186/s13742-015-0047-8
27. Team RC (2019): R: A Language and Environment for Statistical Computing (Version 3.5. 2, R Foundation for Statistical Computing, Vienna, Austria, 2018). *There is no corresponding record for this reference*. Retrieved from https://scholar.google.ca/scholar?cluster=7867649181275189008,12785896441631966750,17433060647574105173&hl=en&as_sdt=0,5&sciodt=0,5
28. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, *et al.* (2019): Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet* 51: 793–803.
29. Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics Consortium. Electronic address: douglas.ruderfer@vanderbilt.edu, Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics Consortium (2018): Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell* 173: 1705–1715.e16.
30. Grasby KL, Jahanshad N, Painter JN, Colodro-Conde L, Bralten J, Hibar DP, *et al.* (2020): The genetic architecture of the human cerebral cortex. *Science* 367. https://doi.org/10.1126/science.aay6690
31. Lee PH, Anttila V, Won H, Feng Y-CA, Rosenthal J, Zhu Z, *et al.* (2019): Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* 179: 1469–1482.e11.
32. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, *et al.* (2013): Discovery and refinement of loci associated with lipid levels. *Nat Genet* 45: 1274–1283.
33. Wuttke M, Li Y, Li M, Sieber KB, Feitosa MF, Gorski M, *et al.* (2019): A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat Genet* 51: 957–972.
34. Orban P, Desseilles M, Mendrek A, Bourque J, Bellec P, Stip E (2017): Altered brain connectivity in patients with schizophrenia is consistent across cognitive contexts. *J Psychiatry Neurosci* 42: 17–26.
35. Bellec P, Carbonell FM, Perlbarg V, Lepage C, Lyttelton O, Fonov V, *et al.* (2011): A neuroimaging analysis kit for Matlab and Octave. *Proceedings of the 17th International Conference on Functional Mapping of the Human Brain* 2735–2746.
36. Fonov VS, Evans AC, McKinstry RC, Almli CR, Collins DL (2009): Unbiased nonlinear average age-

appropriate brain templates from birth to adulthood. *Neuroimage* 47: S102.

37. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012): Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59: 2142–2154.
38. Benhajali Y, Bellec P (2016, November 3): Quality Control and assessment of the NIAK functional MRI preprocessing pipeline. https://doi.org/10.6084/m9.figshare.4204845.v1
39. Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, *et al.* (2014): Machine learning for neuroimaging with scikit-learn. *Front Neuroinform* 8: 14.
40. Moreau CA, Urchs SGW, Kuldeep K, Orban P, Schramm C, Dumas G, *et al.* (2020): Mutations associated with neuropsychiatric conditions delineate functional brain connectivity dimensions contributing to autism and schizophrenia. *Nat Commun* 11: 1–12.
41. e1071.pdf (n.d.): Retrieved from https://cran.r-project.org/web/packages/e1071/e1071.pdf
42. Marshall CR, Howrigan DP, Merico D, Thiruvahindrapuram B, Wu W, Greer DS, *et al.* (2017): Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat Genet* 49: 27–35.
43. Kirov G, Rees E, Walters JTR, Escott-Price V, Georgieva L, Richards AL, *et al.* (2014): The penetrance of copy number variations for schizophrenia and developmental delay. *Biol Psychiatry* 75: 378–385.
44. Moreno-De-Luca D, Sanders SJ, Willsey AJ, Mulle JG, Lowe JK, Geschwind DH, *et al.* (2013): Using large clinical data sets to infer pathogenicity for rare copy number variants in autism cohorts. *Mol Psychiatry* 18: 1090–1095.