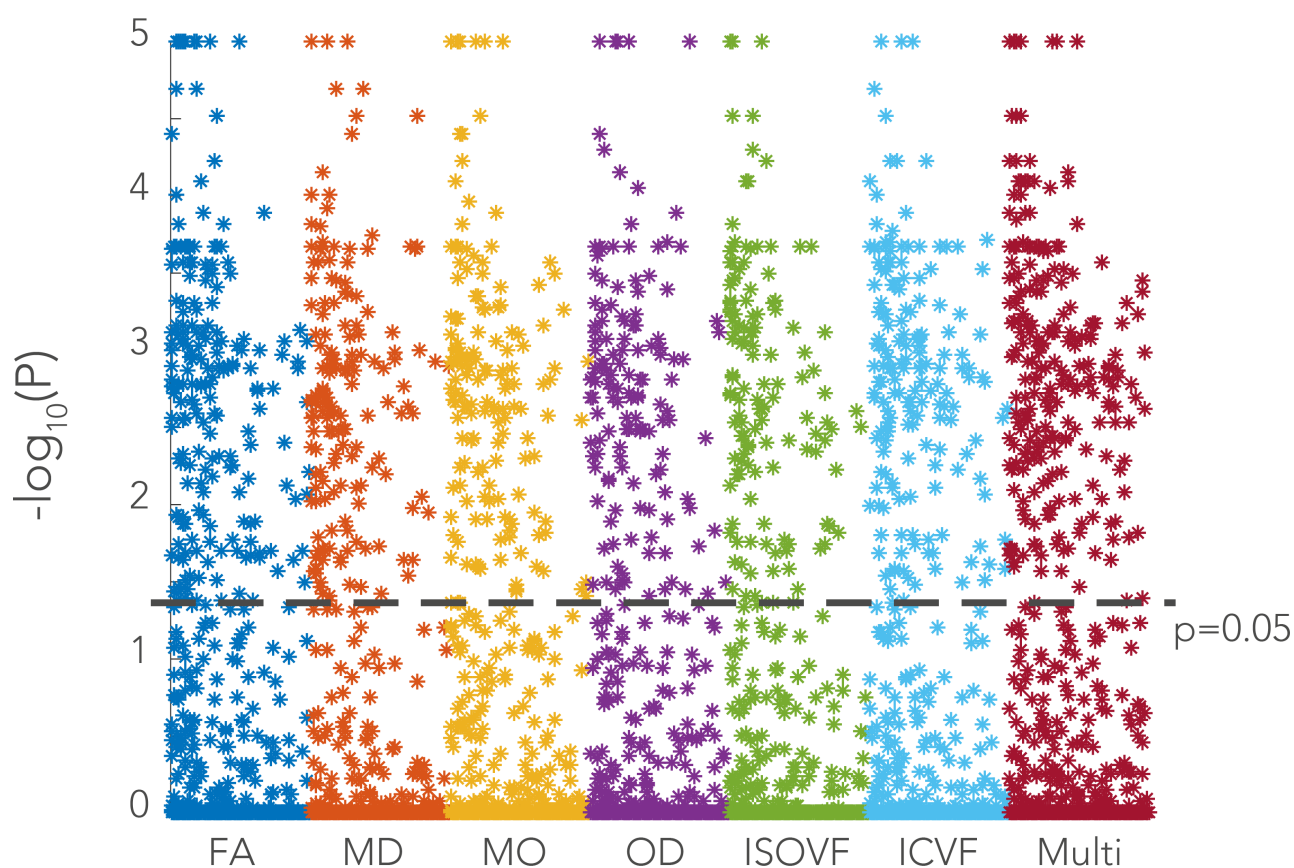


Supplementary Figure 1.

Effect sizes varying with the number of microstructural principal components.

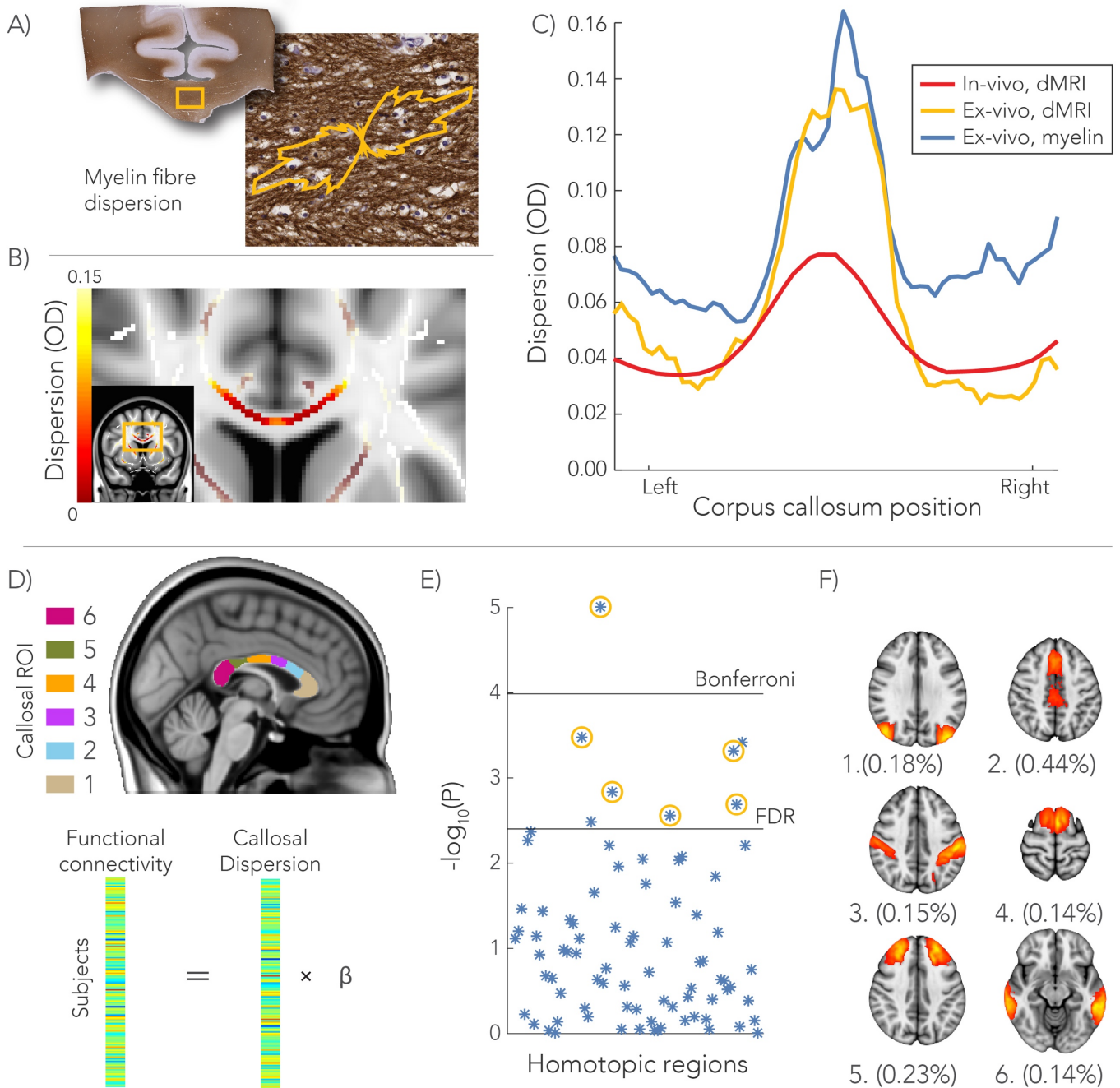
Percentage variance explained by the multimodal regression model – functional connectivity predicted by microstructure – as a function of the number of principal components used as regressors (in blue), averaged across the homotopic region pairs. The green line represents the total variance explained by the models in which the rows (i.e., the subject entries) of the design matrix are randomly permuted. The dashed line indicates the number of principal components used in this study ($p = 30$).



Supplementary Figure 2.

Manhattan plot showing the p-values for the microstructure-function associations in the training cohort (n = 7481 subjects) depicted for each microstructural metrics separately.

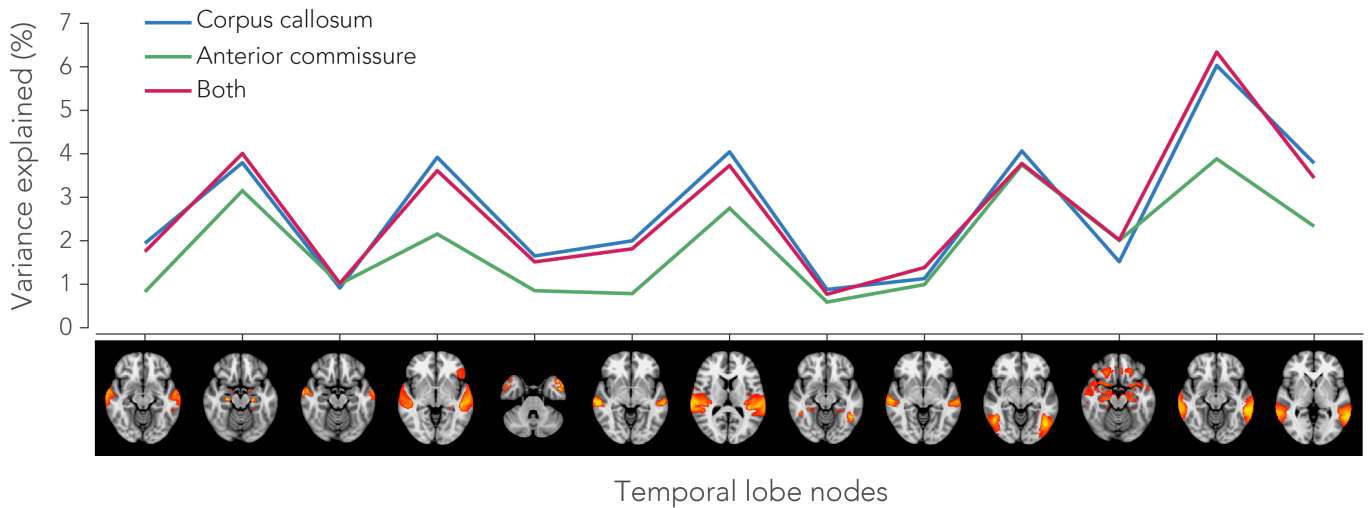
Significance was determined using permutation testing (n = 100,000 permutations, two-sided). Each dot represents a regressor's beta value that corresponds to a microstructural principal component in the regression models expressed as $-\log_{10}(P)$. All p-values are corrected for the family-wise-error, with $p < 0.05$ is considered to be significant (see Fig. 3).



Supplementary Figure 3.

Validation of fiber dispersion in the corpus callosum and its relation to interhemispheric functional connectivity.

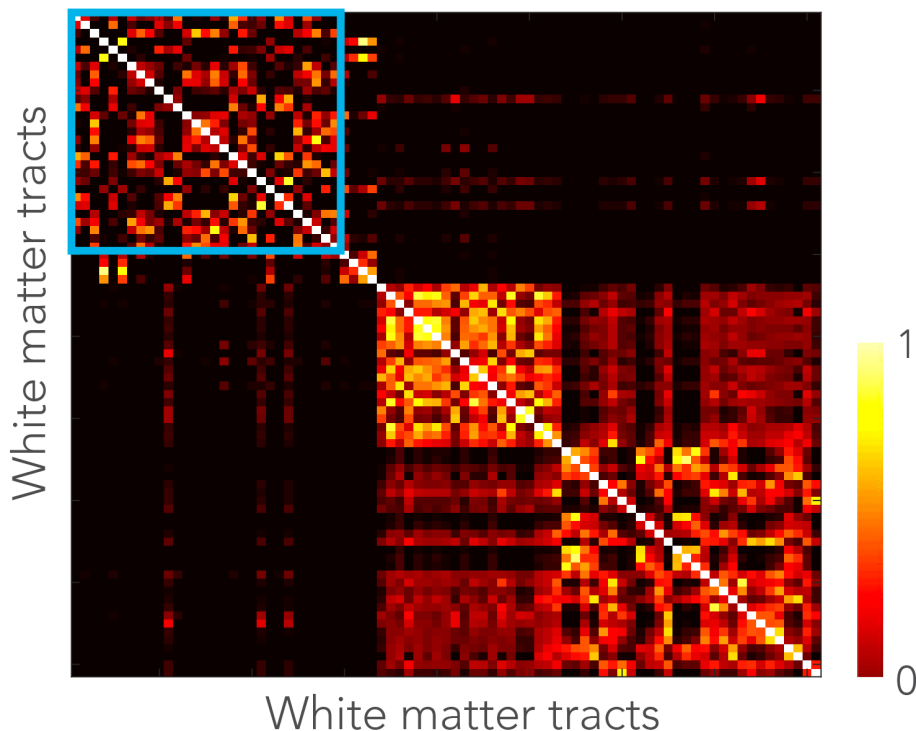
Ex-vivo brain specimens ($n=3$) were scanned for dMRI and subsequently histologically stained for myelin (proteolipid protein, PLP). The pattern of fiber dispersion in the corpus callosum correlates well with dMRI dispersion of both the ex-vivo specimens as well as the in-vivo subjects from UK Biobank. For six homotopic pairs, functional connectivity was found to have a significant association with orientation dispersion in the relevant midline callosal ROI. **A)** Fiber dispersion was obtained from the histological myelin sections using texture analysis to evaluate fiber dispersion obtained from dMRI within the same specimen. **B)** OD map averaged across the in-vivo subjects in the corpus callosum projected onto the white matter skeleton. **C)** Dispersion profiles in the corpus callosum from the ex-vivo datasets in comparison to in-vivo subjects from UK Biobank. **D)** Prediction of functional connectivity from midline callosal dispersion derived from the in-vivo subjects. Dispersion was extracted from the callosal region having the highest overlap with the tract running between the homotopic regions of interest (ROIs). **E)** Manhattan plot showing the association (linear regression with permutation testing, two-sided) between callosal dispersion and functional connectivity for all homotopic regions in the training cohort ($n = 7481$ subjects). Dots above the false discovery rate (FDR) threshold are considered to be significant. The yellow circles also survived the negative control analysis (i.e. strongest correlation with dispersion from the anatomical correct callosal ROI in comparison to all other callosal ROIs). **F)** Spatial ICA maps of the homotopic regions corresponding to the yellow circles (E) with the percentage of functional connectivity variance explained.



Supplementary Figure 4.

Effect sizes for the temporal lobe brain regions.

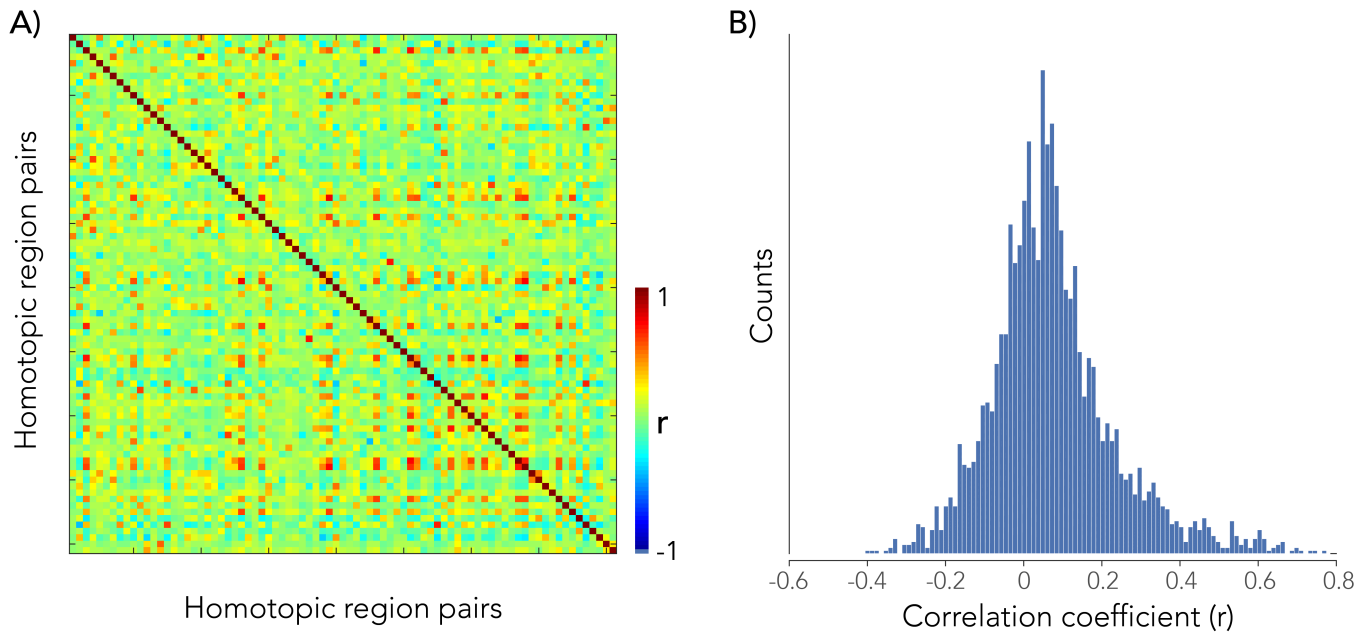
Percentage variance explained in functional connectivity between temporal lobe regions by the microstructural signature from either the corpus callosum, the anterior commissure or both. While temporal lobe regions are also connected through the anterior commissure, microstructural information from this pathway is in general not better in explaining temporal lobe functional connectivity than callosal microstructure, nor does a combination of both pathways.



Supplementary Figure 5.

Spatial overlap among the 81 white matter tracts.

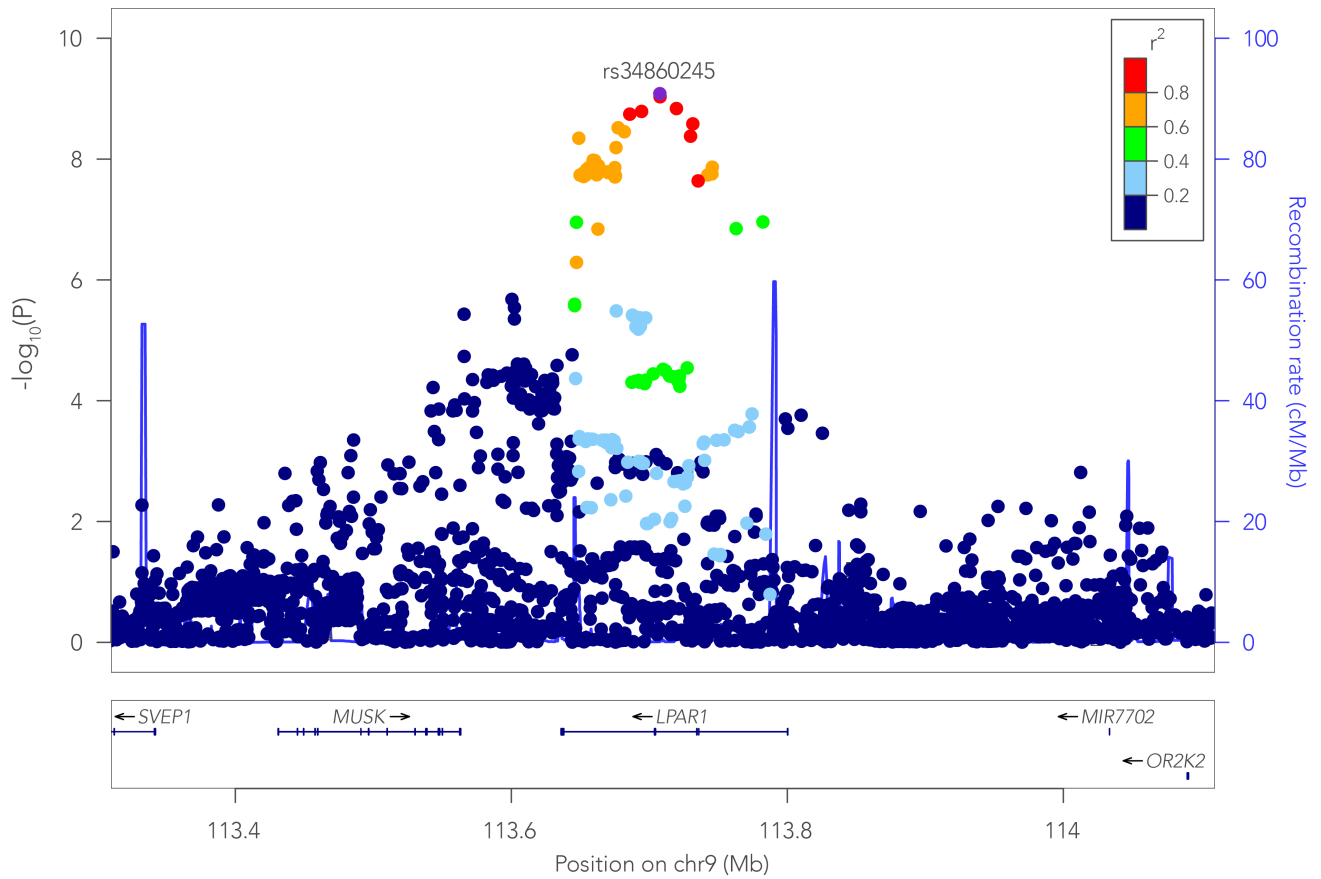
Each entry expresses the Dice similarity index between two white matter tracts. An index of 0 indicate no spatial overlap at all, while tract pairs with an index of 1 are perfectly overlapping. The tracts are sorted in this matrix after k-means clustering. The light-blue rectangle indicates the 30 tracts that had minimal spatial overlap and were used as control ("wrong") tracts in the negative control analysis.



Supplementary Figure 6.

Correlation between model fits using the multimodal microstructure model to predict functional connectivity.

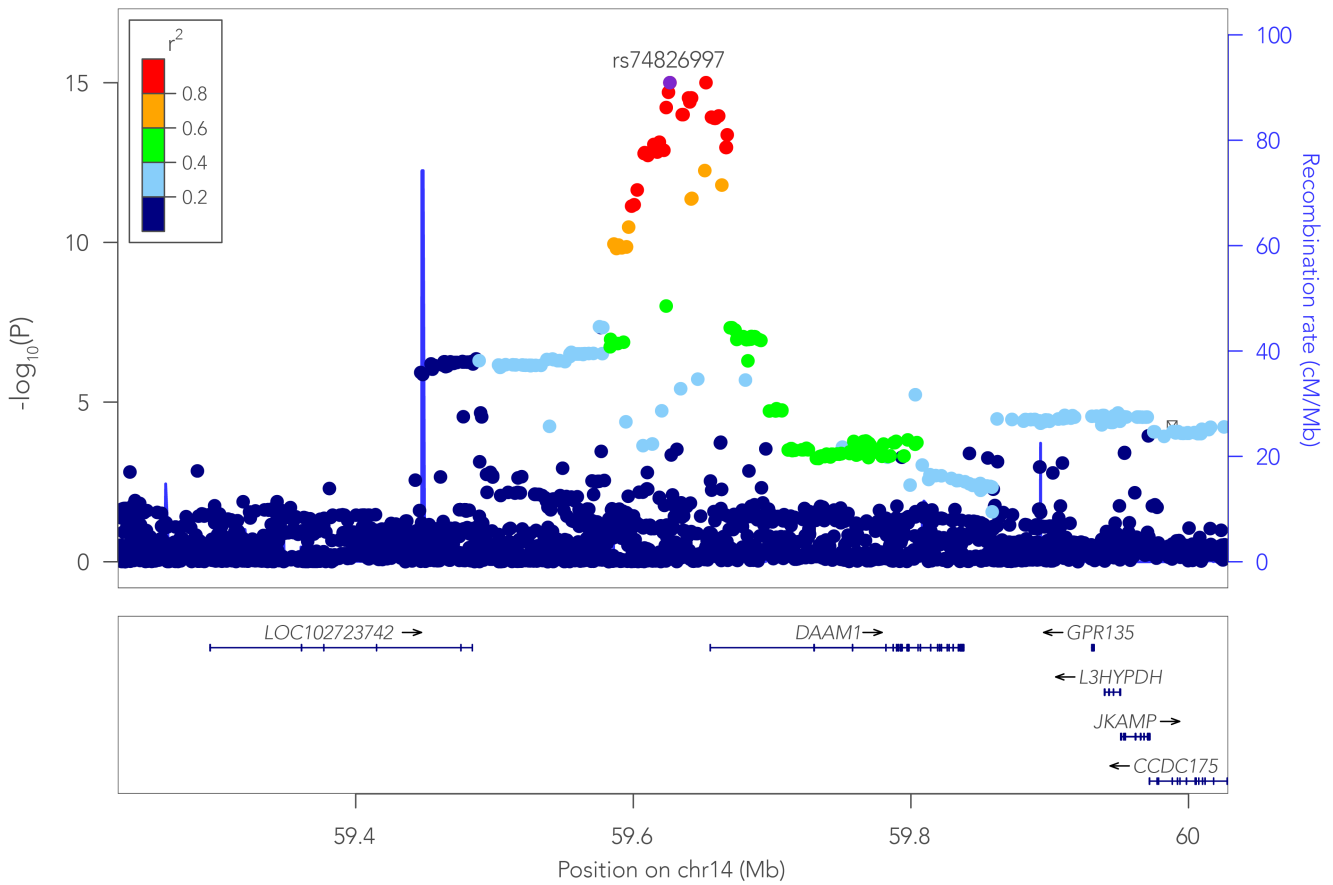
A) Correlation matrix depicting the Pearson correlation coefficient of the model fits between all homotopic region pairs. **B)** Histogram of the off-diagonal correlation coefficients.



Supplementary Figure 7.

Genetic associations around the LPAR1 gene.

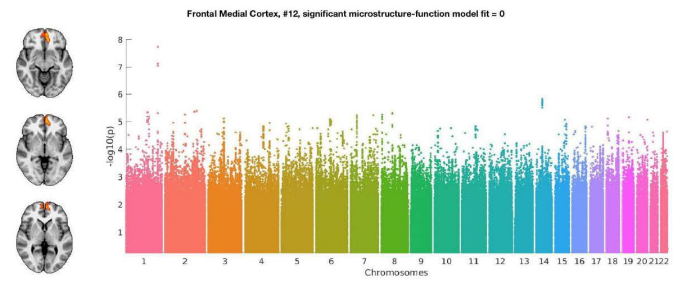
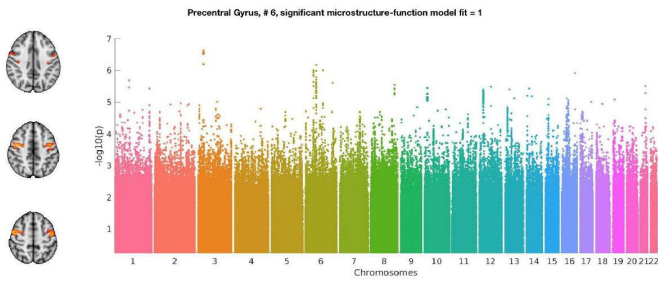
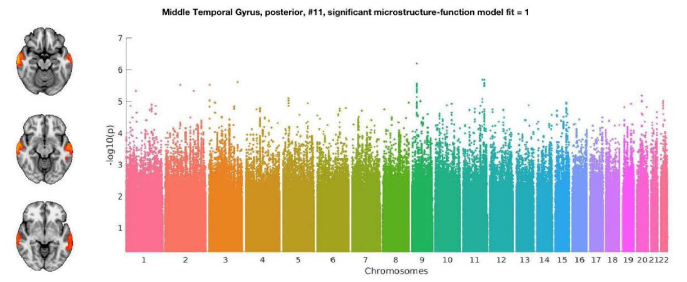
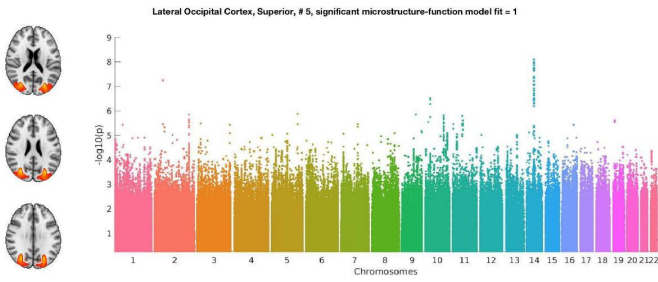
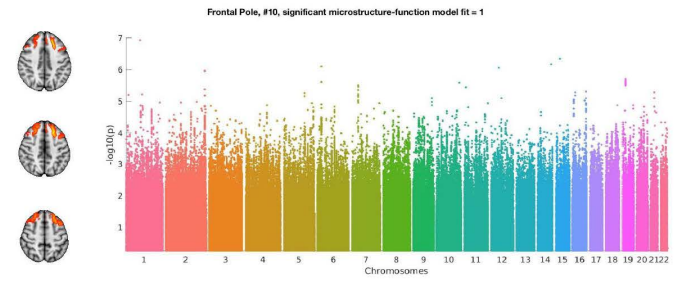
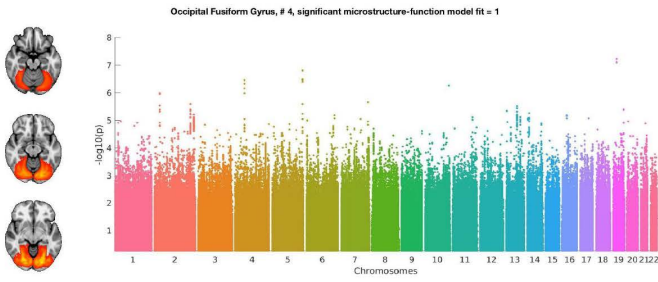
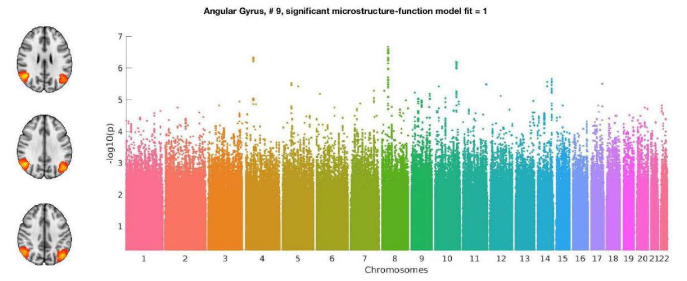
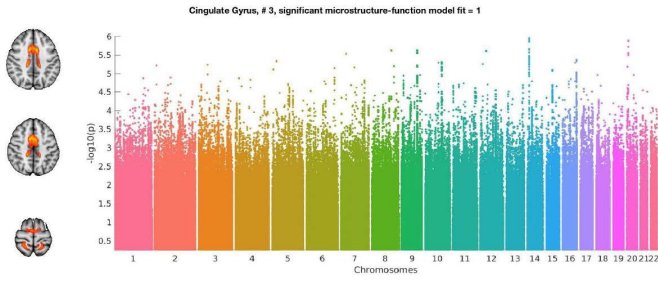
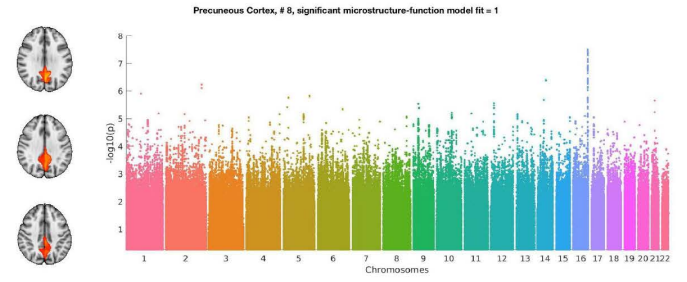
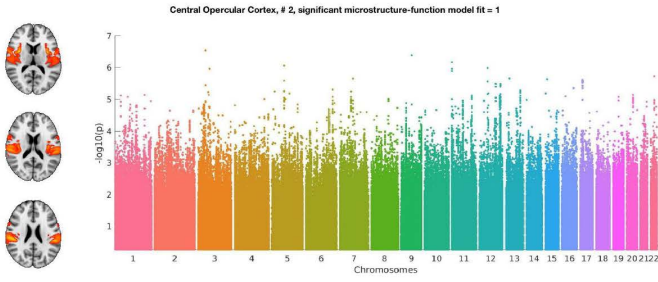
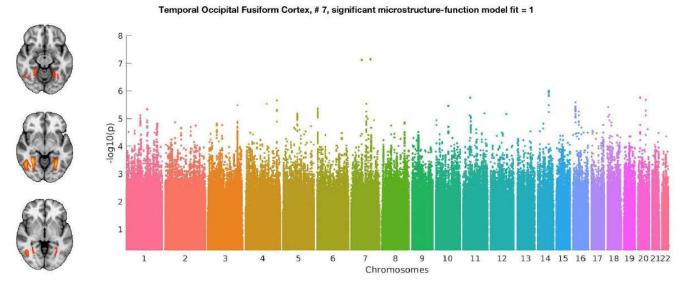
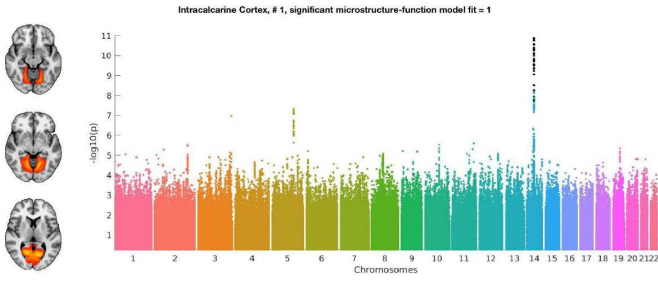
Genetic association (linear regression, two-sided) with the microstructure-function phenotype (i.e. the pattern of functional connectivity that can be predicted from white matter microstructure) centered around SNP rs34860245 in chromosome 9 for the discovery cohort ($n = 7481$ subjects). Points are colour-coded by the local linkage disequilibrium with the top hit SNP (in purple). Associations were estimated using univariate regressions, two-sided). The significance threshold was set to a $-\log_{10}(p\text{-value})$ equal to 7.5 corresponding to a $p\text{-value}$ of $\sim 3 \times 10^{-8}$.

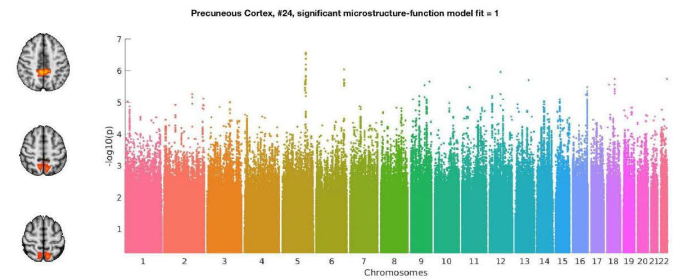
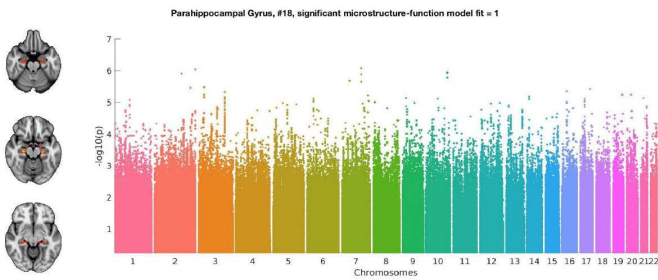
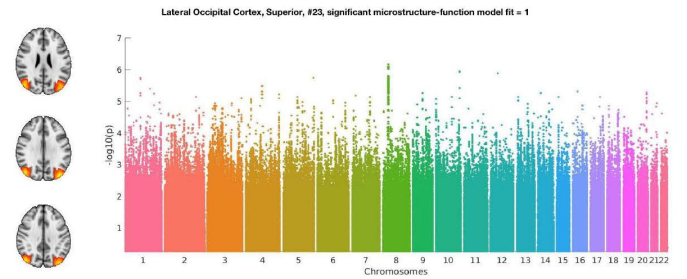
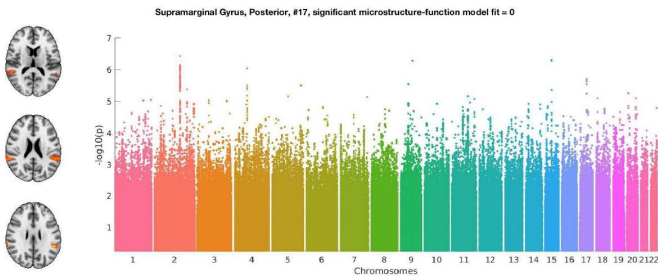
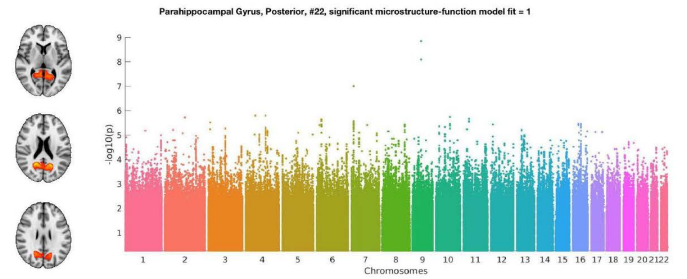
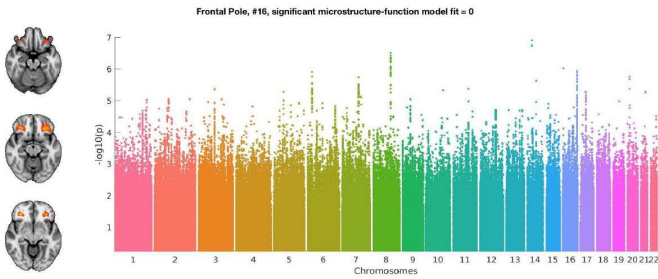
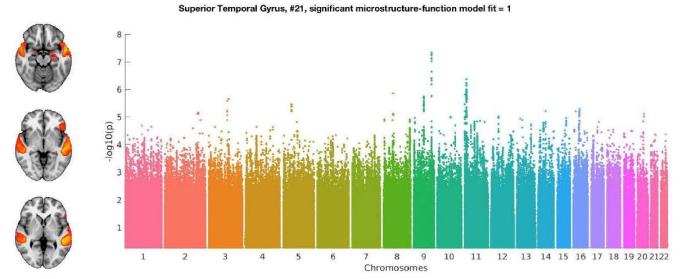
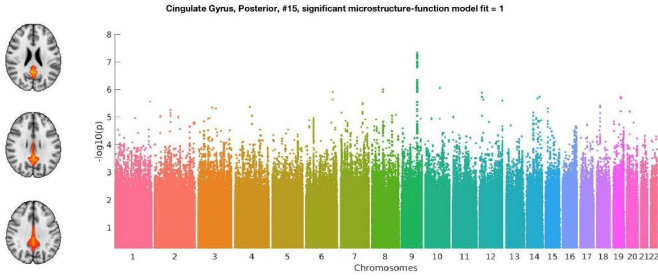
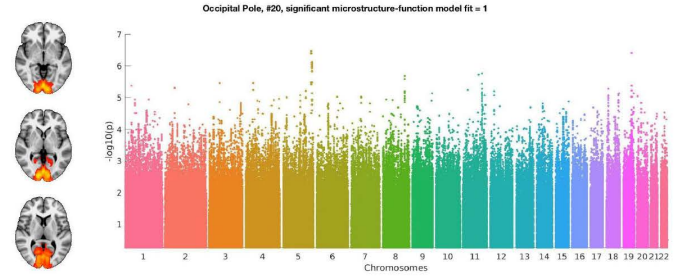
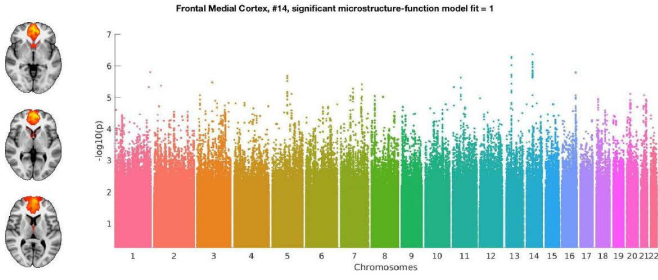
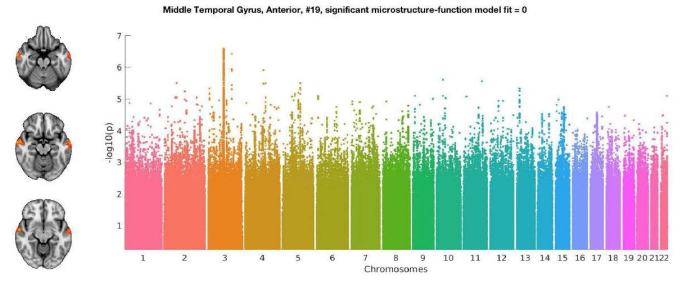
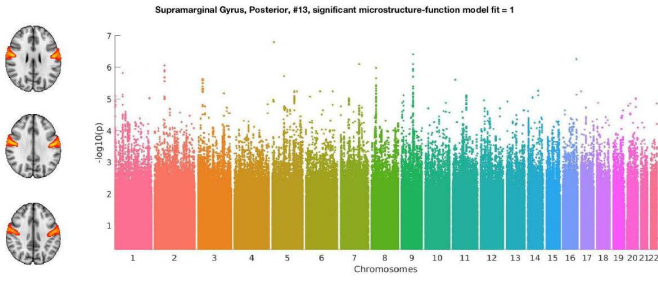


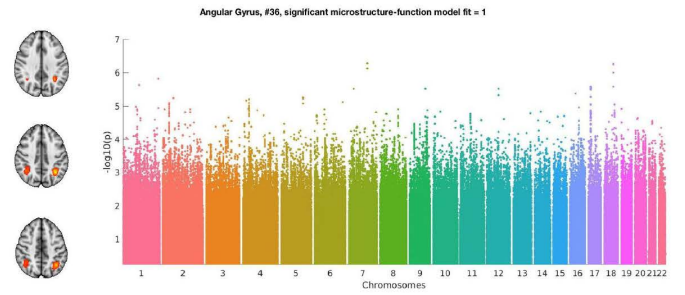
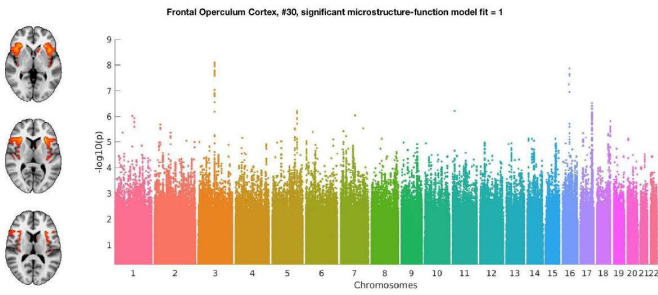
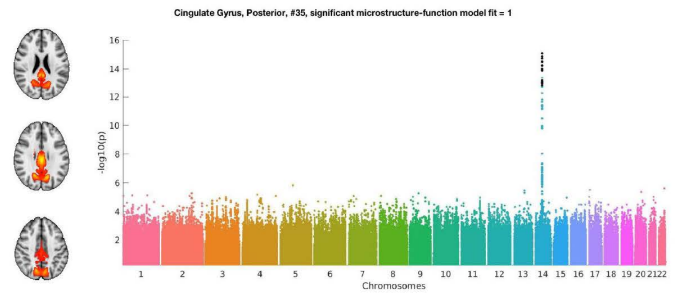
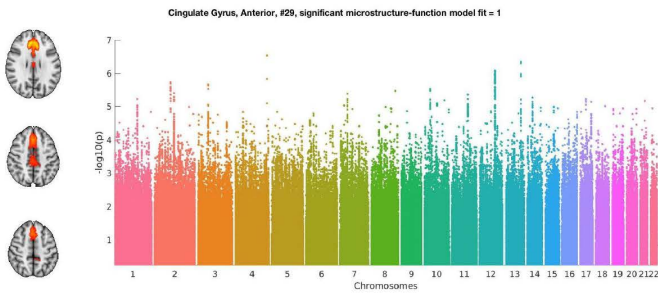
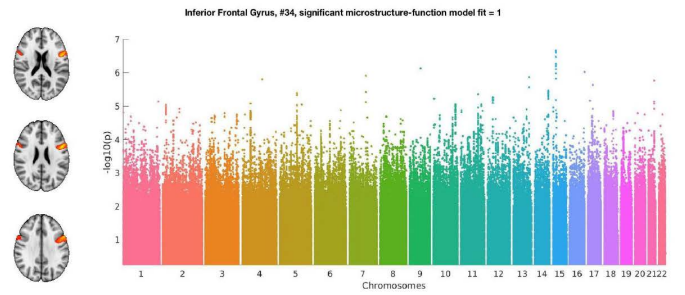
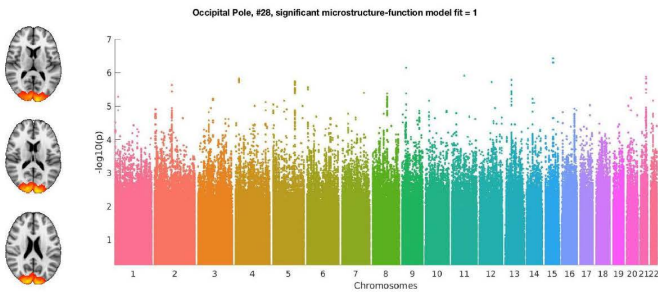
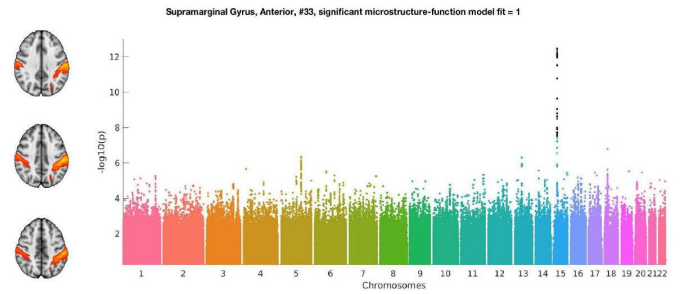
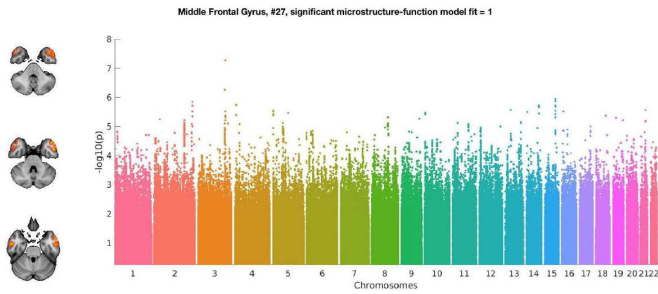
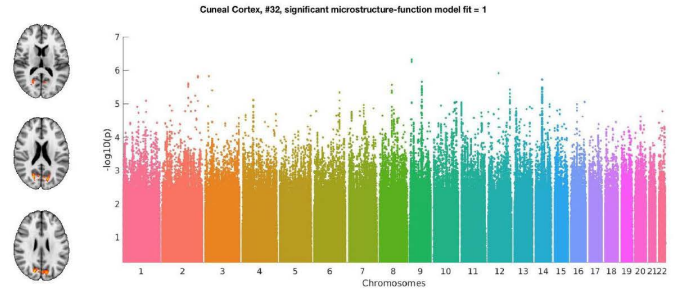
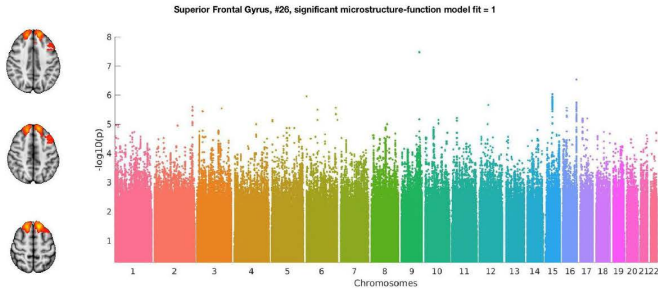
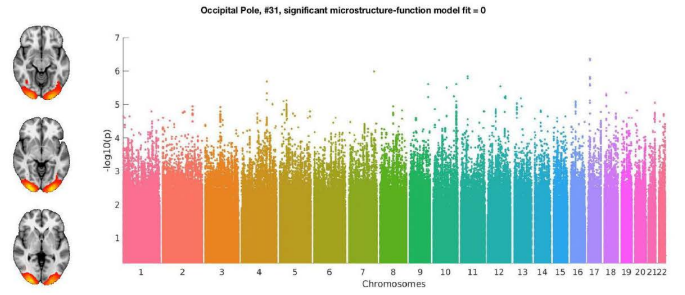
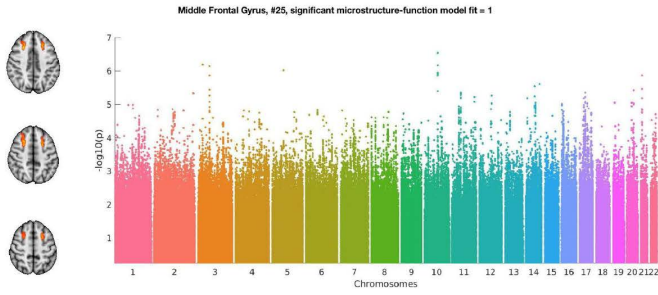
Supplementary Figure 8.

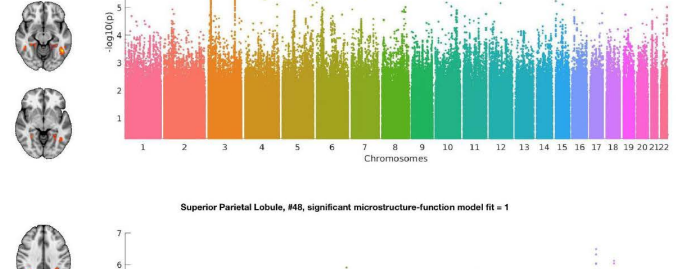
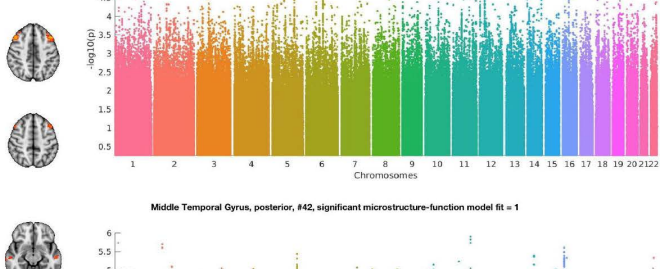
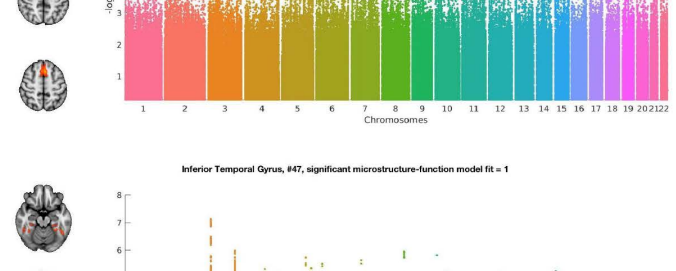
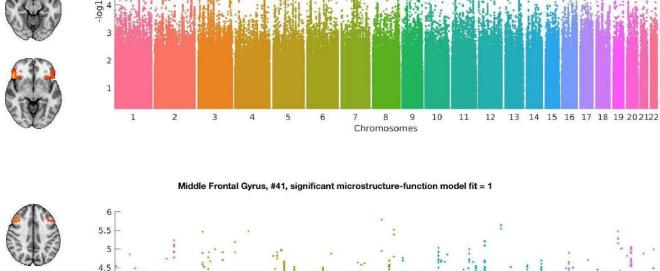
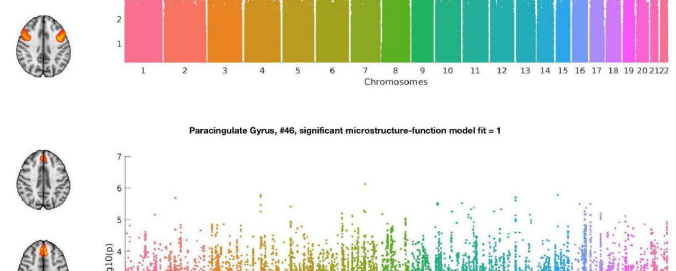
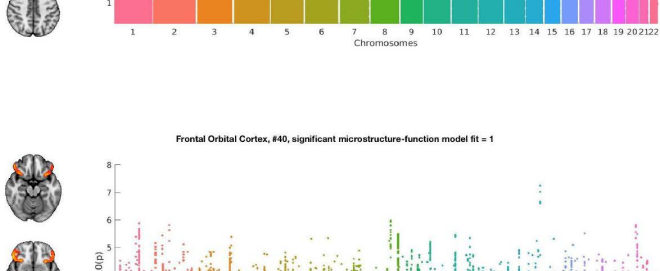
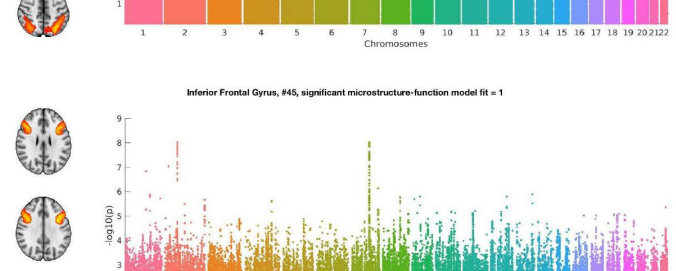
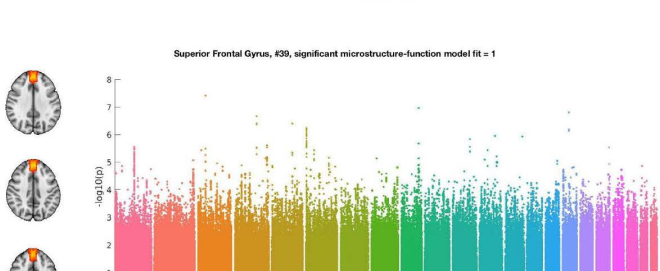
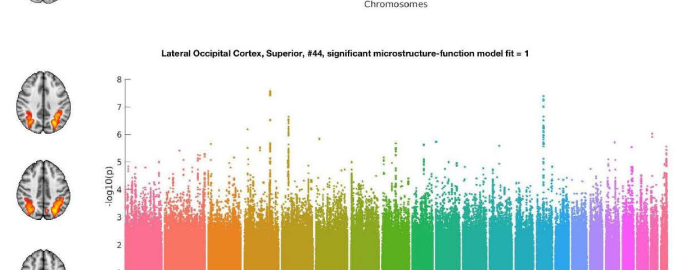
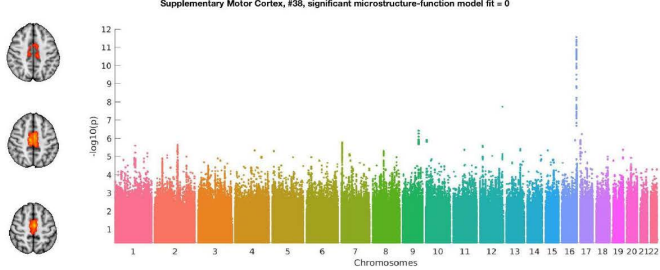
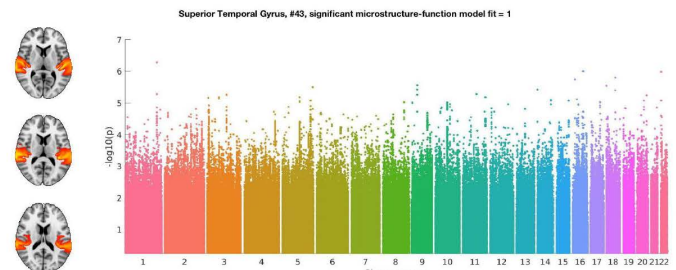
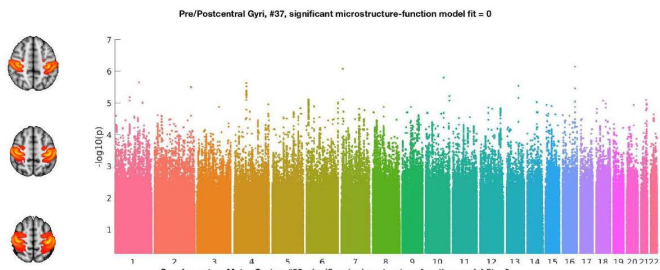
Genetic associations around the DAAM1 gene.

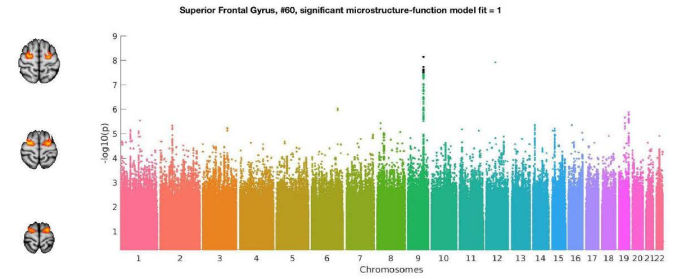
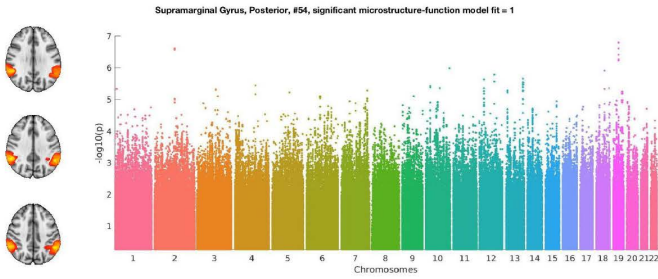
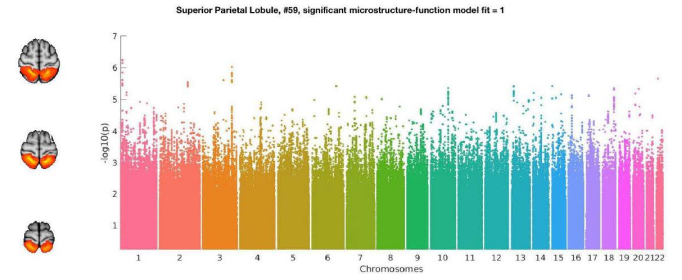
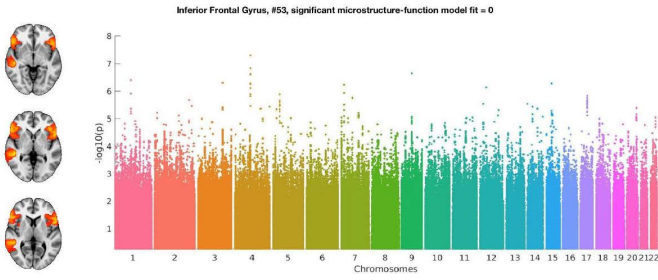
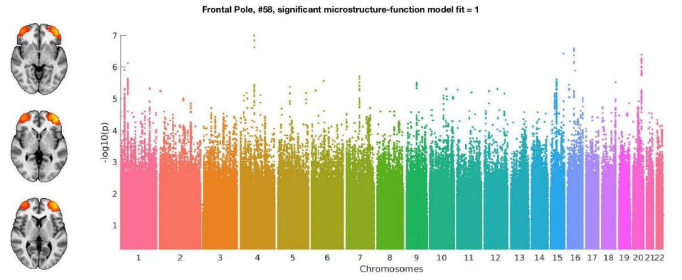
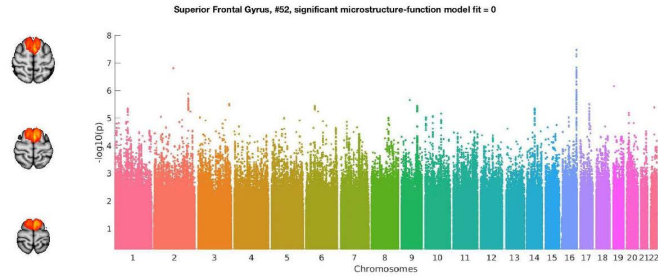
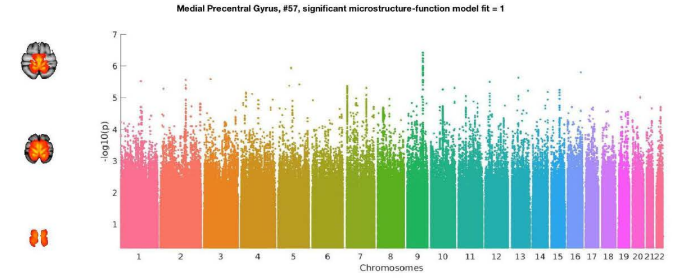
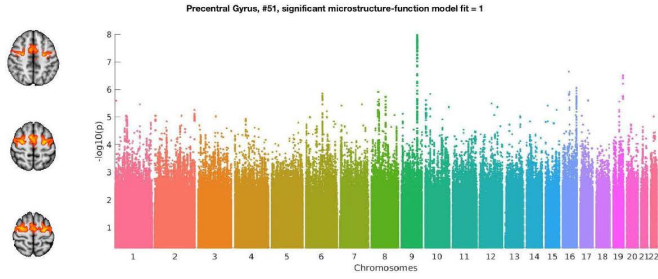
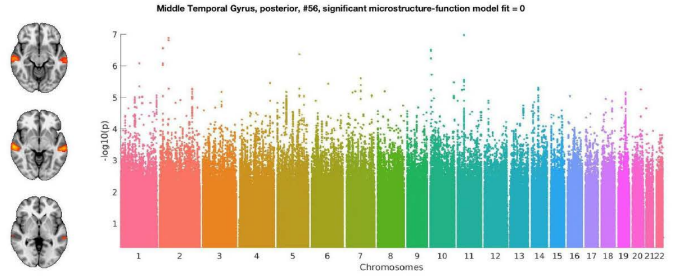
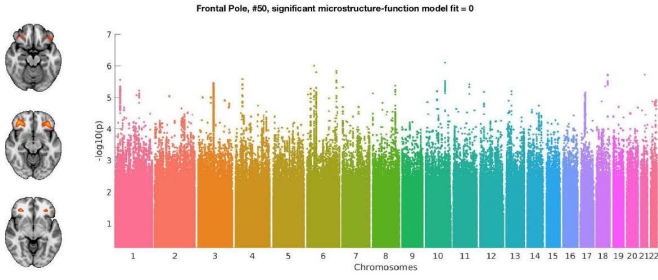
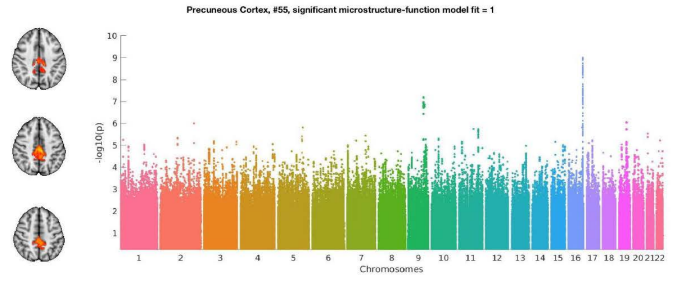
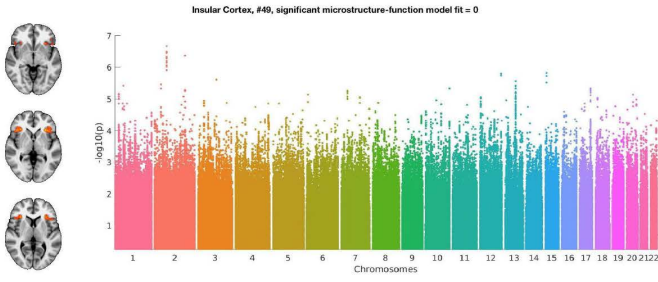
Genetic association (linear regression, two-sided) with the microstructure-function phenotype (i.e. the pattern of functional connectivity that can be predicted from white matter microstructure) centred around SNP rs74826997 in chromosome 9 for the discovery cohort ($n = 7481$ subjects). Points are colour-coded by the local linkage disequilibrium with the top hit SNP (in purple). The significance threshold was set to a $-\log_{10}(p\text{-value})$ equal to 7.5 corresponding to a $p\text{-value}$ of $\sim 3 \times 10^{-8}$.

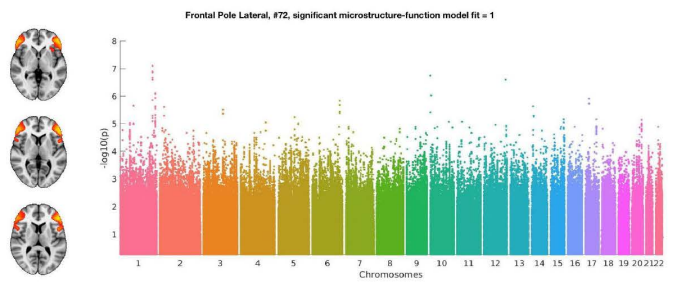
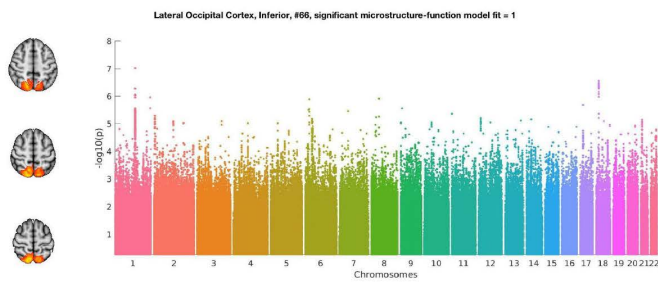
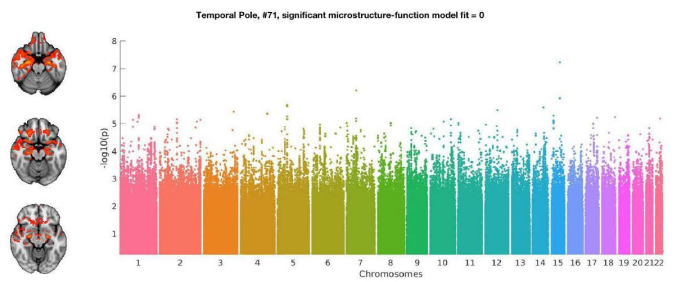
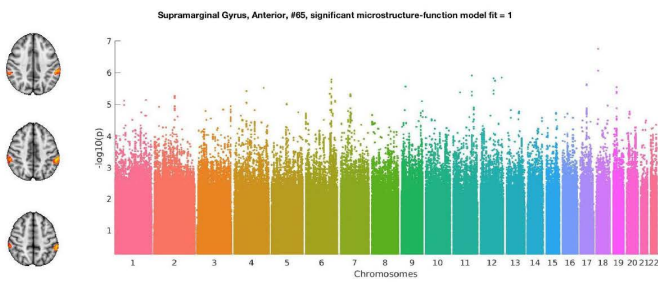
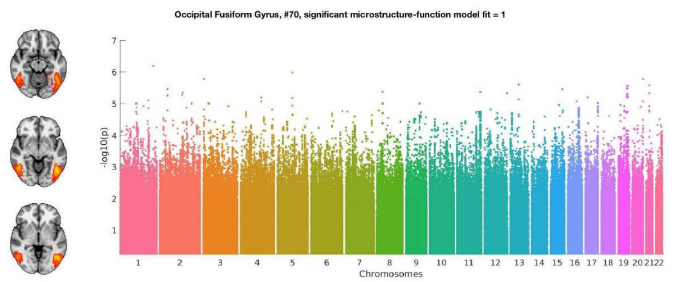
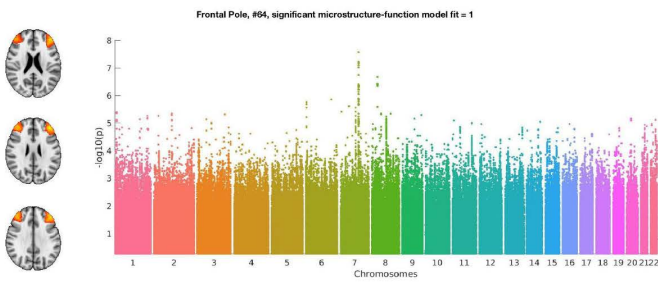
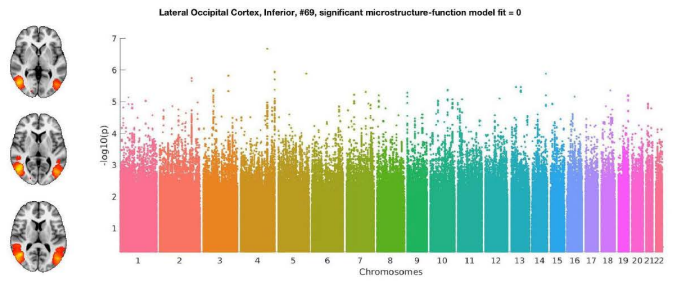
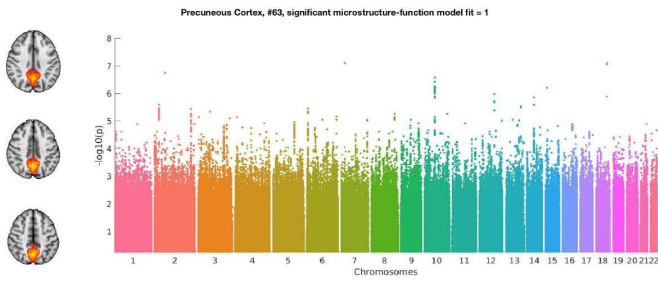
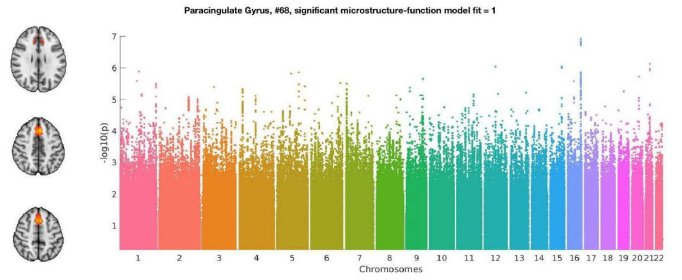
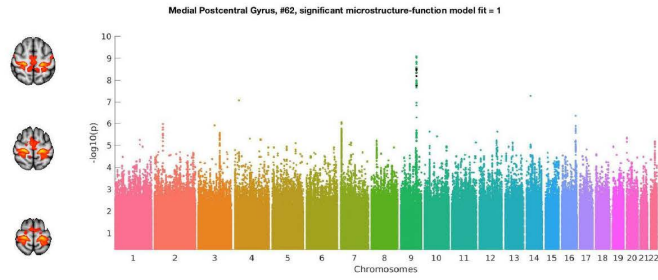
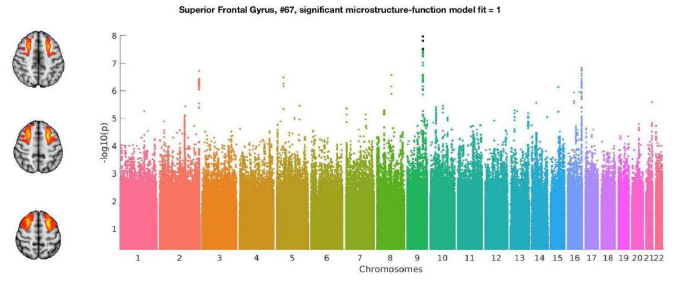
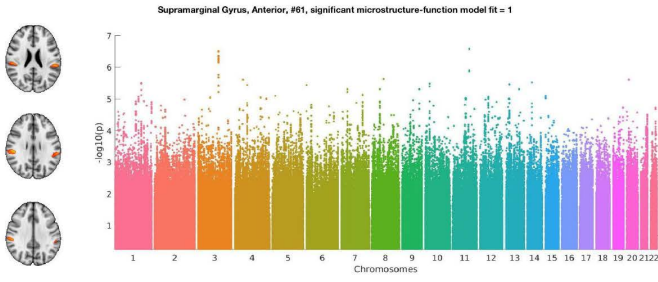










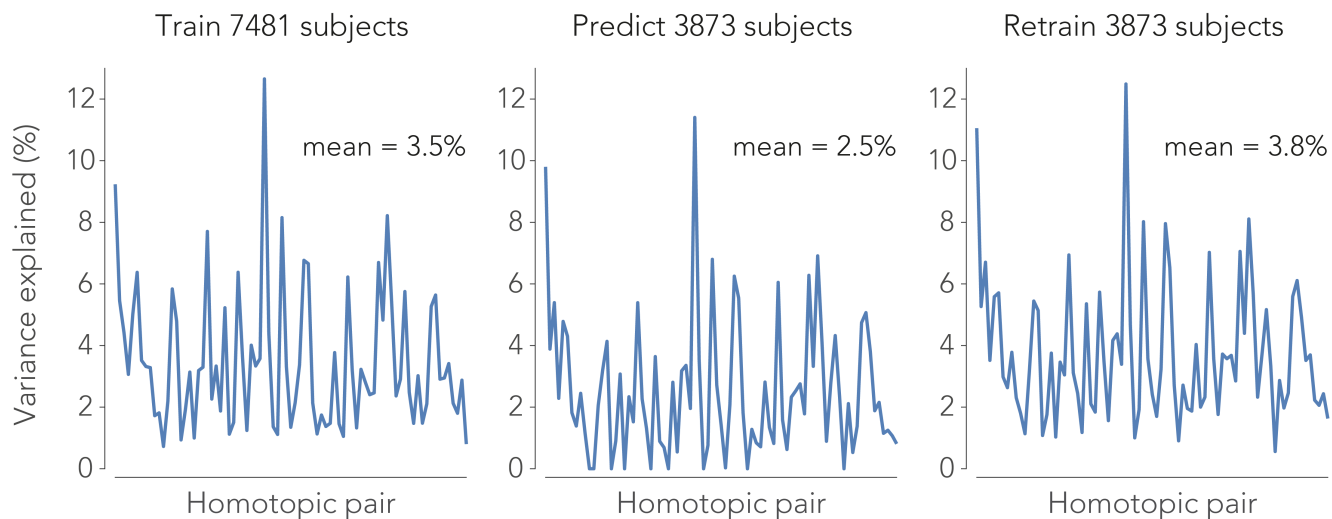




Supplementary Figure 9

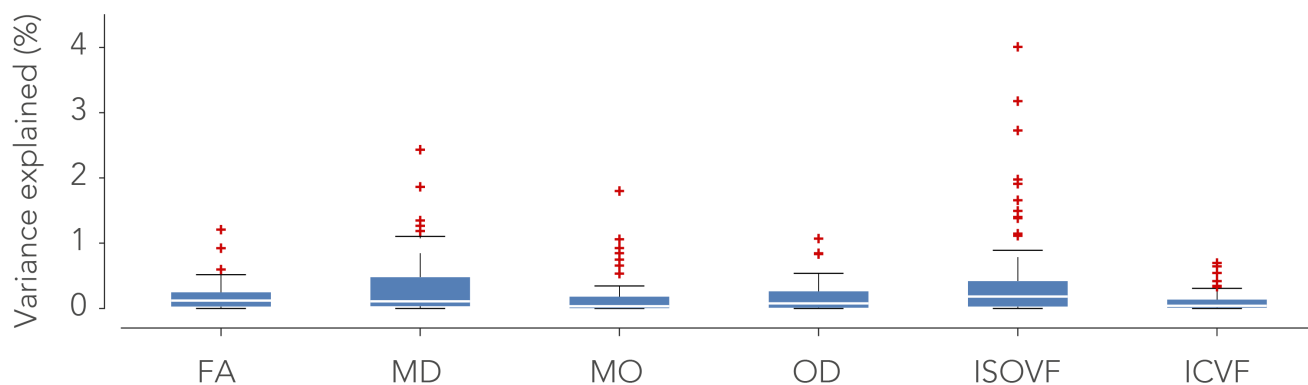
Genome-wide association results for all homotopic region pairs.

Genome-wide associations (linear regression, two-sided) with the microstructure-function phenotype (i.e. the pattern of functional connectivity that can be predicted from white matter microstructure). The Manhattan plots depict the associations with each SNP across all chromosomes expressed as the $-\log_{10}(p\text{-value})$. Each Manhattan plot shows the association of the microstructure-function model fit of a homotopic region pair for the discovery cohort (n = 7481 subjects). The significance threshold was set to a $-\log_{10}(P)$ equal to 7.5 corresponding to a p-value of $\sim 3 \times 10^{-8}$.



Supplementary Figure 10.

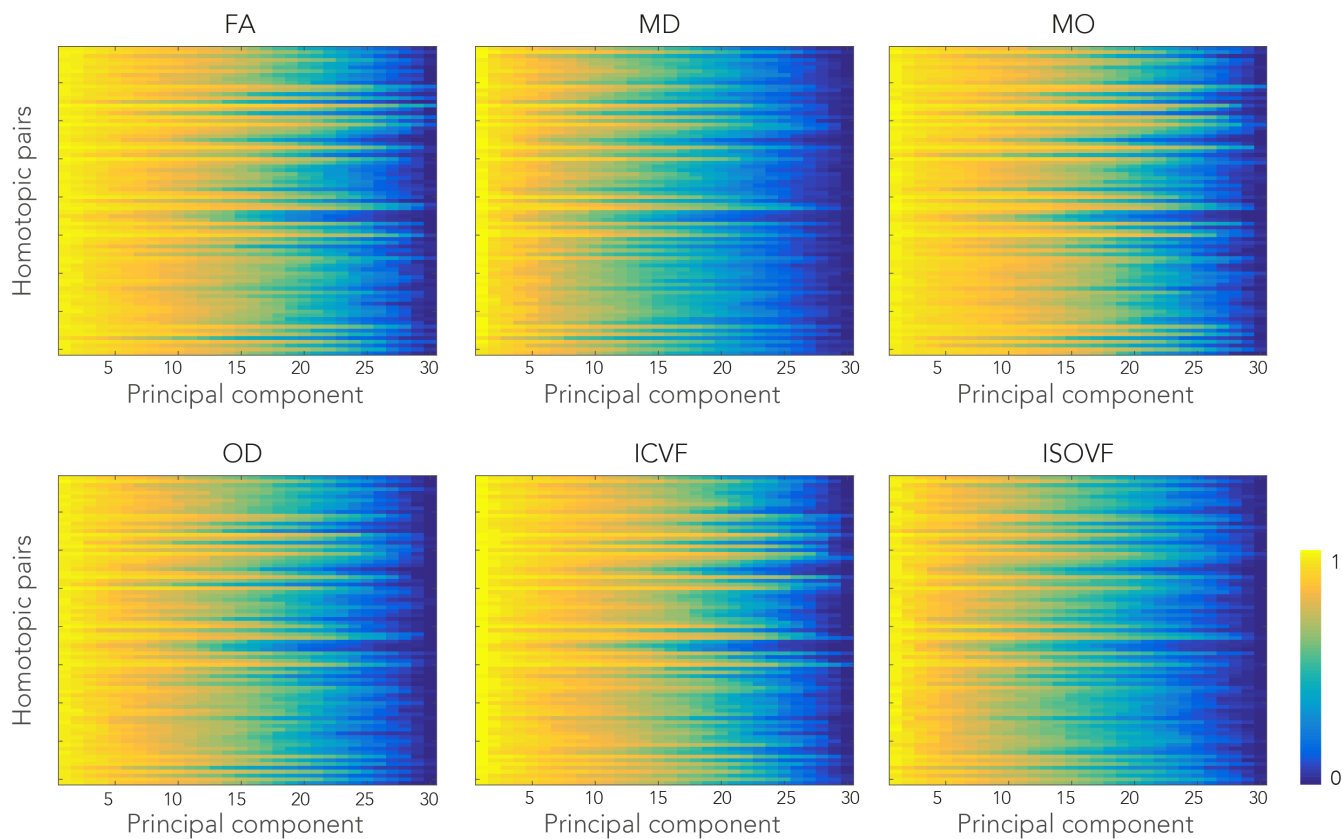
Effect of cohort size and retraining models on the percentage variance explained by the multi-modal microstructure model.



Supplementary Figure 11.

Univariate modelling of white matter microstructure to predict functional connectivity.

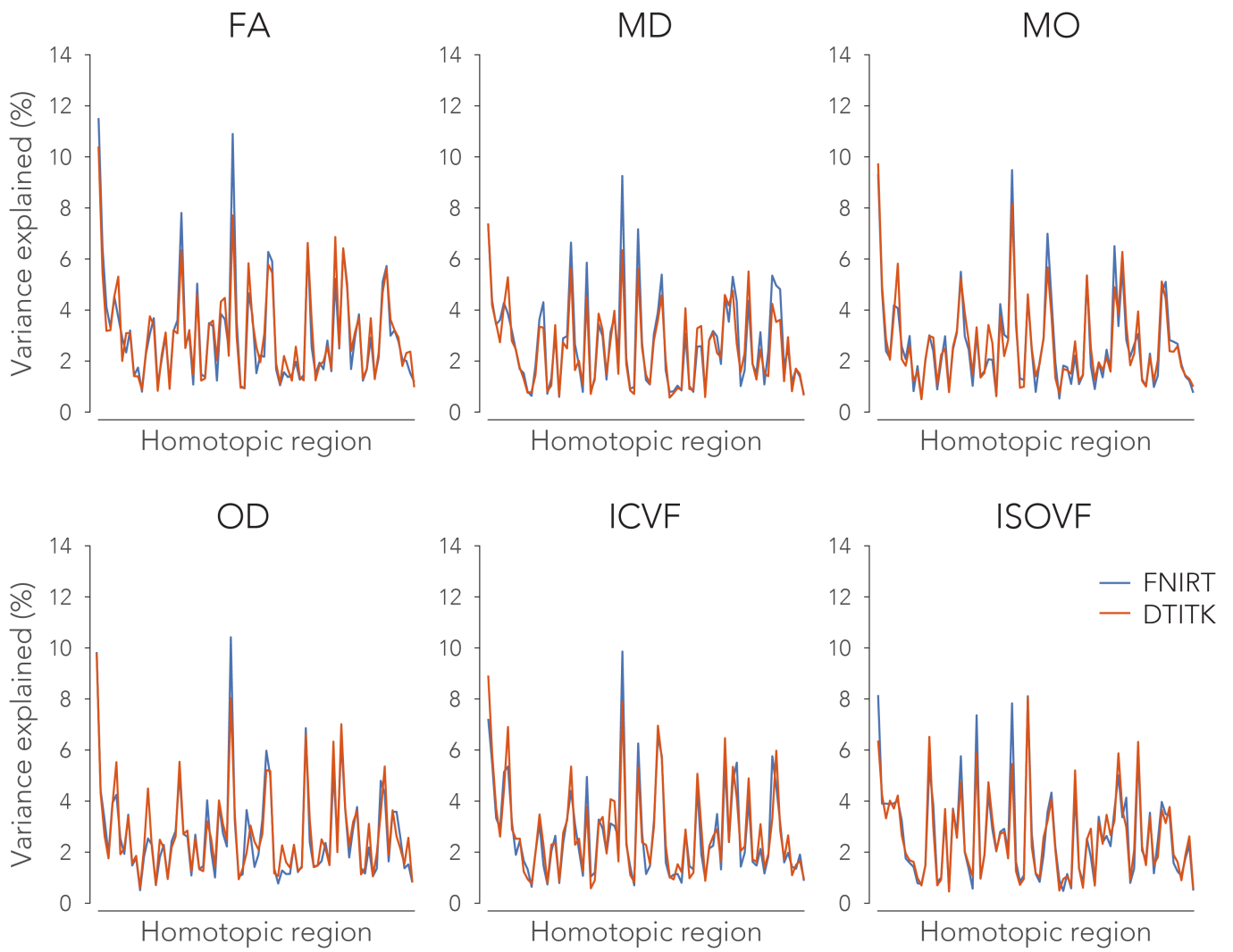
Boxplots of percentage variance explained in functional connectivity between the homotopic region pairs by the microstructural mean of a white matter tract as opposed to the rich microstructural signature generated using principal components analysis for the training cohort ($n = 7481$ subjects). The centre line depicts the median variance explained values across the homotopic region pairs; box limits, the 25th and 75th percentiles; the whiskers extend to the most extreme data points excluding outliers (marked with a + symbol).



Supplementary Figure 12.

Canonical correlation analysis between microstructural principal components derived from two registration methods to align the dMRI volumes.

7481 subjects were aligned using either FNIRT-based FA registration or using DTITK registration incorporating the full tensor. After applying the warfields to each microstructural volume, the microstructural metrics were extracted from the tract skeletons connecting a given homotopic region pair for all subjects. Next, a principal components analysis was performed on these microstructural matrices (see Fig. 2 for an overview). To evaluate whether FNIRT and DTITK yield principal components that carry the same information, a canonical correlation analysis was performed to maximize the overlap between the two subspaces. Each entry in the matrices above represents the correlation coefficient (i.e., the cosine of the principal angle) between the principal components generated FNIRT and DTITK for a given white matter tract. Especially the top principal components show a high degree of similarity, whereas principal components carrying less variance are probably more susceptible to noise or mis-registration. The effect of both registration approaches on model performance is given in Supplementary Fig. 13.



Supplementary Figure 13.

Model performance for the two registration approaches; optimized FNIRT and DTITK.

In terms of percentage variance explained in functional connectivity by white matter microstructure, the difference between the registration algorithms to align the dMRI data is negligible for all microstructural metrics. See Supplementary Fig. 12 caption for more information.

Supplementary Table 1.

Brain areas corresponding to the homotopic regions according to the Harvard-Oxford cortical structural atlas. The spatial maps can be found in Figure 1.A. Percentage variance explained – for the training and replication cohort – by the multimodal microstructure models to predict functional homotopic connectivity is given for each area. The coordinates of the brain area in the left hemisphere is given for the Montreal Neurological Institute (MNI) 2 mm standard space atlas. Figure 5 depicts the explained variance on a brain surface for each of the homotopic region pairs.

Area	Name	MNI (x,y,z)	Variance explained (%)		Area	Name	MNI (x,y,z)	Variance explained (%)	
			Training	Replication				Training	Replication
1	Intracalcarine Cortex	51, 26, 42	9.24	9.79	42	Middle Temporal Gyrus, posterior	74, 51, 33	2.14	0.03
2	Central Opercular Cortex	70, 57, 43	5.46	3.88	43	Superior Temporal Gyrus	71, 52, 40	3.38	2.07
3	Cingulate Gyrus	49, 57, 61	4.42	5.38	44	Lateral Occipital Cortex, Superior	61, 33, 58	6.77	6.24
4	Occipital Fusiform Gyrus	57, 25, 34	3.06	2.28	45	Inferior Frontal Gyrus	66, 70, 52	6.66	5.53
5	Lateral Occipital Cortex, Superior	59, 26, 50	5.00	4.78	46	Paracingulate Gyrus	47, 75, 60	2.13	1.82
6	Precentral Gyrus	66, 60, 60	6.38	4.29	47	Inferior Temporal Gyrus	63, 39, 29	1.13	0.00
7	Temporal Occipital Fusiform Cortex	59, 35, 31	3.52	1.82	48	Superior Parietal Lobule	61, 36, 58	1.74	1.28
8	Precuneus Cortex	48, 36, 49	3.32	1.39	49	Insular Cortex	64, 75, 37	1.37	0.84
9	Angular Gyrus	66, 30, 55	3.28	2.44	50	Frontal Pole	62, 80, 30	1.48	0.71
10	Frontal Pole	58, 75, 60	1.72	1.16	51	Precentral Gyrus	59, 61, 61	3.77	2.81
11	Middle Temporal Gyrus, posterior	76, 51, 31	1.81	0.00	52	Superior Frontal Gyrus	51, 78, 58	1.46	1.34
12	Frontal Medial Cortex	50, 91, 37	0.73	0.00	53	Inferior Frontal Gyrus	69, 72, 39	1.05	0.82
13	Supramarginal Gyrus, Posterior	71, 58, 49	2.19	2.05	54	Supramarginal Gyrus, Posterior	71, 38, 54	6.23	6.04
14	Frontal Medial Cortex	49, 86, 43	5.84	3.19	55	Precuneus Cortex	48, 44, 58	3.21	1.58
15	Cingulate Gyrus, Posterior	47, 41, 49	4.79	4.13	56	Middle Temporal Gyrus, posterior	75, 52, 30	1.33	0.63
16	Frontal Pole	61, 80, 29	0.94	0.00	57	Medial Precentral Gyrus	50, 49, 68	3.23	2.32
17	Supramarginal Gyrus, Posterior	72, 42, 47	1.93	0.90	58	Frontal Pole	63, 88, 37	2.81	2.52
18	Parahippocampal Gyrus	56, 52, 30	3.14	3.07	59	Superior Parietal Lobule	56, 39, 65	2.40	2.75
19	Middle Temporal Gyrus, Anterior	76, 58, 28	1.00	0.00	60	Superior Frontal Gyrus	56, 60, 68	2.46	1.78
20	Occipital Pole	50, 22, 36	3.19	2.33	61	Supramarginal Gyrus, Anterior	72, 46, 49	6.70	6.27
21	Superior Temporal Gyrus	70, 58, 32	3.30	1.52	62	Medial Postcentral Gyrus	58, 52, 62	4.83	3.32
22	Parahippocampal Gyrus, Posterior	53, 38, 38	7.70	5.38	63	Precuneus Cortex	48, 35, 58	8.22	6.90
23	Lateral Occipital Cortex, Superior	64, 25, 50	2.26	2.27	64	Frontal Pole	64, 81, 49	5.36	4.07
24	Precuneus Cortex	49, 38, 62	3.33	1.31	65	Supramarginal Gyrus, Anterior	74, 45, 57	2.36	0.89
25	Middle Frontal Gyrus	56, 70, 60	1.88	0.00	66	Lateral Occipital Cortex, Inferior	51, 30, 65	2.91	2.79
26	Superior Frontal Gyrus	54, 81, 57	5.23	3.64	67	Superior Frontal Gyrus	57, 68, 64	5.75	4.32
27	Middle Frontal Gyrus	71, 63, 21	1.12	0.89	68	Paracingulate Gyrus	47, 74, 57	2.48	2.32
28	Occipital Pole	53, 18, 43	1.51	0.69	69	Lateral Occipital Cortex, Inferior	69, 29, 40	1.47	0.00
29	Cingulate Gyrus, Anterior	47, 74, 50	6.38	0.00	70	Occipital Fusiform Gyrus	67, 31, 31	3.02	2.11
30	Frontal Operculum Cortex	63, 68, 36	3.64	2.80	71	Temporal Pole	60, 63, 24	1.48	0.53
31	Occipital Pole	59, 18, 33	1.24	0.55	72	Frontal Pole Lateral	67, 77, 39	2.11	1.38
32	Cuneal Cortex	50, 24, 50	4.01	3.17	73	Middle Frontal Gyrus	58, 83, 51	5.27	4.73
33	Supramarginal Gyrus, Anterior	68, 44, 57	3.34	3.35	74	Middle Temporal Gyrus, Occipital	75, 47, 28	5.64	5.06
34	Inferior Frontal Gyrus	71, 66, 49	3.58	1.96	75	Frontal Pole	57, 90, 45	2.91	3.80
35	Cingulate Gyrus, Posterior	49, 38, 51	12.66	11.39	76	Middle Temporal Gyrus, Occipital	73, 38, 37	2.95	1.88
36	Angular Gyrus	62, 35, 58	4.40	3.51	77	Cingulate Gyrus, Posterior	47, 46, 55	3.41	2.15
37	Pre/Postcentral Gyri	63, 49, 65	1.36	0.00	78	Angular Gyrus	70, 32, 47	2.13	1.15
38	Supplementary Motor Cortex	48, 57, 61	1.11	0.76	79	Frontal Orbital Cortex	57, 83, 29	1.80	1.25
39	Superior Frontal Gyrus	48, 85, 54	8.15	6.79	80	Frontal Pole	52, 92, 33	2.88	1.07
40	Frontal Orbital Cortex	66, 76, 33	3.32	2.72	81	Frontal Pole	54, 95, 40	0.80	0.81
41	Middle Frontal Gyrus	65, 73, 58	1.35	1.43					

Supplementary Table 2.

Genome-wide associations (linear regression, two-sided) with the residuals of microstructure-function phenotype (i.e. the pattern of functional connectivity that cannot be explained by white matter microstructure). Listed is the rsid of the SNP showing the most significant association. Additionally, the brain area is reported (see Supplementary Table 1), the nearest gene of the SNP, the base-pair position, the SNP alleles, the minor allele frequency (maf) and the p-values of the discovery (n = 7481 subjects) and the replication GWAS (n = 3873 subjects) are given. A significance threshold is given for a $-\log_{10}(\text{p-value})$ equal to 7.5 corresponding to a p-value of $\sim 3 \times 10^{-8}$. Significance threshold for the replication GWAS was determined using Bonferroni correction ($p < 2.9 \times 10^{-3}$).

chr	area	rsid	nearest gene	position	ref allele	minor allele	maf	p-value	replication p-value
10	30	rs7080472	PLCE1	96012950	G	G	0.42	7.48E-10	5.98E-05

Supplementary Table 3.

Genome-wide associations (linear regression, two-sided) with the first principal component of the multi-modal microstructure phenotype of each white matter tract. Listed are rsids of the SNPs showing the most significant association that were replicated in the replication cohort. Additionally, the white matter tract is reported (connecting the areas as listed in Supplementary Table 1), the nearest gene of each SNP, the base-pair position, the SNP alleles, the minor allele frequency (maf) and the p-values of the discovery (n = 7481 subjects) and the replication GWAS (n = 3873 subjects) are given. A significance threshold is given for a $-\log_{10}(\text{p-value})$ equal to 7.5 corresponding to a p-value of $\sim 3 \times 10^{-8}$. Significance threshold for the replication GWAS was determined using Bonferroni correction ($p < 3.6 \times 10^{-4}$).

chr	tract	rsid	nearest gene	position	ref allele	minor allele	maf	p-value	replication p-value
3	3	rs13083859	GMNC	190662703	G	G	0.38	2.60E-08	8.17E-06
	4	rs10937450		190652328	G	G	0.38	3.69E-10	9.66E-08
	20	rs10937450		190652328	G	G	0.38	1.37E-09	6.40E-07
	25	rs13066753		190647187	C	C	0.39	9.87E-09	3.38E-07
	28	rs10937450		190652328	G	G	0.38	1.02E-08	2.14E-07
	31	rs10937450		190652328	G	G	0.38	5.08E-10	6.77E-08
	34	rs13066753		190647187	C	C	0.39	2.39E-08	2.03E-05
	60	rs13100545		190657895	A	A	0.38	1.11E-08	1.34E-05
	62	rs13066848		190668204	T	T	0.38	3.14E-08	7.22E-06
5	12	rs13180395	VCAN	82716119	T	T	0.30	1.67E-09	4.01E-08
	14	rs13180395		82716119	T	T	0.30	7.99E-09	1.27E-07
	75	rs4470745		82789647	A	A	0.41	8.86E-09	6.42E-06
	81	rs4470745		82789647	A	A	0.41	1.32E-09	1.40E-07

Supplementary Table 4.

List of confound variables for the microstructure-function models.

Cumulative count	Confound variable name
1	Age
2	Age squared
3	Age * sex
4	Age squared * sex
5	Age (<i>quantile normalised</i>)
6	Age squared (<i>quantile normalised</i>)
7	Age * sex (<i>quantile normalised</i>)
8	Age squared * sex (<i>quantile normalised</i>)
9	rfMRI head motion
10	tfMRI head motion
11	Head size scaling
12	rfMRI head motion squared
13	tfMRI head motion squared
17	4 confounds relating to bed position in scanner (x,y,z,table)
21	4 confounds relating to bed position in scanner (x,y,z,table)
31	10 confounds modelling slow date-related drift
32	rfMRI head motion (<i>quantile normalised</i>)
33	tfMRI head motion (<i>quantile normalised</i>)
34	Head size scaling (<i>quantile normalised</i>)
38	4 confounds relating to bed position in scanner (x,y,z,table) (<i>quantile normalised</i>)
42	4 confounds relating to bed position in scanner (x,y,z,table) squared (<i>quantile normalised</i>)
52	10 confounds modelling slow date-related drift (<i>quantile normalised</i>)
53	Imaging centre
54	Sex
55	White matter lesion volume
57	2 dMRI EDDY movement parameters
58	Inverse T1 SNR
59	Inverse REST SNR
60	White matter lesion volume (<i>quantile normalised</i>)
62	2 dMRI EDDY movement parameters (<i>quantile normalised</i>)
63	Inverse T1 SNR (<i>quantile normalised</i>)
64	Inverse REST SNR (<i>quantile normalised</i>)