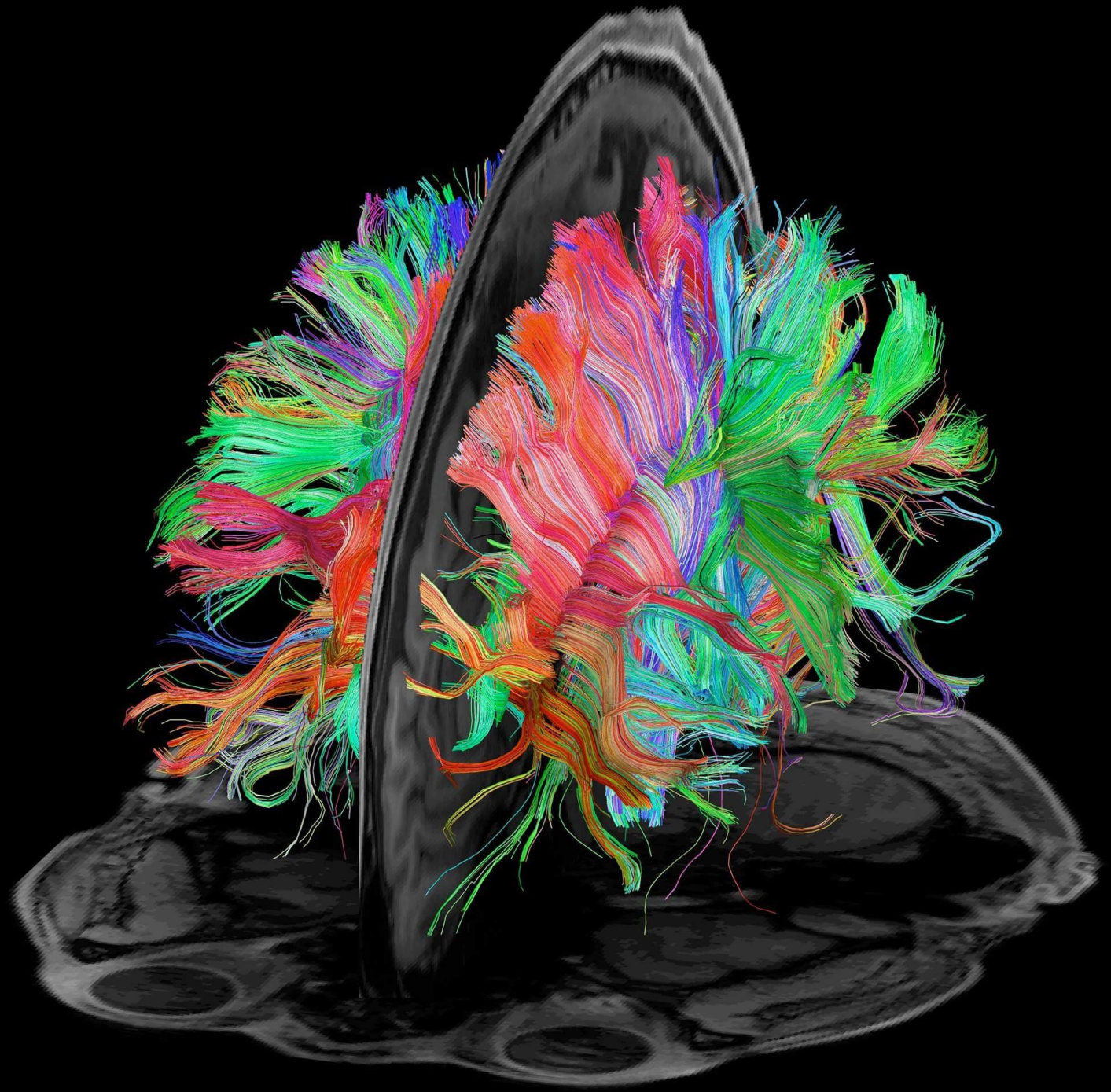


Genetic architecture of the white matter connectome of the human brain

Zhiqiang Sha, Ph.D.
Language and Genetics Department,
Max Planck Institute for Psycholinguistics

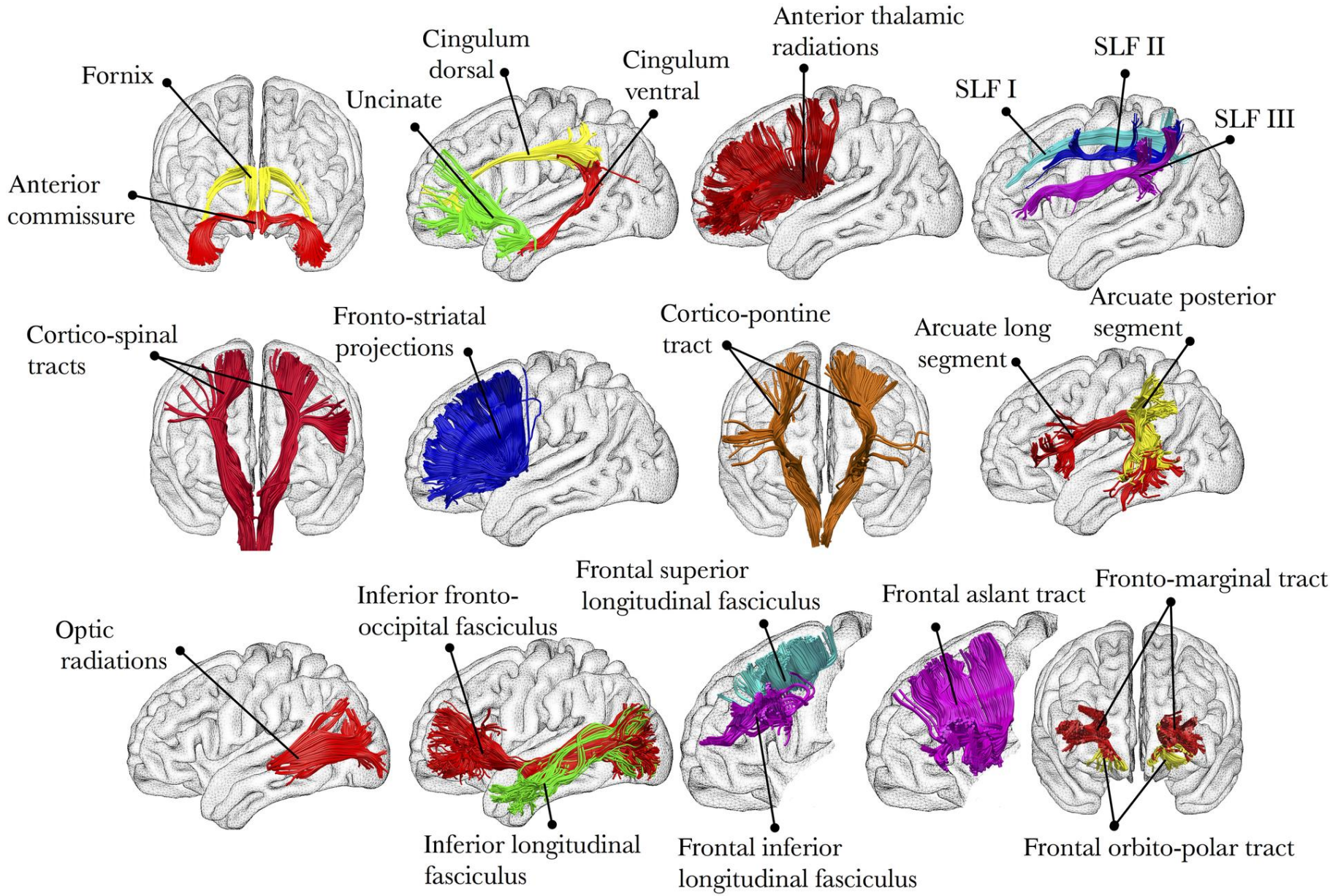
10.1101/2022.05.10.4
bioRxiv
THE PREPRINT SERVER FOR BIOLOGY

Background

White matter fiber tracts connecting neurons of different regions

Background

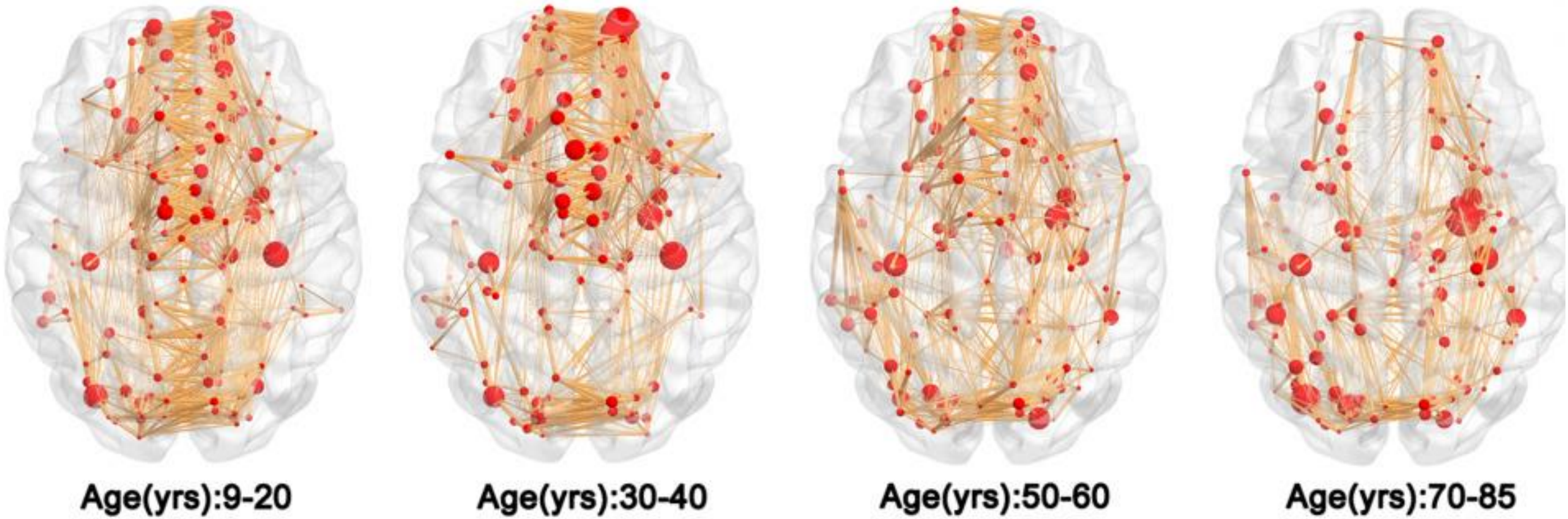
Major fiber tracts connecting cerebral cortex, subcortical regions and cerebellum



Wikipedia

Background

White matter connectivity changes during neurodevelopment and aging

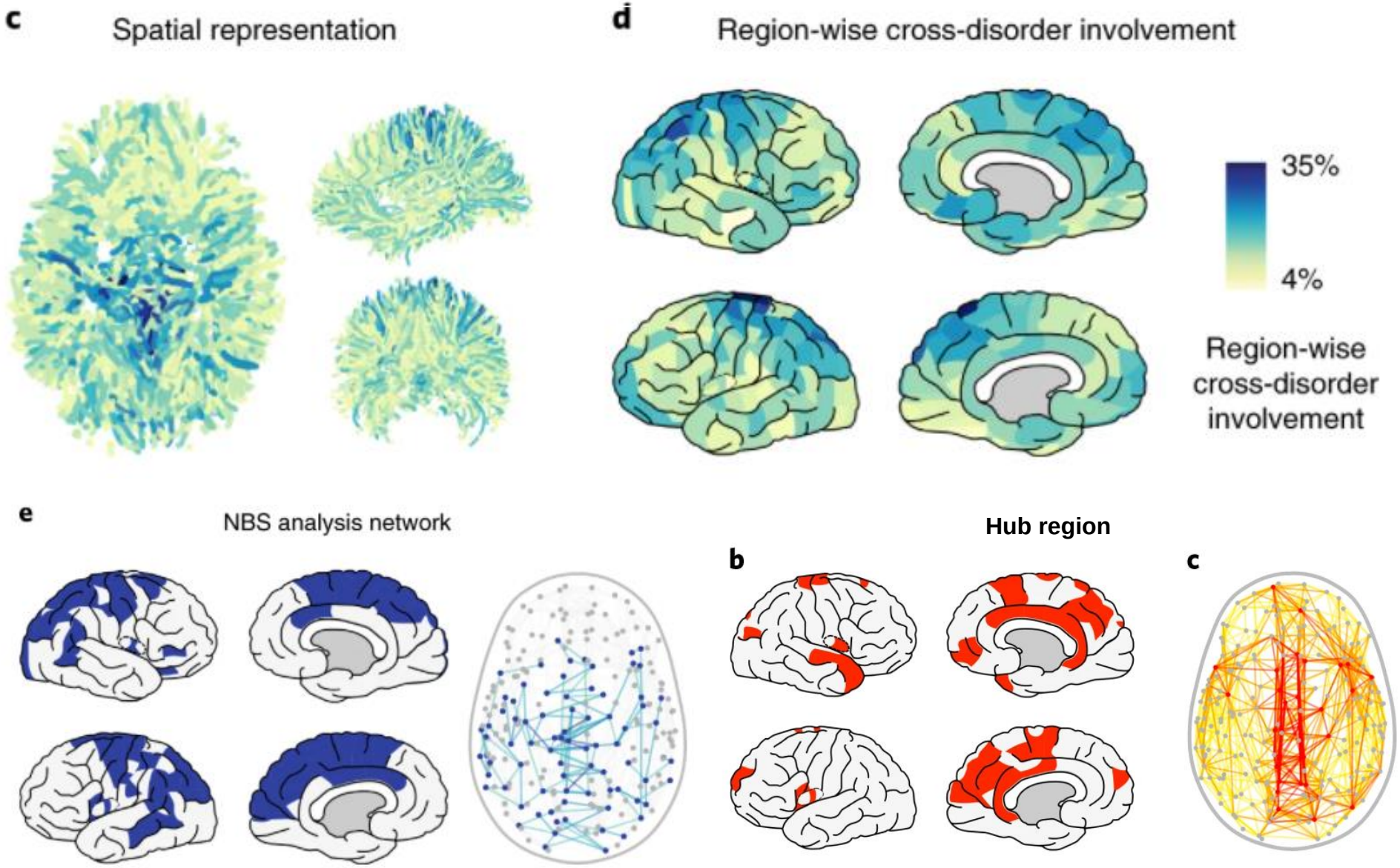


Dynamic changes of white matter connectome across lifespan

Zhao et al., 2015 HBM

Background

Disrupted white matter connectivity across neuropsychiatric disorders

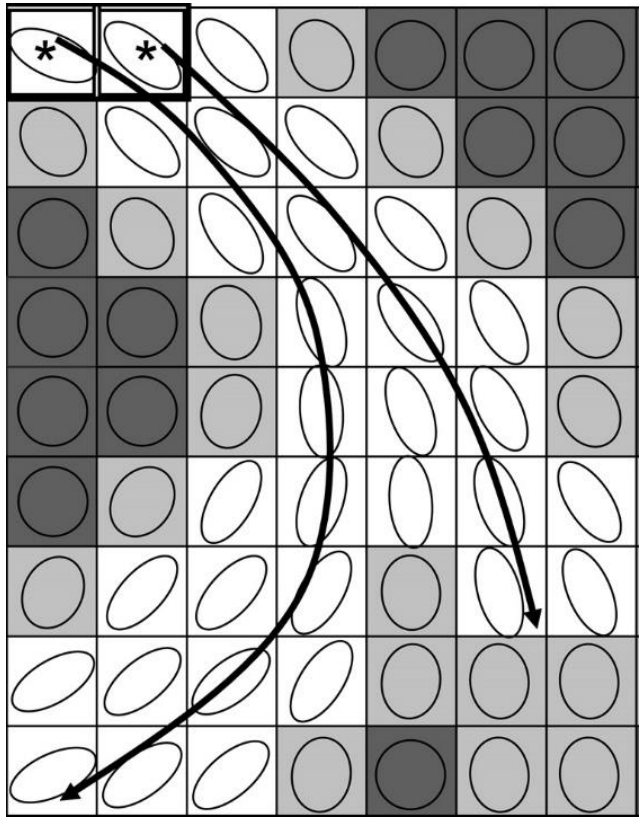


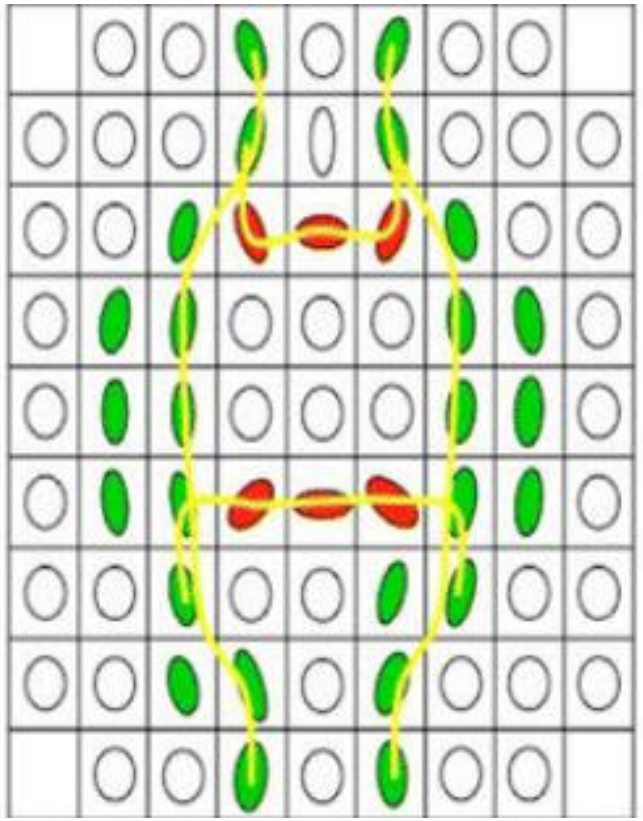
De Lange et al., 2019, NHB

Background

Tractography technology tracks the white matter tracts in the human brain

DTI is sensitive to water molecule diffusion which is constrained to be especially parallel to nerve fibers





Fiber tractography is to reconstruct fiber tracts (i.e. white matter connectivity) by linking the voxels that have the same orientation, while counting the streamlines between the start point and end point.

Mori et al., 1999, Ann Neurol
Mori et al., 2006, Neuron

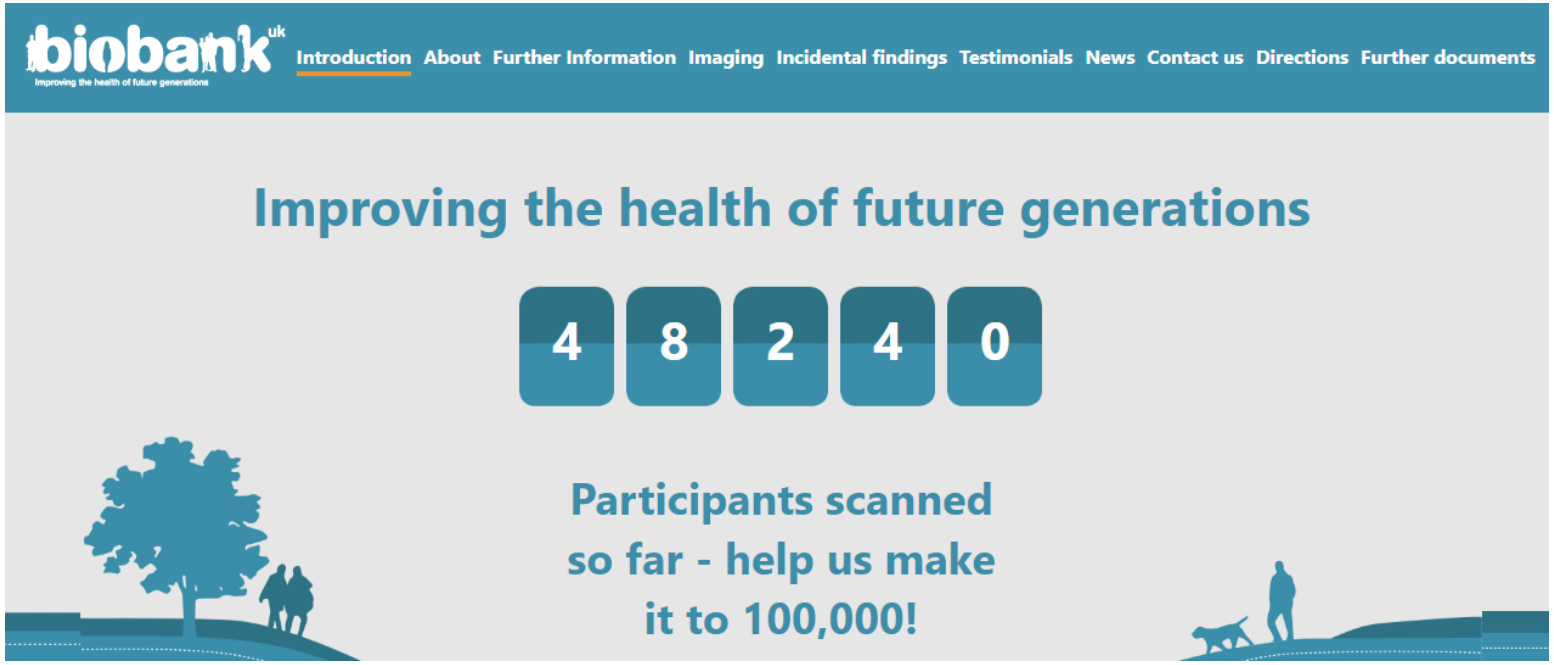
Background

UK Biobank (UKB) provides a great opportunity with large sample size to explore genetics of brain structure.

For the genetic information, UKB has collected about 500,000 subjects (ages 40-80 years), including genotypes, exome sequences and whole genome sequences.

For the imaging study, UKB aims to conduct detailed MRI imaging scans of over 100,000 participants in Stockport, Newcastle, Bistol and Reading centers.

Now, they have scanned about 48,000 subjects with brain imaging and genetic information.



Aims of the present study

Here we carried out a brain-wide tractography on 30,810 participants from UK Biobank.

- **Estimate heritability and perform multivariate GWAS analysis of the tract measures**
- **Functional annotation of genetic variants that associate with structural connectome**
- **Assess the association of regional connectivity with other clinical traits, including psychiatric disorders (schizophrenia, bipolar disorder, autism, attention-deficit hyperactivity disorder), neurological disorders (Alzheimer's disease, amyotrophic lateral sclerosis, and epilepsy) and handedness.**
- **Genetic influences on fiber tracts linking core language regions.**

Methods

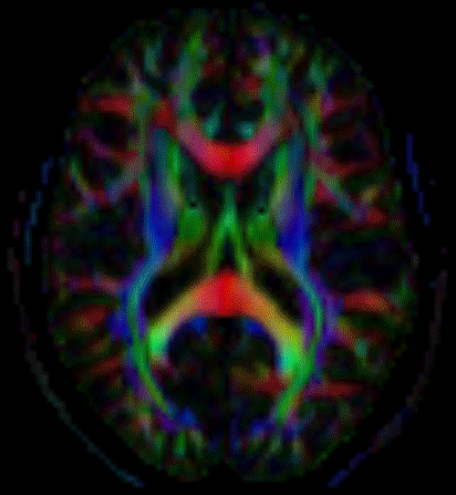
Individualized structural network construction

T1-weighted image

Anatomical parcellation

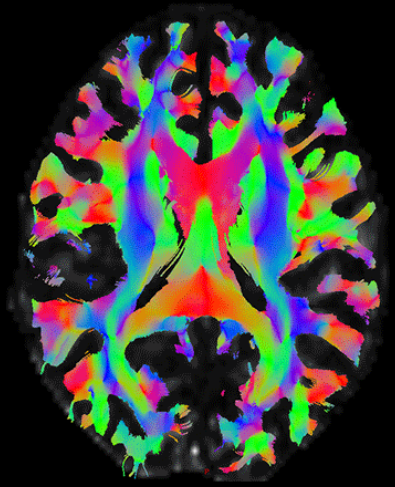
Node definition

Diffusion-weighted image

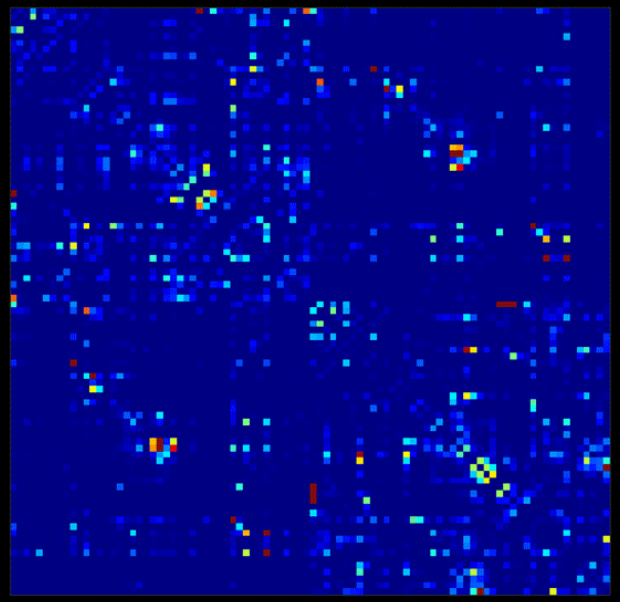


Fiber tractography

Edge definition

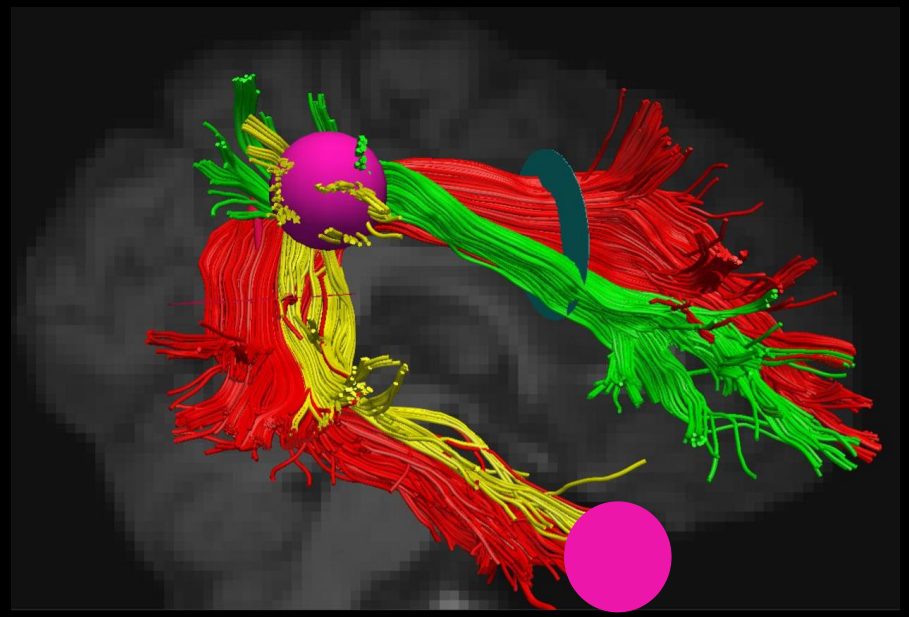


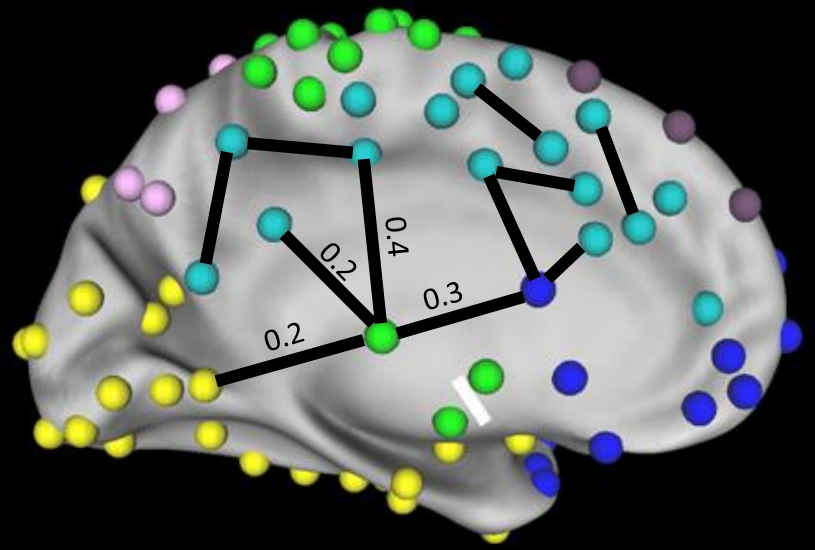
White matter structural connectivity matrix



Methods

Node and edge-level measures per individual





Two types of measurement per individual:

- **Edge-level**
  Count the number of streamlines linking each pair of regions, while adjusting for the brain size by dividing by the average volume of both regions.

- **Node-level**
  Degree centrality of each region: Sum the connectivity linking to a given region

Methods

Genetics of structural connectome of the human brain

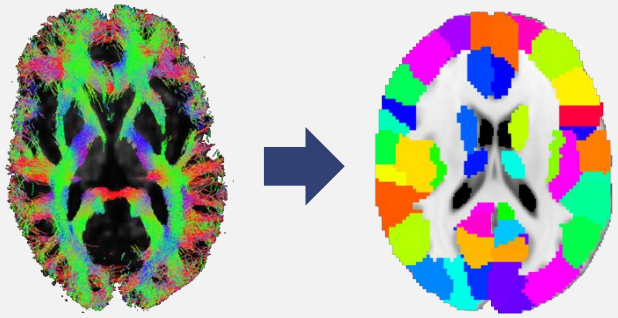
1. Sample quality control

- Exclude outliers based on heterozygosity (PC corrected heterozygosity>0.19) and genotype missingness (missing rate>0.05).
- Exclude subjects with a mismatch of their self-reported and genetically inferred sex.
- Exclude the subjects without "white British ancestry".
- Exclude the subjects with relatedness of kinship coefficient>0.0442.

2. Genetic quality control

- Exclude variants with minor allele frequencies<1%.
- Exclude variants with INFO<0.7.
- Exclude variants with Hardy-Weinberg equilibrium $p<1x10^{-7}$.

3. Brain measures



Automated Anatomic Labeling atlas

- Node level (degree)
- Edge level (connectivity)

We finally included 30,810 participants

	Male	Female
Participants	14,636	16,174
Age	64.56 ± 7.59	63.20 ± 7.34

Methods

SNP heritability and multivariate GWAS analysis

1. **SNP heritability:** Estimate the proportion of variance in a phenotype explained by all SNPs using GCTA.

(1) Remove cryptic relatedness (cutoff: 0.025).
(2) Construct genetic relationship matrix.
(3) Calculate genome-based restricted maximum likelihood (GREML).

2. **Multivariate GWAS** (separately for node- and edge-level metrics):

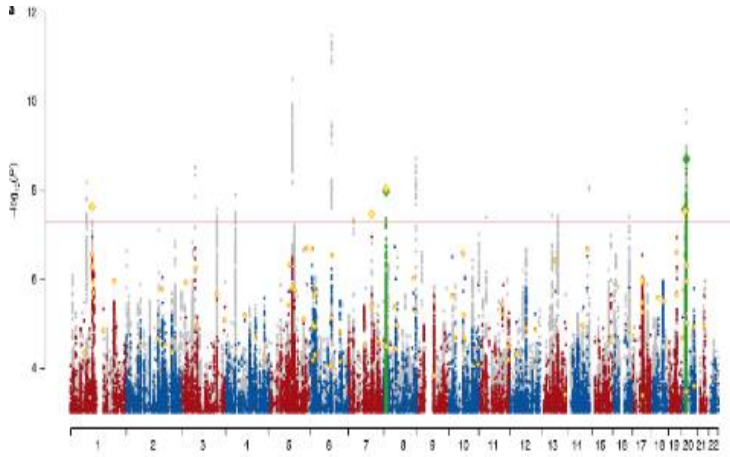
(1) For each significantly heritable trait, test **univariate association** with additive genetic model for each SNP.
(2) Separately for each SNP, test its univariate association with each brain measure, convert the P values to z scores, then use **Mahalanobis distance** to compute a single multivariate association X^2 (null distribution obtained through permutation; MOSTest software).
(3) Identify **brain traits** that make the greatest contributions to overall multivariate association by calculating the mean of unsigned z-scores across lead SNPs.

3. **Covariates used in the GWAS analysis:** age, nonlinear age, 10 genetic principle components, assessment center, genotype measurement batch and sex.

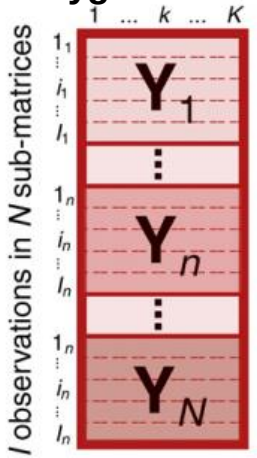
Methods

Associations of regional connectivity with different polygenic scores

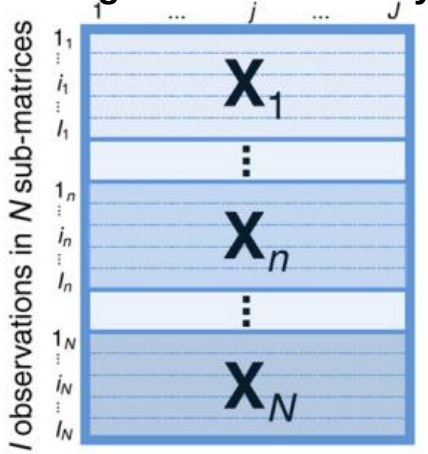
GWAS summary statistics



Y
Polygenic score



X
Regional connectivity



Canonical correlation model

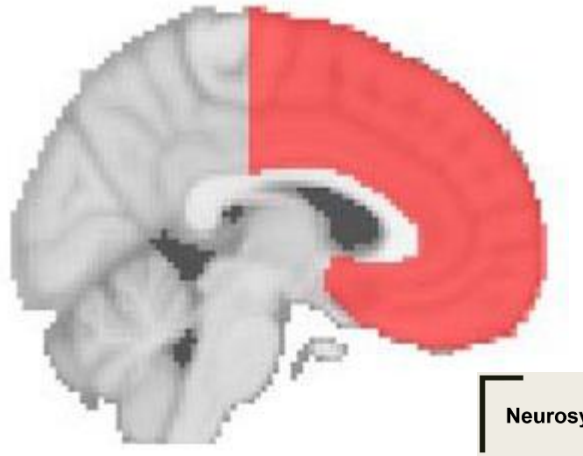
$$u_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1j}x_j$$

$$u_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2j}x_j$$

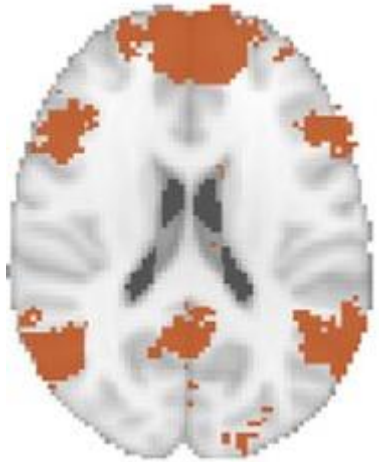
$$\vdots$$

$$u_N = a_{N1}x_1 + a_{N2}x_2 + \cdots + a_{Nj}x_j$$

1) Region with high loadings (|r|>0.2)



2) Co-activation map



Neurosynth

10,000 fMRI studies

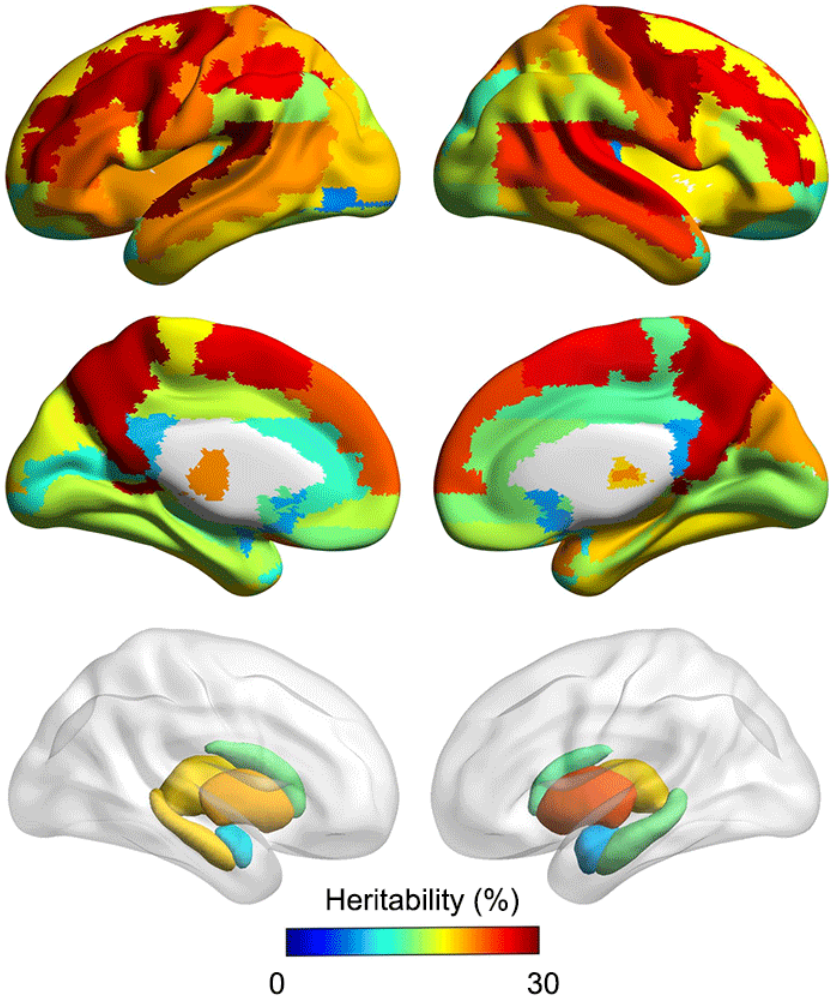
3) Cognitive annotations

Neurosynth.org

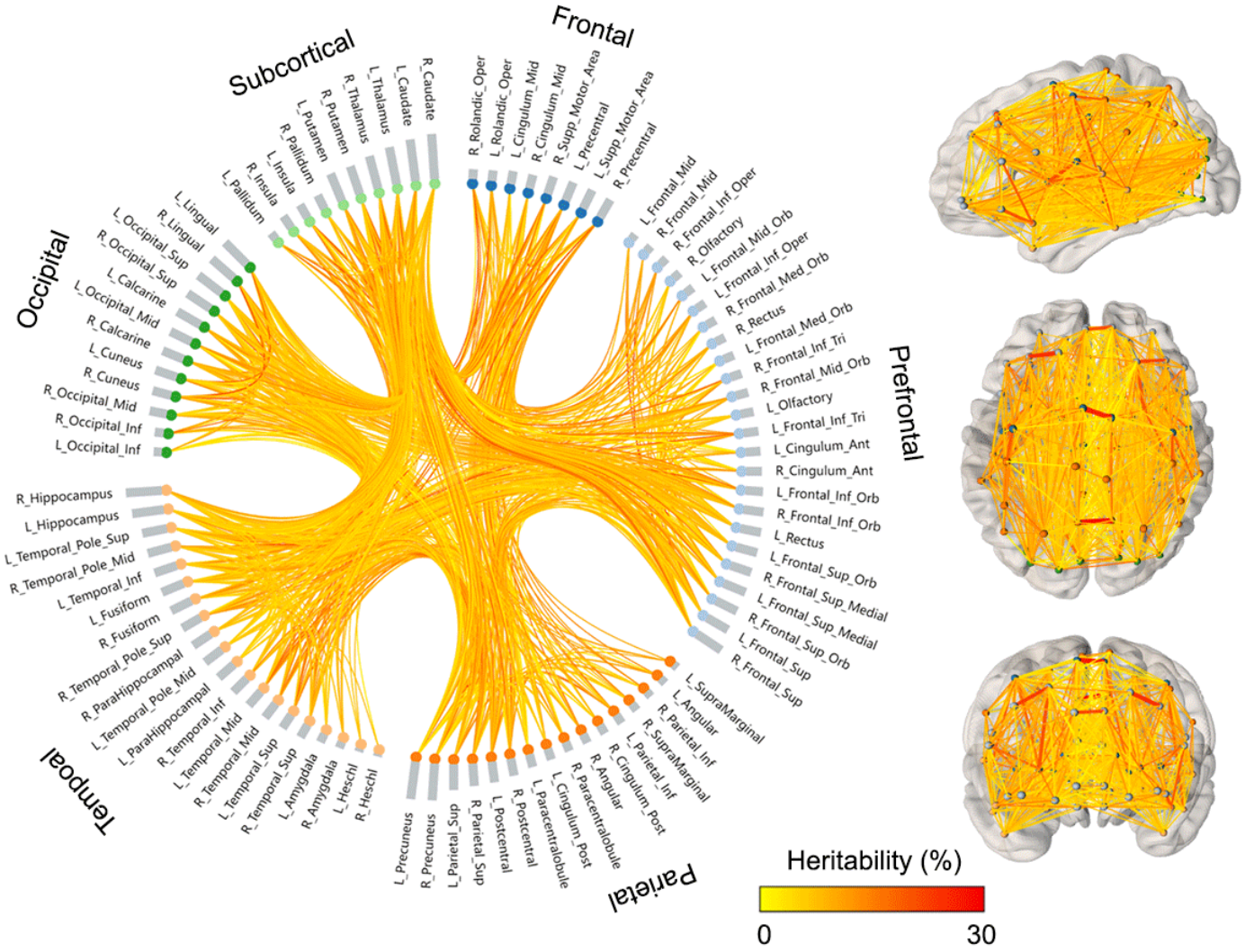
Results

SNP heritability analysis

Heritability of node-level measures



Heritability of edge-level measures

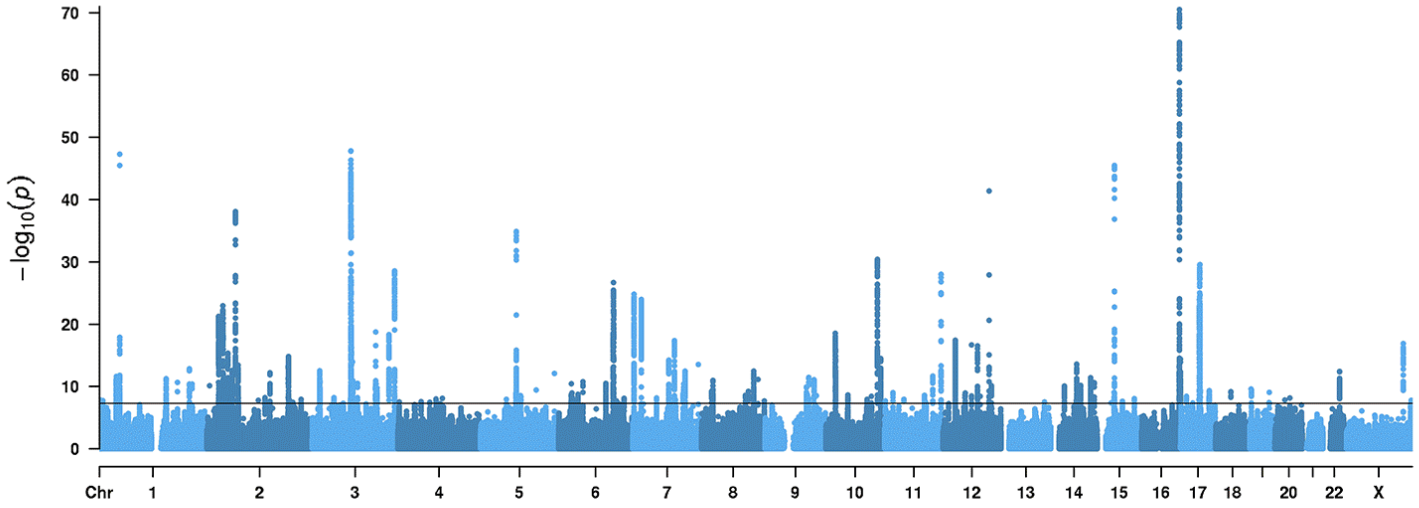


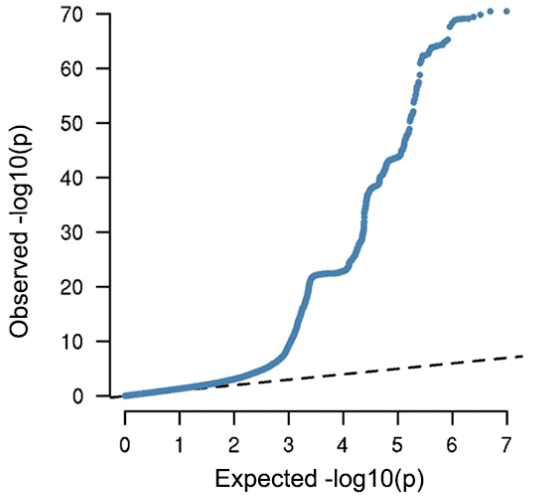
851 edge-level measures showed significantly heritable after Bonferroni correction

Results

Multivariate GWAS analysis

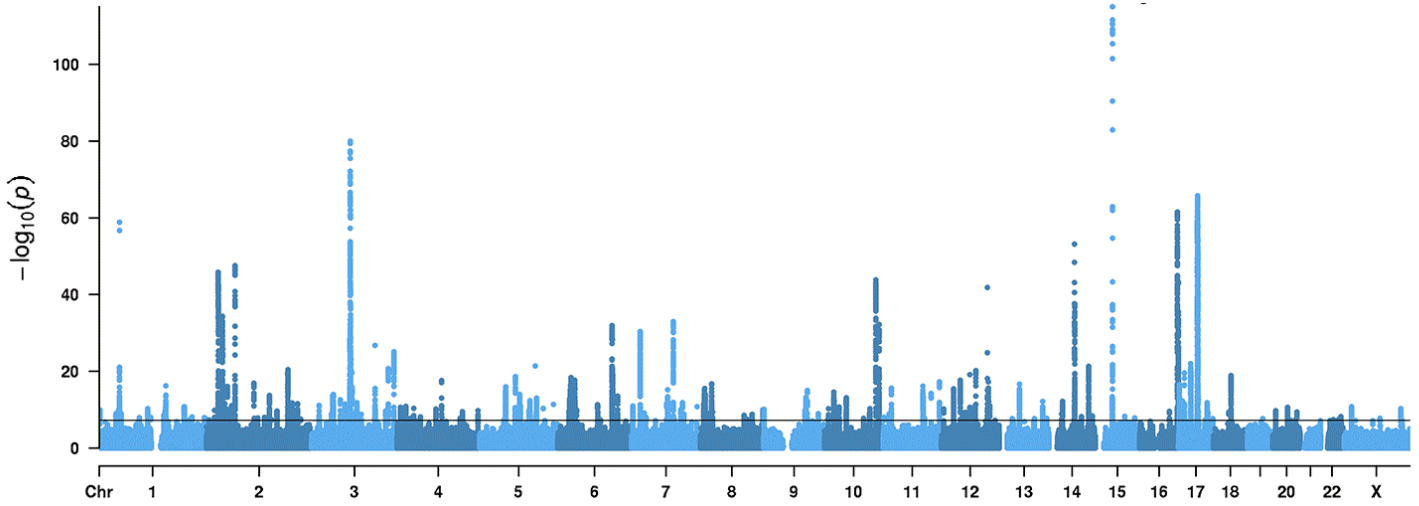
Multivariate GWAS of node-level measures

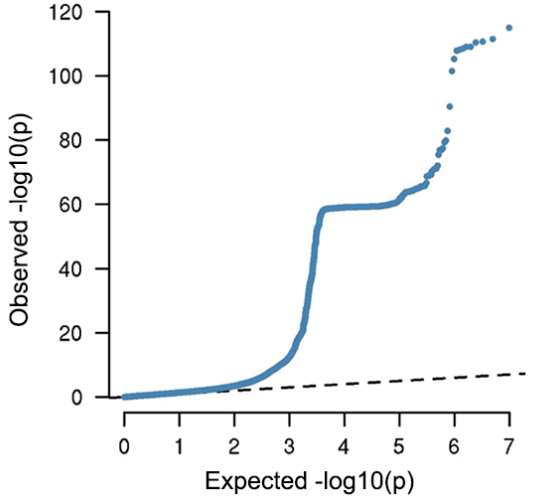




154 lead SNPs in 128 genomic loci

Multivariate GWAS of edge-level measures



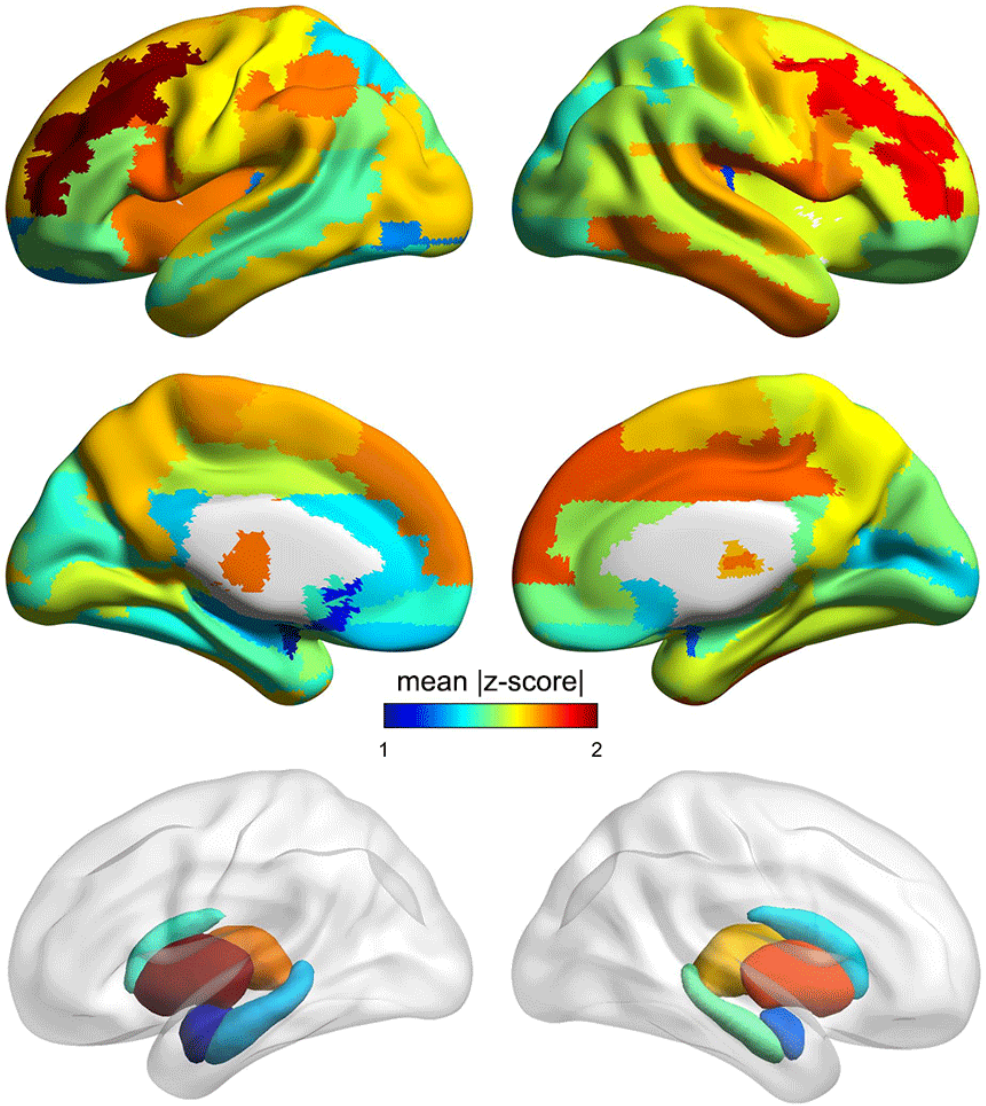


231 lead SNPs in 181 genomic loci

Results

Brain regions contributing to the multivariate associations

Regions driving significant associations

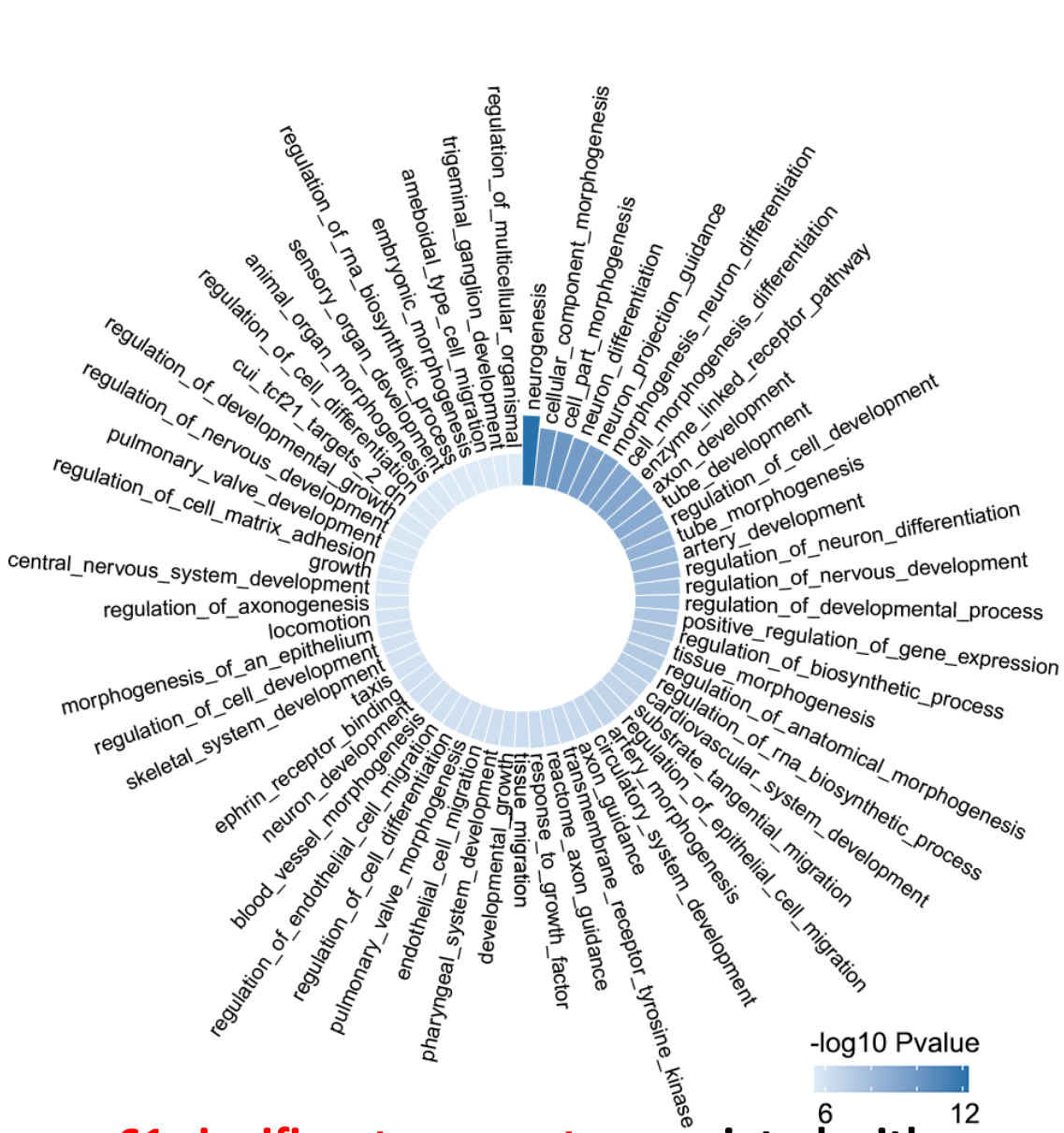


Connectivity driving significant associations

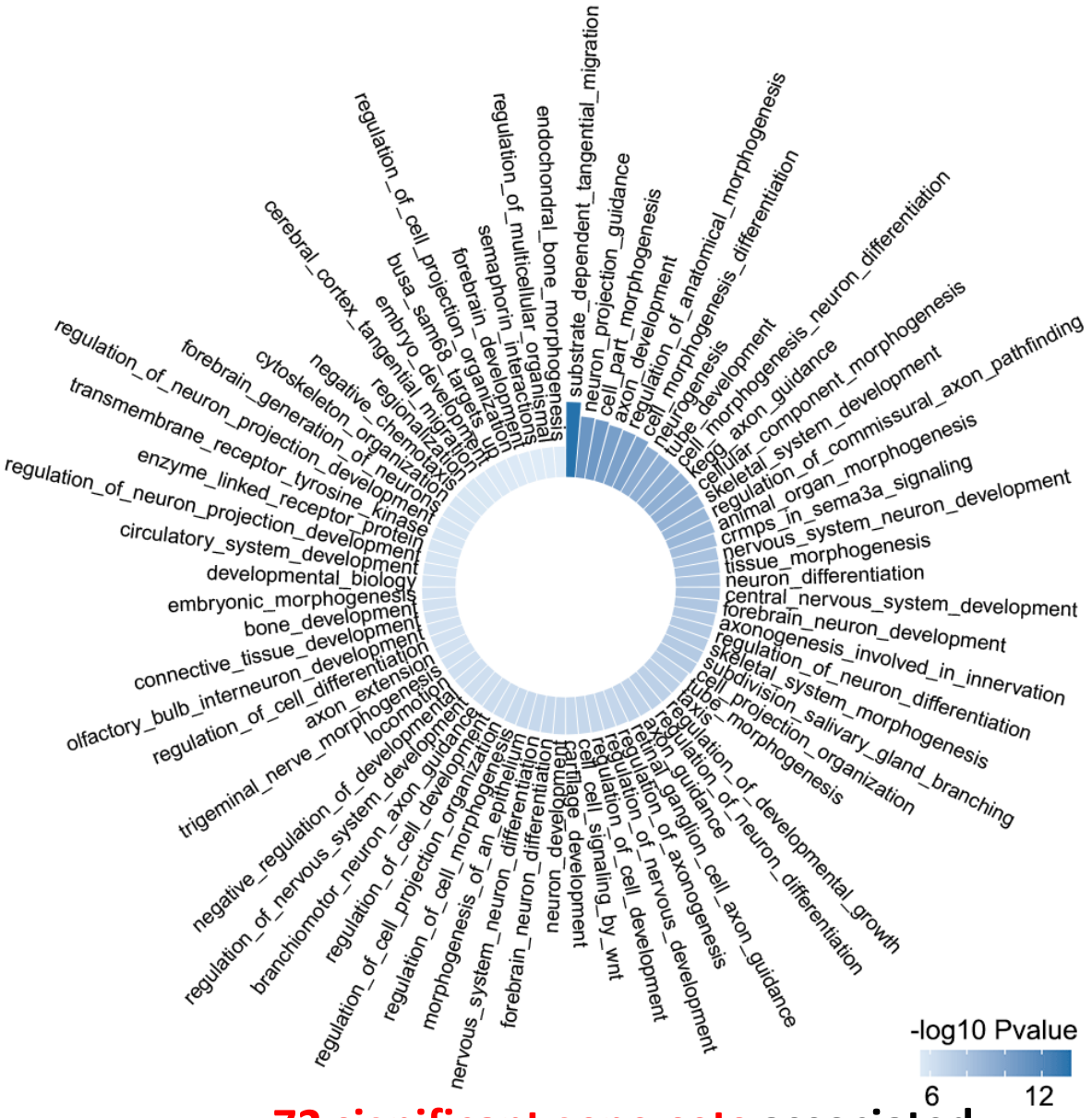
No.	Region1	Region2	Mean |z-score|
1	R_Precuneus	L_Precuneus	1.59
2	R_Calcarine	L_Calcarine	1.52
3	R_Temporal_Mid	R_Precentral	1.52
4	R_Putamen	R_Postcentral	1.52
5	L_Thalamus	L_Calcarine	1.49
6	L_Temporal_Sup	L_Insula	1.48
7	L_Angular	L_Precentral	1.47
8	L_Temporal_Mid	L_Precentral	1.47
9	R_Lingual	L_Calcarine	1.47
10	L_Frontal_Mid	L_Frontal_Sup	1.47
11	L_Temporal_Mid	L_Temporal_Sup	1.45
12	R_Frontal_Mid	R_Frontal_Sup	1.44
13	R_Putamen	R_Frontal_Sup	1.44
14	R_Thalamus	R_Calcarine	1.43
15	R_Frontal_Sup_Medial	L_Frontal_Sup_Medial	1.42
16	R_Lingual	R_Cuneus	1.42
17	L_Frontal_Inf_Tri	L_Frontal_Sup	1.41
18	R_Supp_Motor_Area	L_Supp_Motor_Area	1.41
19	L_Frontal_Inf_Tri	L_Frontal_Mid	1.41
20	R_Occipital_Sup	L_Occipital_Sup	1.41

Results

Functional annotations of structural connectome-related genes



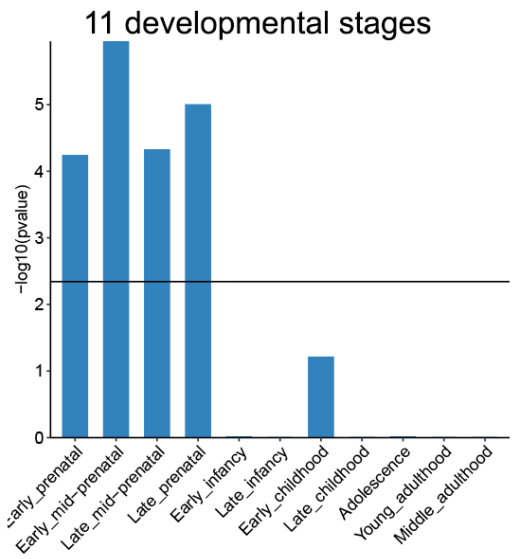
61 significant gene-sets associated with node-level measures

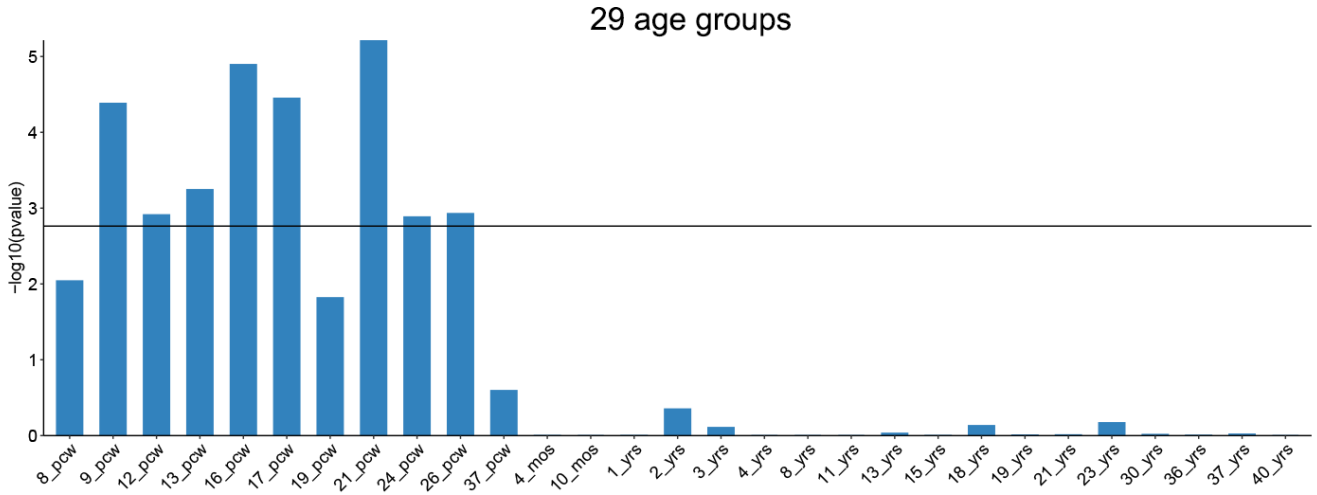


72 significant gene-sets associated with edge-level measures

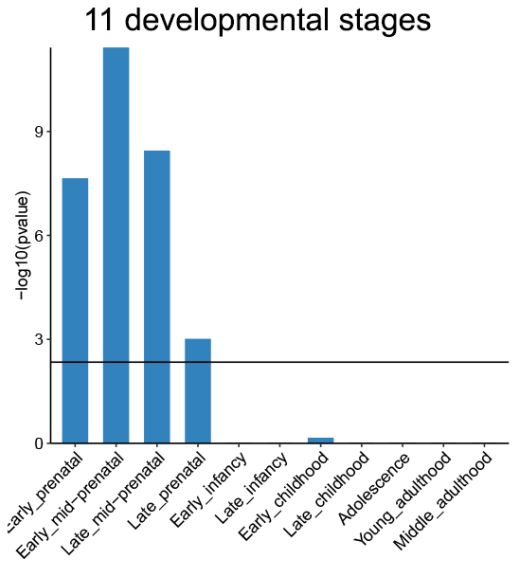
Results Neurodevelopment annotations of structural connectome-related genes

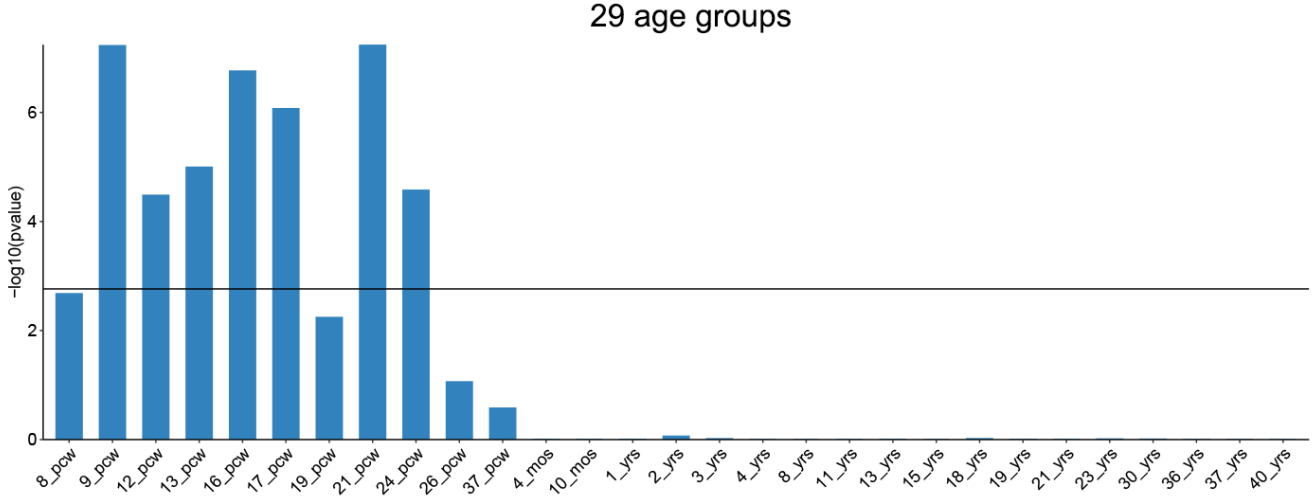
Node-level connectivity





Edge-level connectivity

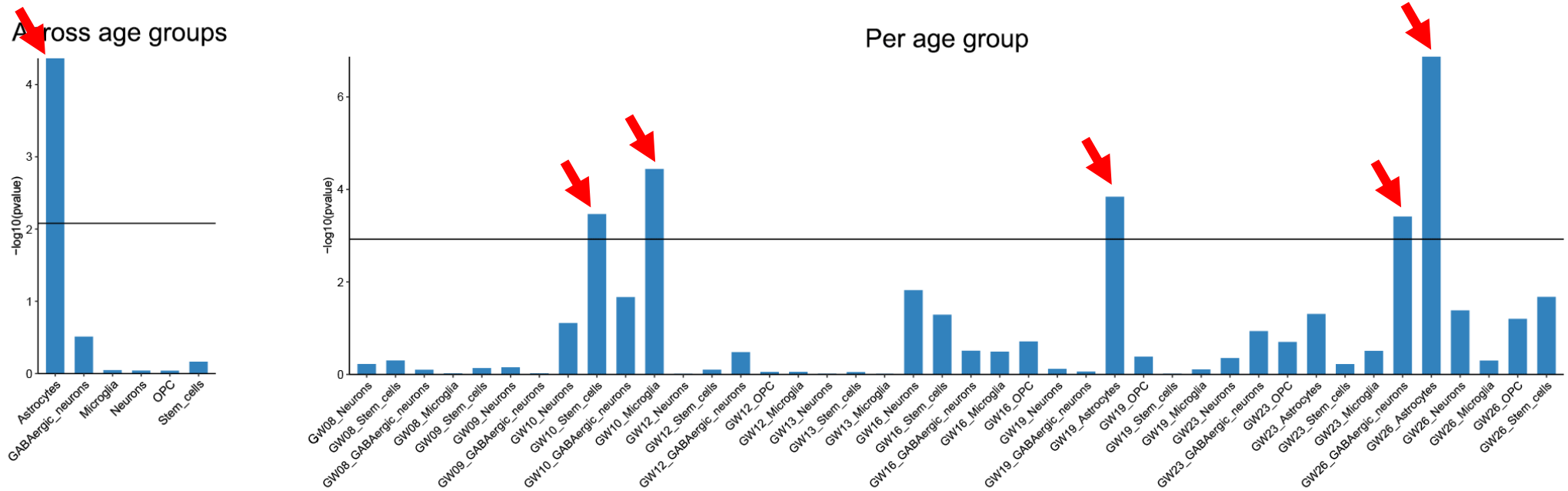




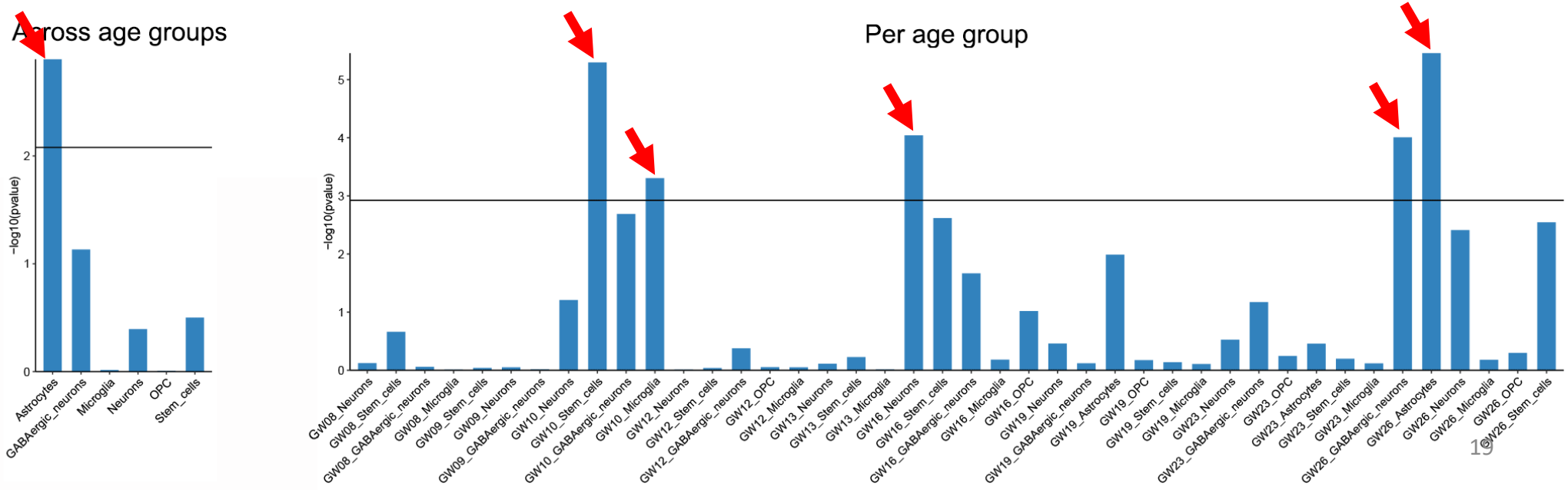
Results

Cell-type annotations of structural connectome-related genes

Node-level connectivity

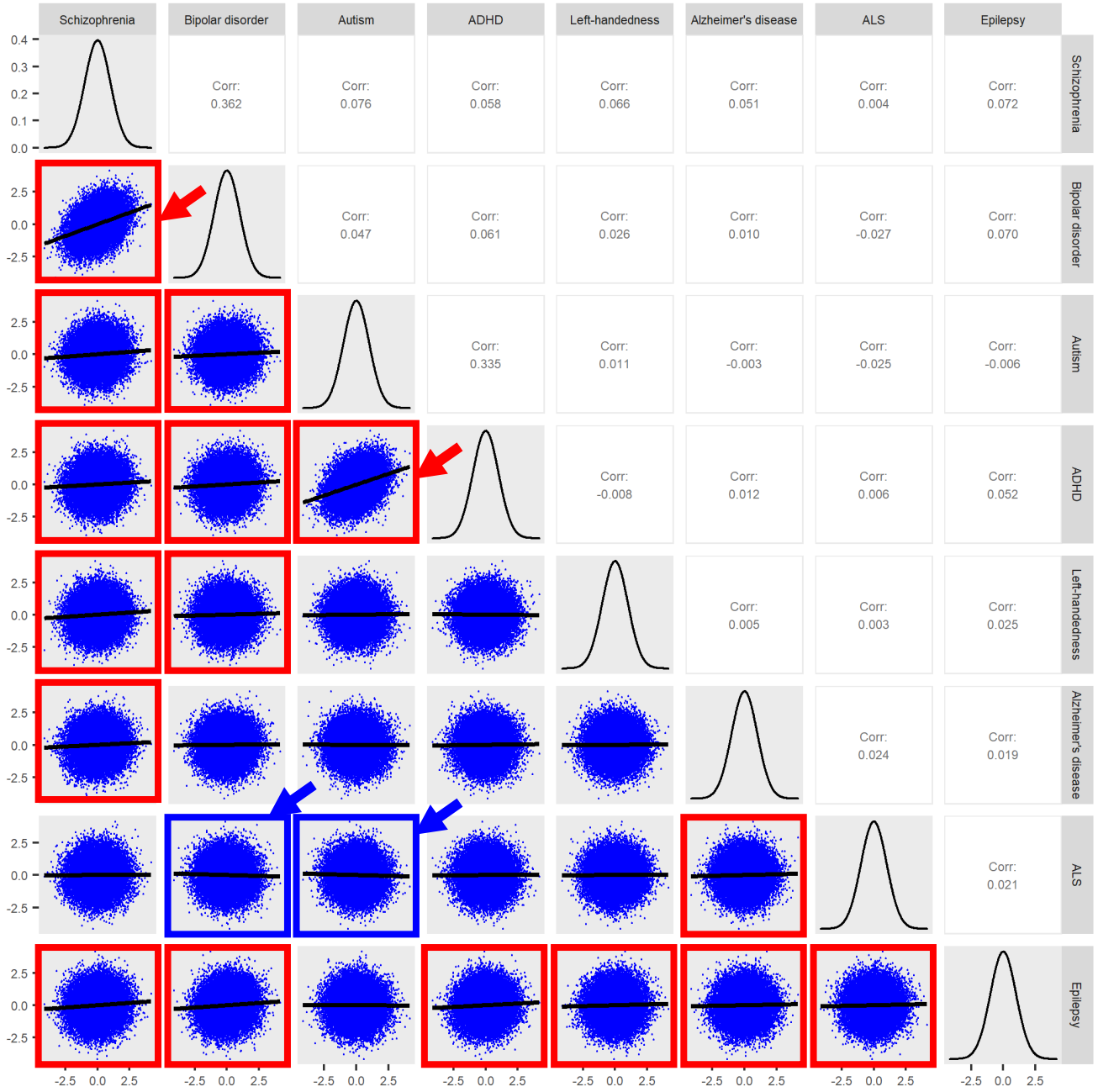


Edge-level connectivity



Results

Correlations between different polygenic scores

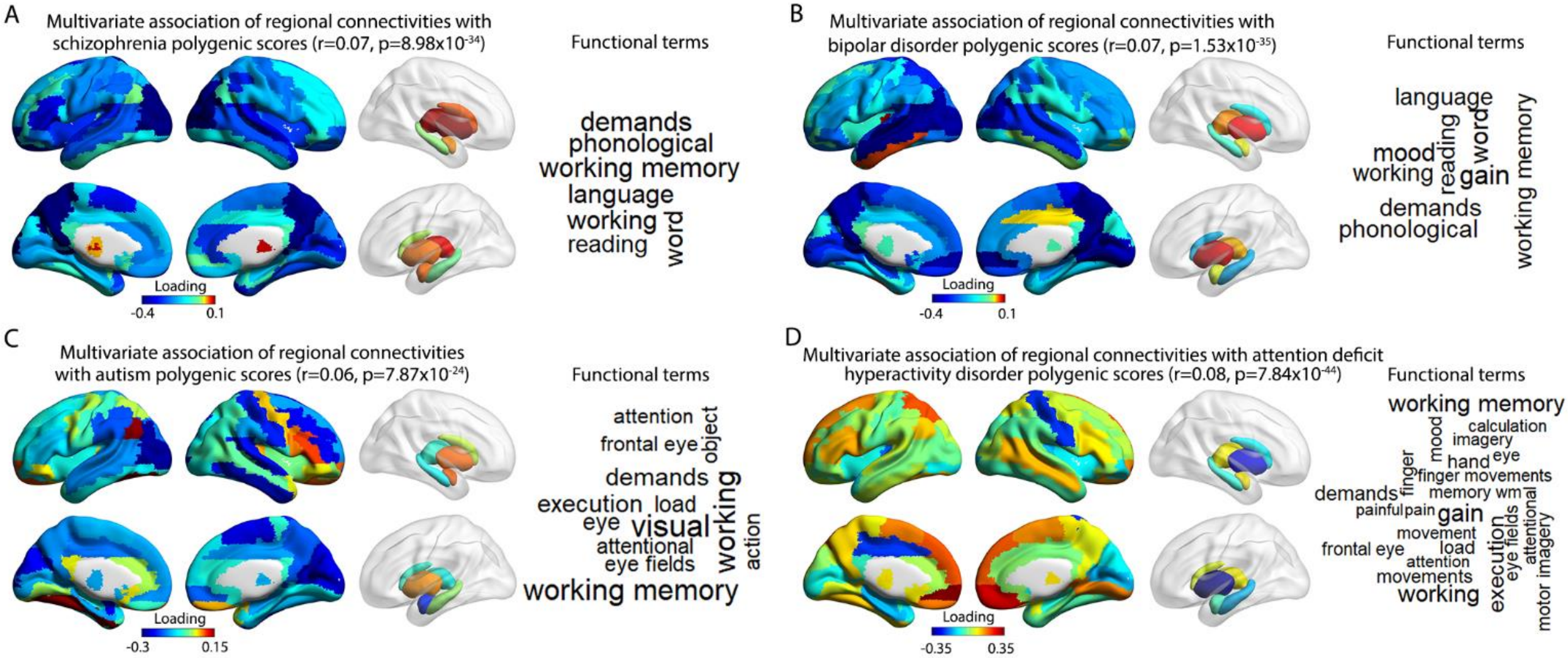


- **Sixteen correlations were positive**, with the highest between polygenic scores for schizophrenia and bipolar disorder, and between attention deficit/hyperactivity disorder and autism.

- **Two correlations were negative**, between polygenic scores for amyotrophic lateral sclerosis and bipolar disorder, and between amyotrophic lateral sclerosis and autism.

Results

Multivariate association of regional connectivity with polygenic scores



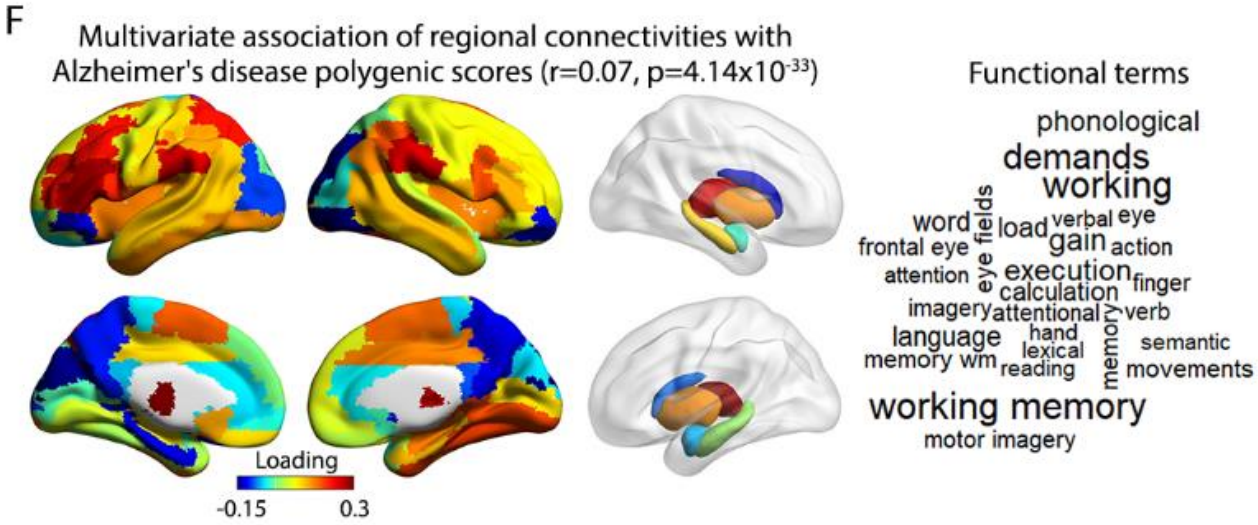
Increased polygenic risk for schizophrenia, bipolar disorder, autism associated with generally reduced connectivity of cortical regions, increased subcortical connectivity

Results

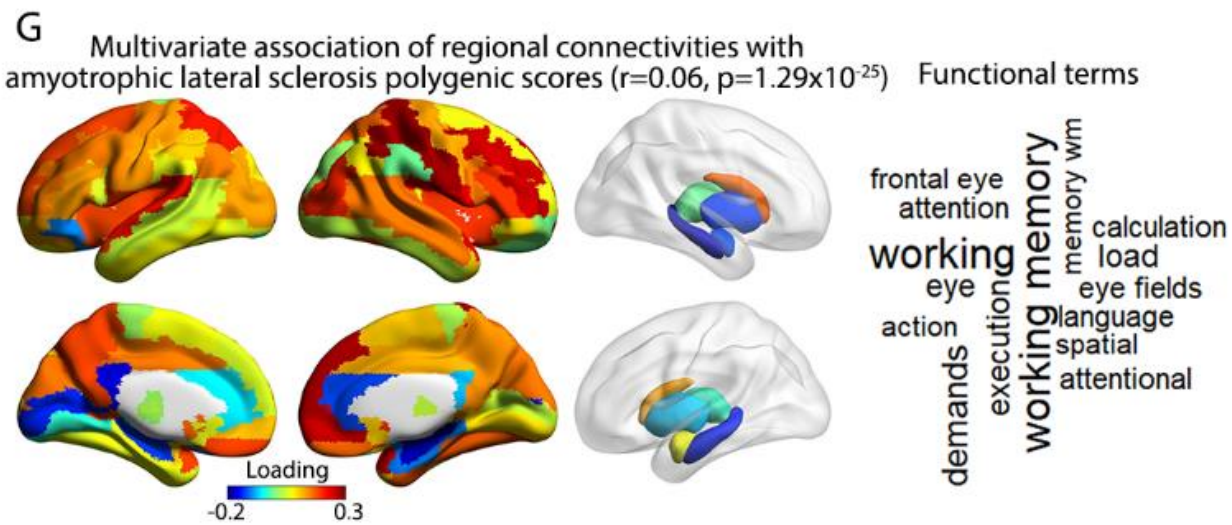
Multivariate association of regional connectivity with polygenic scores

- **Increased polygenic risk for Alzheimer's disease, amyotrophic lateral sclerosis associated with generally increased connectivity of cortical regions**
  - **Plasticity > compensatory changes?**
  - **Individuals expressing their polygenic risk might participate less?**

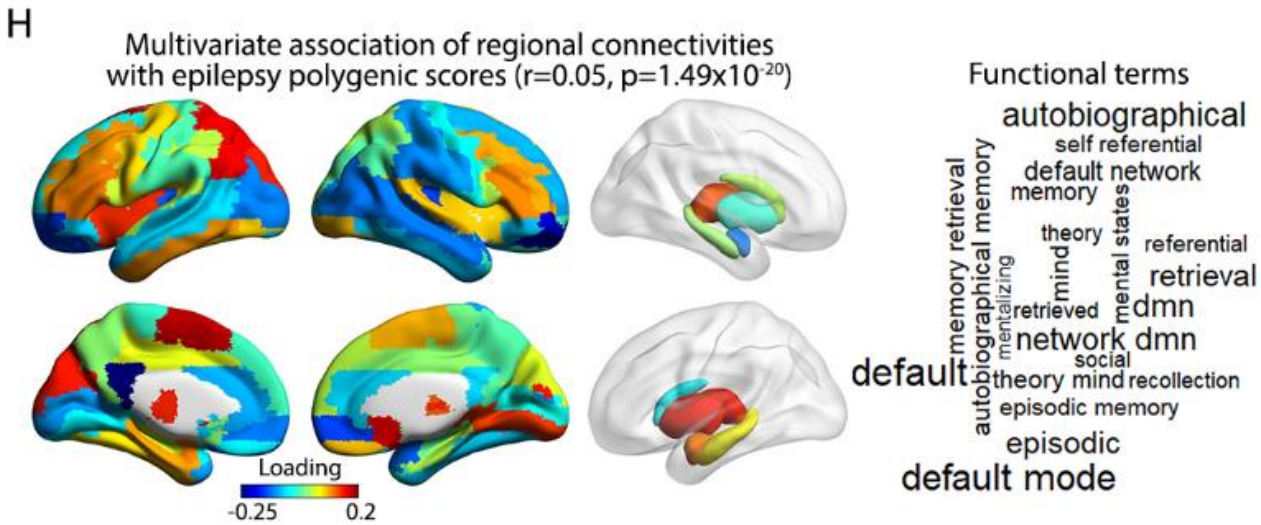
F



G



H

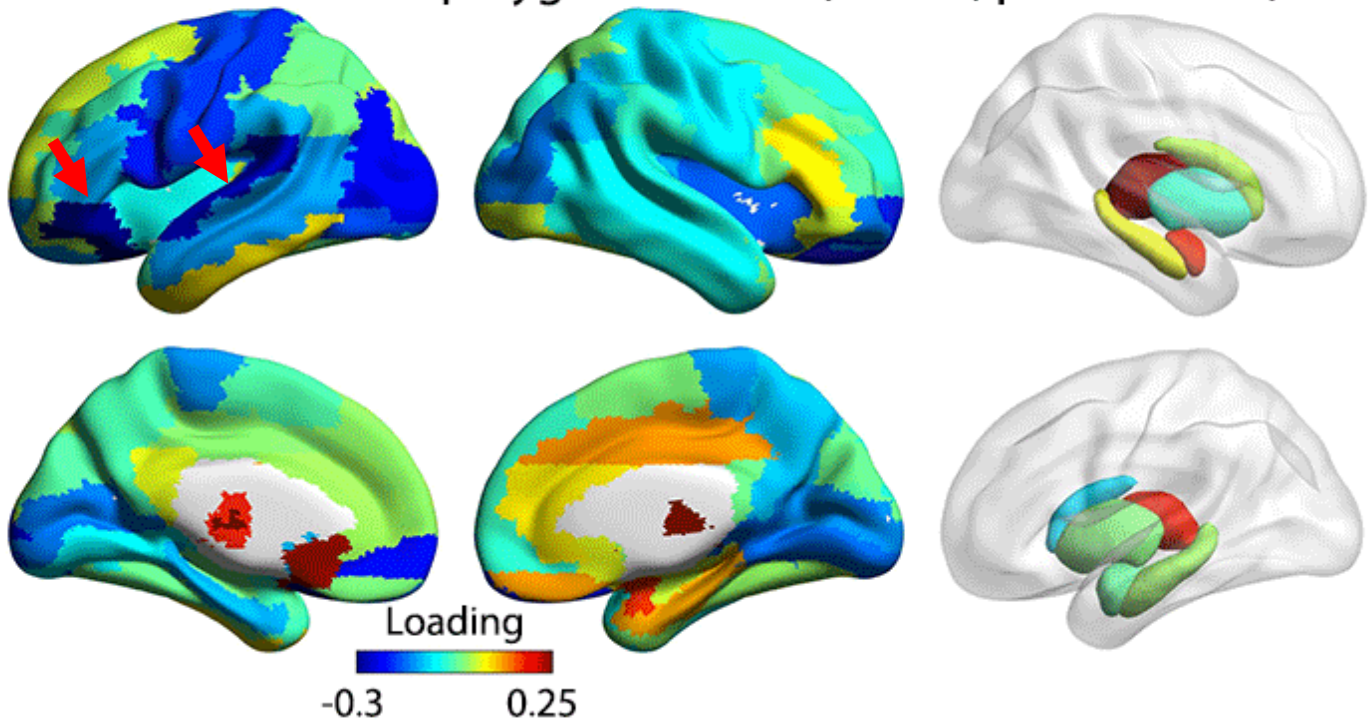


Results

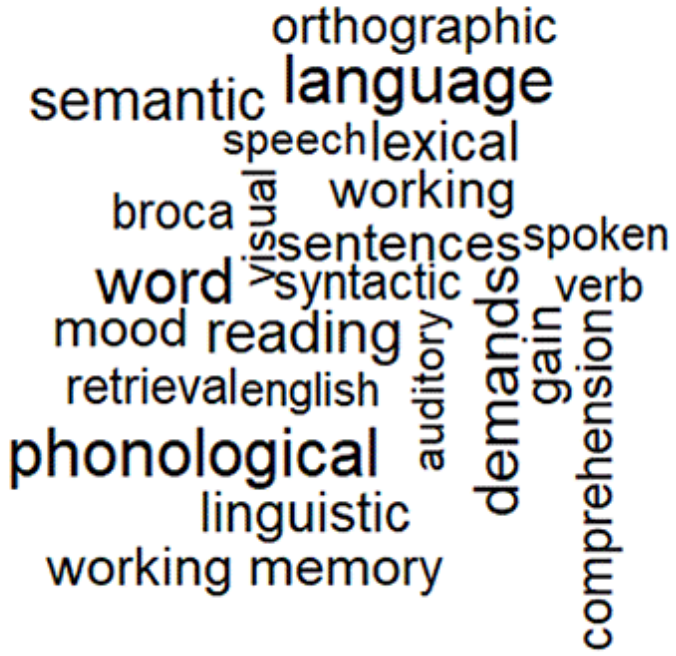
Multivariate association of regional connectivity with handedness polygenic score

- **Polygenic influence on handedness associated with connectivity of language-related regions**
  - **Consistent with findings based on anatomical asymmetry of the cerebral cortex (Sha et al. PNAS 2021)**
  - **Developmental and evolutionary links between left-hemispheric specializations for handedness and language**

Multivariate association of regional connectivities with left-handedness polygenic scores ($r=0.07$, $p=1.74\times10^{-31}$)

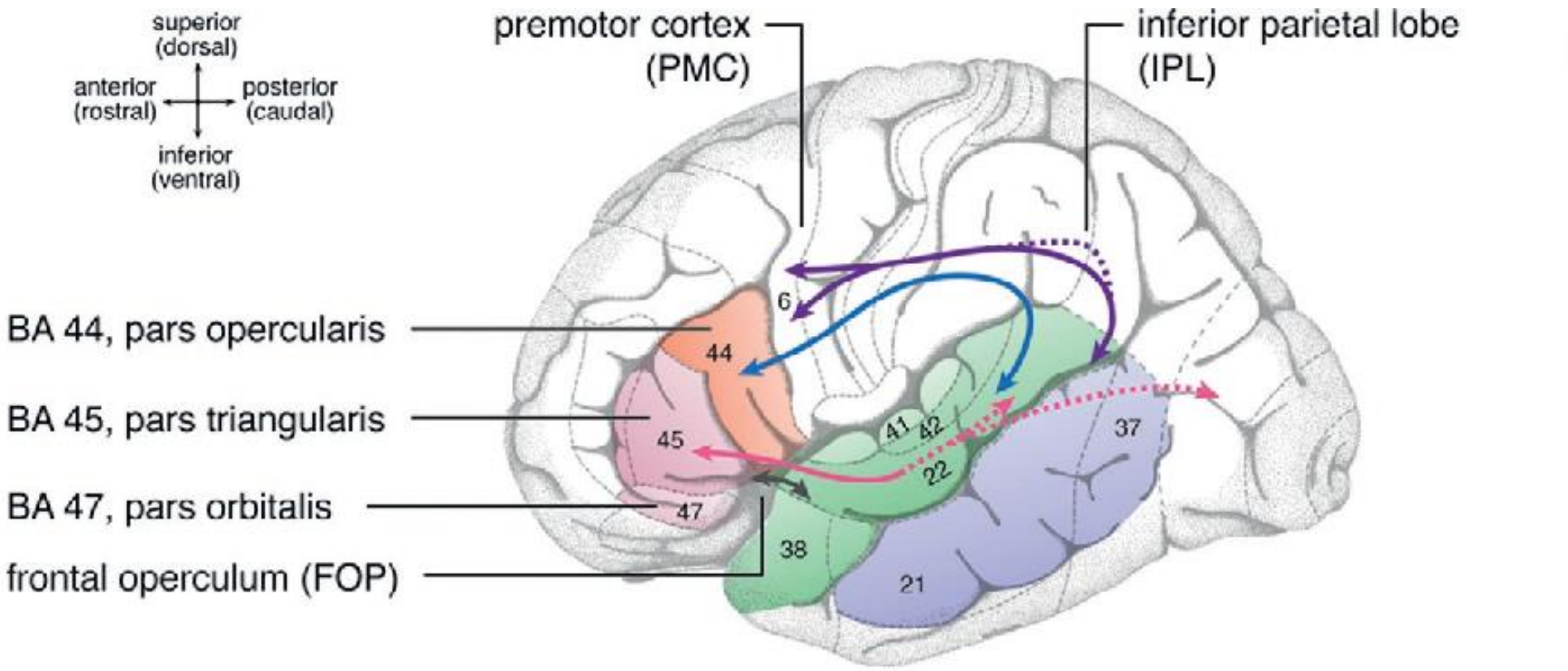


Functional terms



Results

White matter connectivity linking left-hemispheric language regions



inferior frontal gyrus (IFG)

superior temporal gyrus (STG)

middle temporal gyrus (MTG)

Dorsal Pathways

PMC to pSTG/MTG

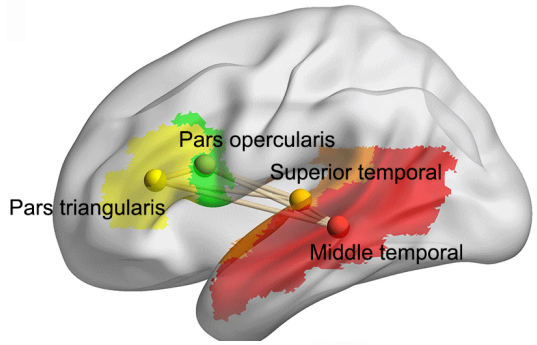
BA 44 to pSTG

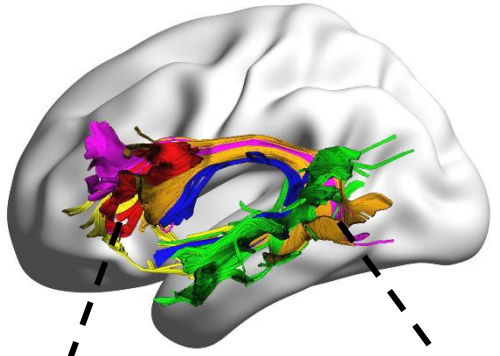
Ventral Pathways

BA 45/47 to STG/MTG

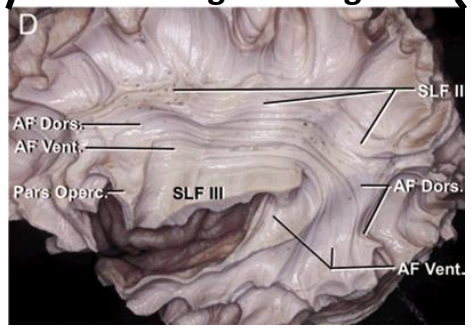
FOP to aSTG

Regions of our interest





Histological image



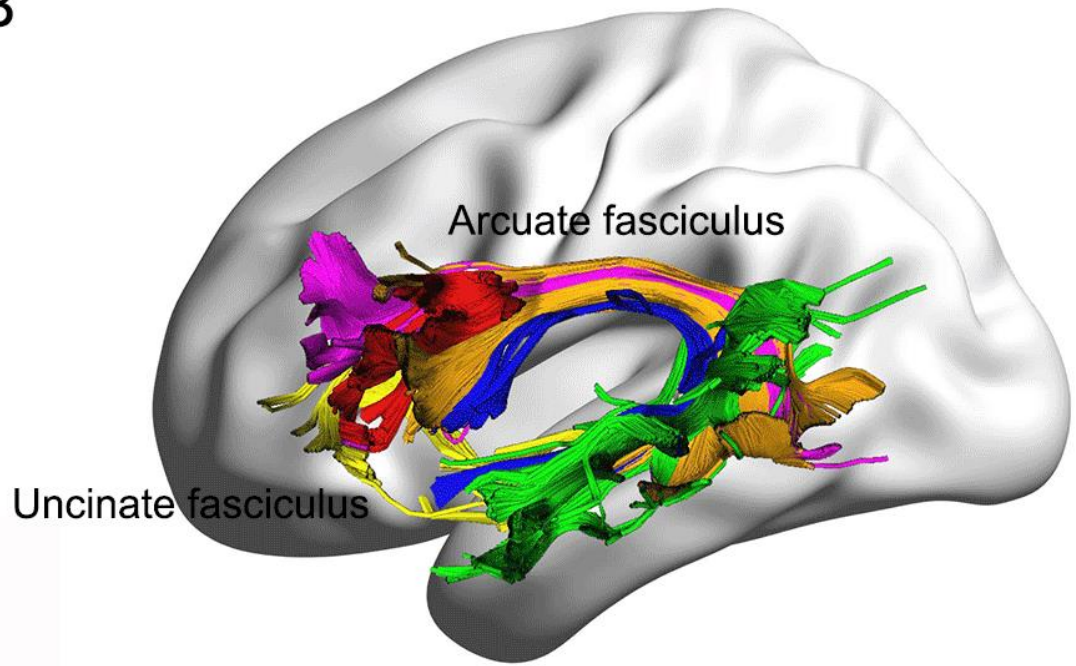
Friederici et al., 2015, Handbook of clinical neurology

Yagmulu et al. 2016, J Neurosurgery

Results

Language-related tracts and closest genes to lead SNPs

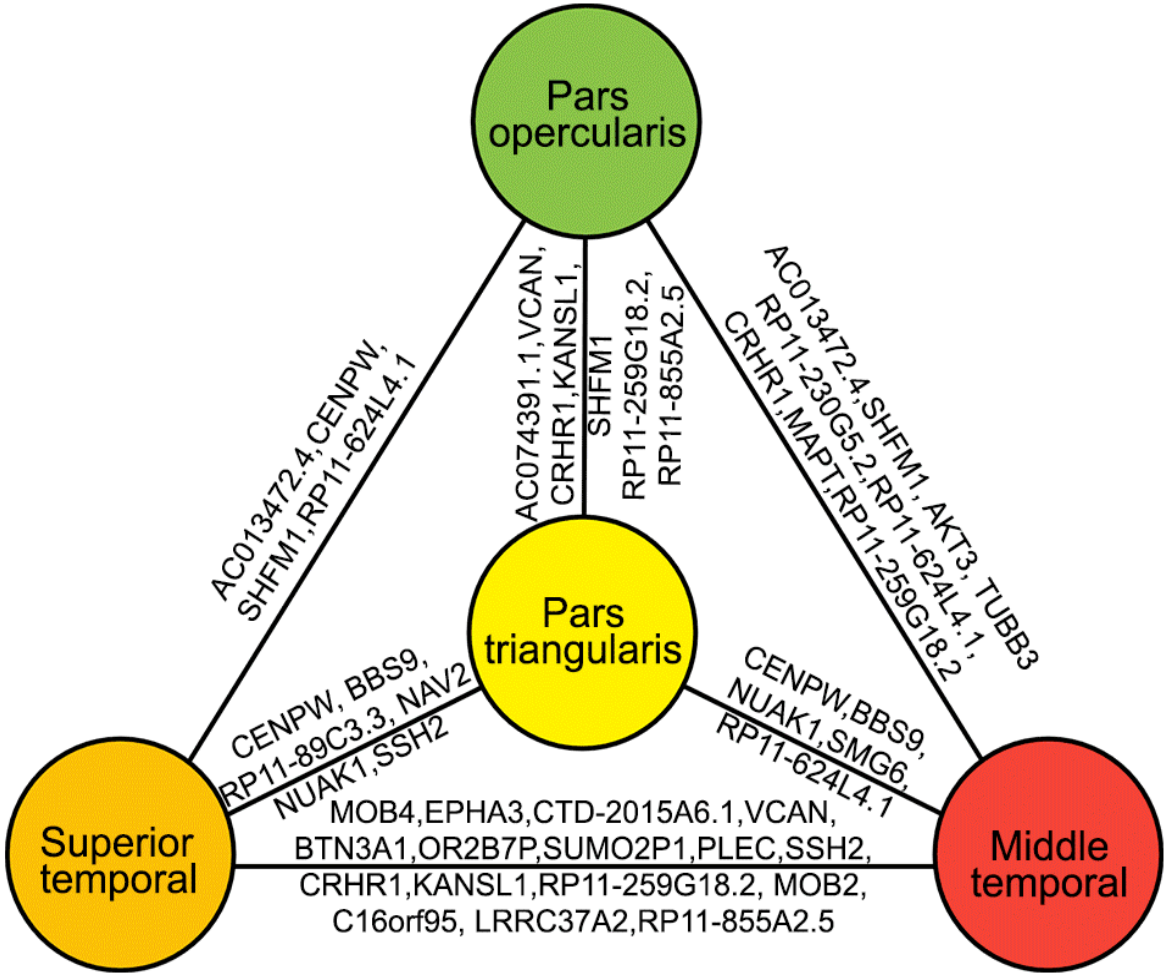
B



- Connectivity between the pars opercularis and pars triangularis cortex
- Connectivity between the middle temporal and superior temporal cortex
- Connectivity between the pars opercularis and middle temporal cortex
- Connectivity between the pars opercularis and superior temporal cortex
- Connectivity between the pars triangularis and middle temporal cortex
- Connectivity between the pars triangularis and superior temporal cortex

C

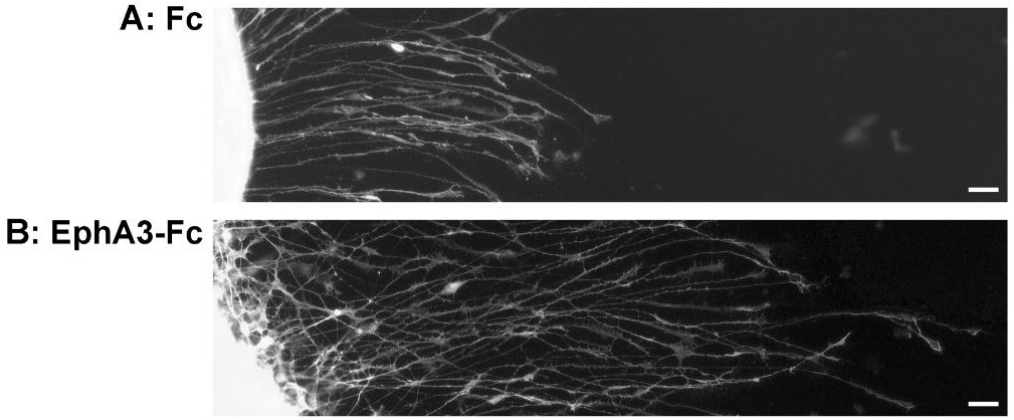
Closest genes to lead SNPs associated with core language network fiber tracts

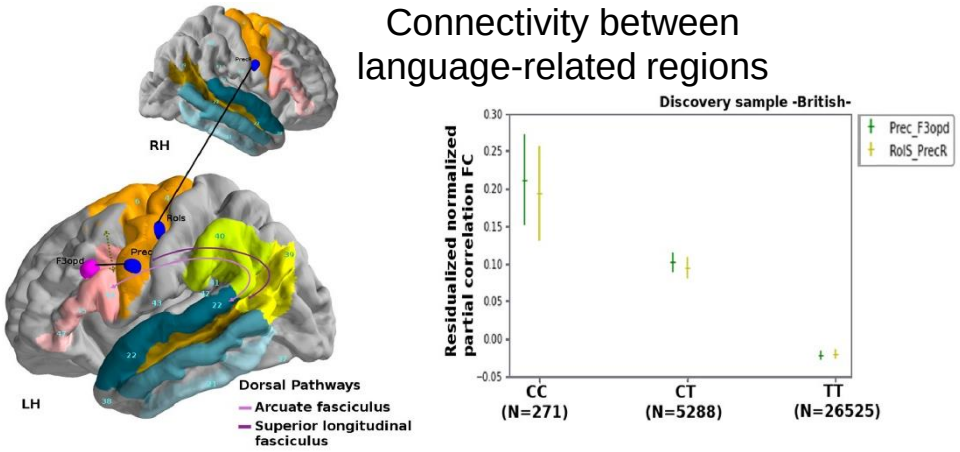


Results

Language-related tracts and closest genes to lead SNPs

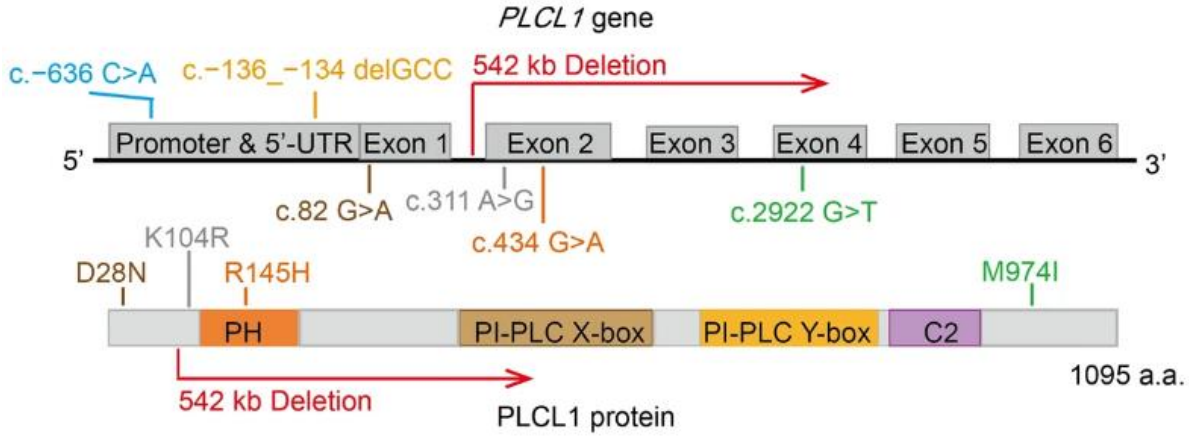
rs12636275 is an intron of ***EPHA3***, regulating **axon projection maps** and **language-related connectivity**

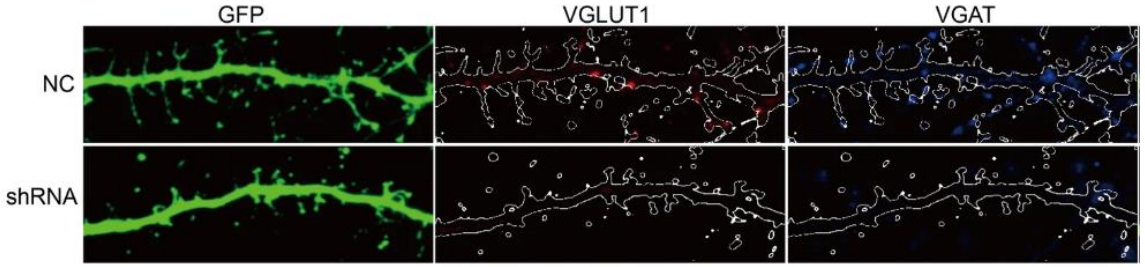




Connectivity between language-related regions

rs7580864 is an intron of ***PLCL1***, associated with **autism** and **GABA signaling pathway**



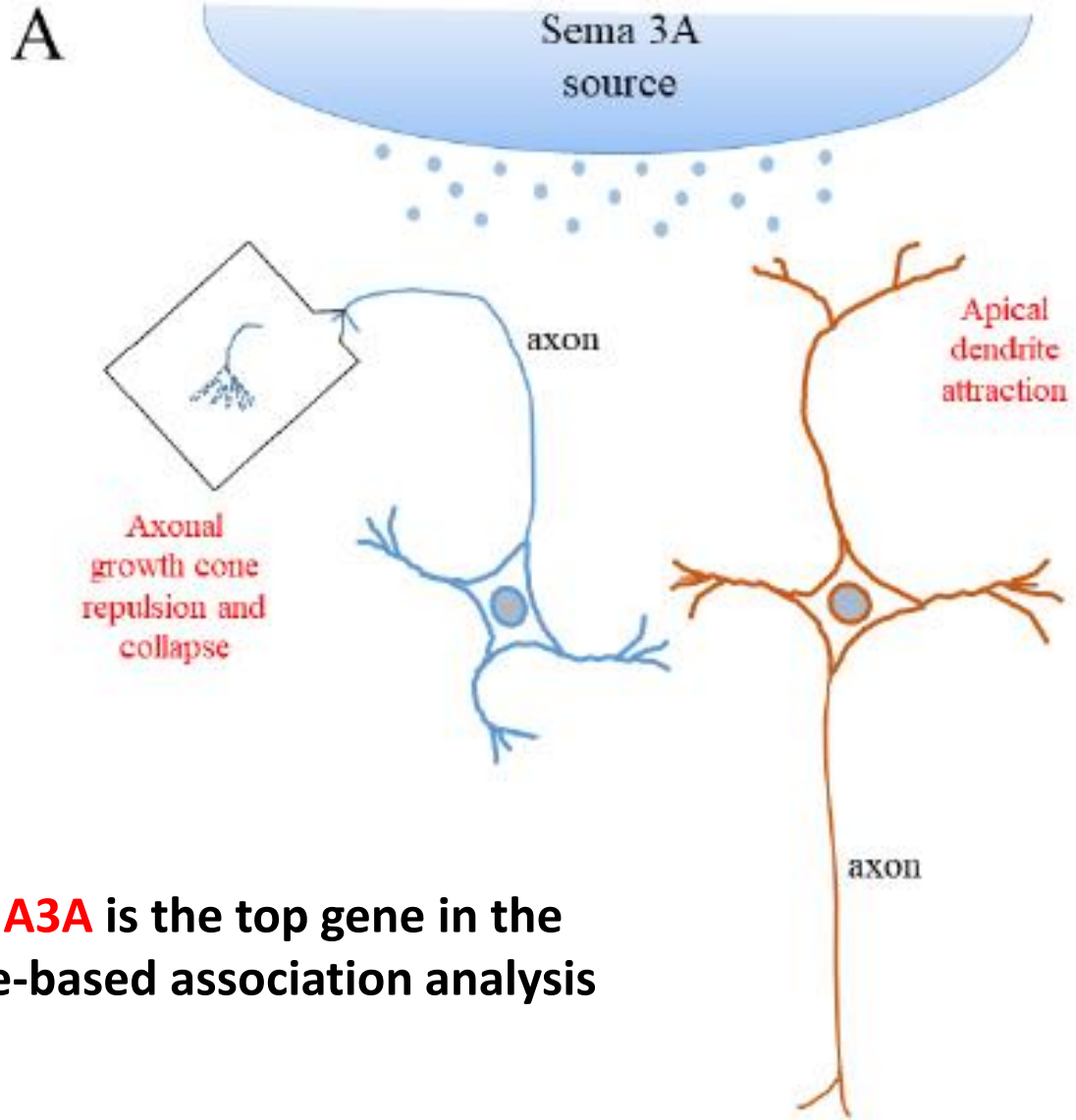


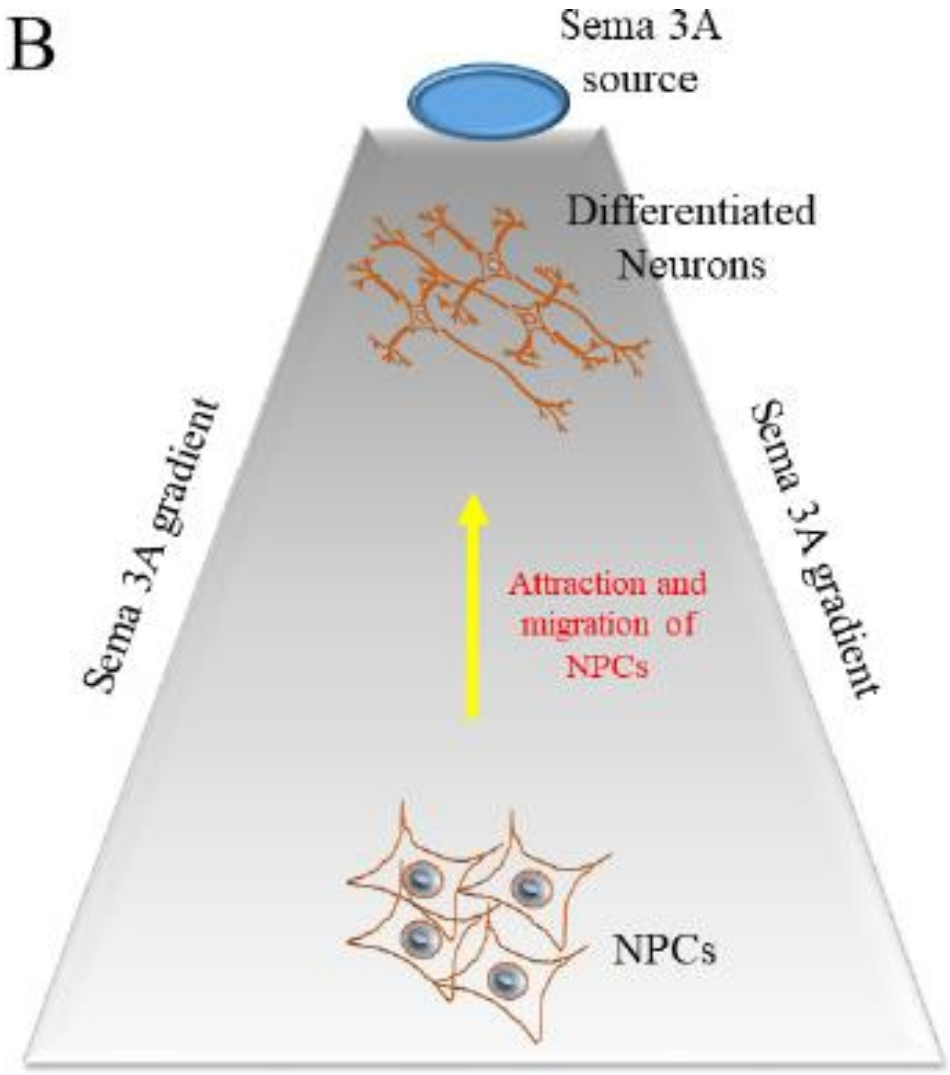
Ortalli et al., 2012; Mekki et al., 2022; Zhang et al., 2021

Discussion

Neural differentiation is important for the white matter network formation.

Axon pruning might shape the precise white matter connectivity.

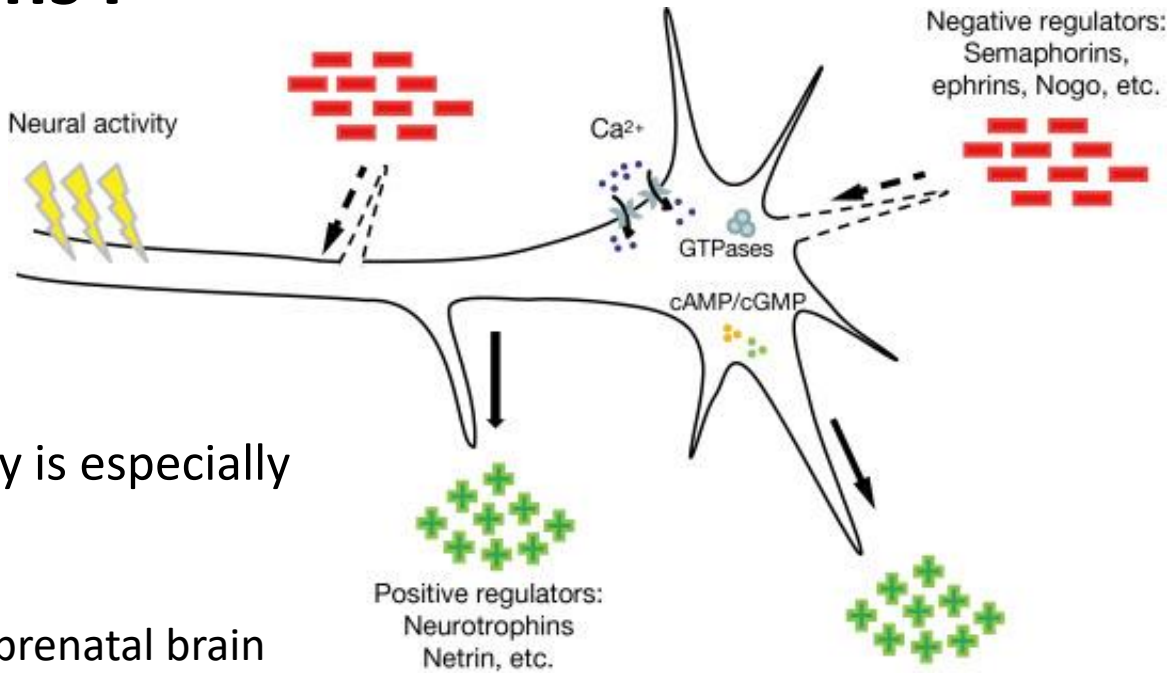




SEMA3A is the top gene in the gene-based association analysis

Tang et al., 2018, Neural Regeneration Research

Conclusions I



- Inter-individual variation in adult white matter connectivity is especially influenced by genes that are:
  - active in the prenatal developing brain
  - upregulated in stem cells, astrocytes, microglia, neurons of prenatal brain
  - involved in neurodevelopmental processes including neural migration, neural projection guidance and axon development

- Roles of glial cells in neurodevelopment are not well understood
  - Astrocytes can express positional guidance cues, e.g. SEMA3A, required for neuronal circuit formation (mediating attraction or repulsion of the growth cone at the axonal tip)
  - Embryonic microglia associate with developing axons and can affect nerve bundle formation

Ma & Gibson (in Rubenstein & Rakic 2013)

Conclusions II

- Polygenic scores for various psychiatric and neurological disorders showed significant associations with white matter connectivity
  - Each implicating distinct sets of brain regions with trait-relevant functional profiles
- Polygenic risk for disorders likely to manifest:
  - partly through affecting the development of large-scale structural brain networks
  - particularly during prenatal brain development.

Acknowledgements

Clyde Francks, Simon E. Fisher and Dick Schijven

UK Biobank Imaging team (Steve Smith and colleagues)

Thank you for your attention!