



HAL
open science

The genetic architecture of the human language connectome

Yasmina Nozha Mekki

► **To cite this version:**

Yasmina Nozha Mekki. The genetic architecture of the human language connectome. Neuroscience. Université Paris-Saclay, 2022. English. NNT : 2022UPAST019 . tel-03649334

HAL Id: tel-03649334

<https://theses.hal.science/tel-03649334v1>

Submitted on 22 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The genetic architecture of the human language connectome

*L'architecture génétique du connectome du langage dans
le cerveau humain*

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 575, Electrical, Optical, Bio-Physics and Engineering
(EOBE)

Spécialité de doctorat : Sciences de l'information et de la communication
Graduate School : Sciences de l'ingénierie et des systèmes, Référent : Faculté des
sciences d'Orsay

Thèse préparée dans l'unité de recherche **BAOBAB (Université Paris-Saclay,
CEA, CNRS)**, sous la direction de **Vincent Frouin**, directeur de recherche,
le co-encadrement de **Cathy Philippe**, docteure et de **Vincent Guillemot**, docteur.

Thèse soutenue à Paris-Saclay, le 15 février 2022, par

Yasmina Nozha Mekki

Composition du jury

Christophe Pallier

Directeur de recherche, CEA NeuroSpin, Université Paris Saclay

Président

Hugues Aschard

Directeur de recherche, Institut Pasteur

Rapporteur & Examineur

Marie-Hélène Grosbras

Directrice de recherche, Université Aix Marseille

Rapporteuse & Examinatrice

Vincent Frouin

Directeur de recherche, CEA NeuroSpin, Université Paris Saclay

Directeur de thèse

Abstract

Language is an essential component of human life because of its central role in social interaction, cultural transmission, and for structuring our thoughts. Through developmental disorders of spoken and written language studies such as dyslexia, a growing number of genes have been identified and labelled as implicated in language etiology. Case-control studies suffering from a lack of statistical power, struggle to uncover significant variants. In parallel, neuroimaging studies have significantly contributed to the understanding of structural and functional aspects of language in the human brain. The recent availability of large-scale cohorts comprising both neuroimaging and genetics data such as the UK Biobank and the Human Connectome Project, have made it possible to study language in the general population via image-derived phenotypes (IDP). Indeed, these IDPs give access to biologically relevant measurements of individual variability and are consequently suitable for the search of genetic associations. In this thesis, we aim to first, highlight the inter-individual variation in brain structure and/or function of language; second, take advantage of this variability to study the genetic basis of the brain's infrastructure for language. For this, we used task-free functional connectivity (FC), extracted from perisylvian cortical regions as well as from anatomical connectivity (AC) measured using well-known language related white matter bundles determined from diffusion MRI. First, we present an in-depth study of language heritability using human language connectome as endophenotypes extracted from the two aforementioned large-scale cohorts. The results found are consistent with the heritability estimates reported in the literature and support that the language-related brain organization is partly underpinned by genetic. Secondly, using the significantly heritable FC endophenotypes, we performed a multivariate genome-wide association study on 32,186 participants from UK Biobank. Twenty genomic loci were found significantly associated with language FCs, out of which three were replicated in an independent sample (the UK Biobank non-British sample, N=4,754). The functional annotation highlighted noticeably the *EPHA3* gene associated with FCs of the fronto-parieto-temporal semantic network, and the *THBS1* gene with a potential role in the perceptual-motor interaction required for language processing. Finally, using the significantly heritable AC endophenotypes, we performed a multivariate genome-wide association study on 31,775 participants from UK Biobank. 278 genomic loci were found significantly associated with language ACs including those previously associated with FC endophenotypes. As a conclusion, this work contributed to uncovering the neurobiology

of human language using state-of-the-art genomic approaches and by developing novel language endophenotypes. The genes pinpointed in this work open new leads for investigations of neurobiology of language using animal and cellular models. Furthermore, as more and more cohorts are being built in the general population, this work strengthens our belief that multi-modal data collection in large-scale cohorts allow us to increase statistical power and to circumvent the small effect sizes in neuroimaging in order to investigate open neuroscience questions. Many other neuroscience fields could probably benefit from this type of methodology.

Résumé

Le langage est une composante essentielle de la vie humaine en raison de son rôle central dans les interactions sociales, la transmission culturelle, et la structuration de nos pensées. Grâce à l'étude des troubles du développement du langage parlé et écrit, tel que la dyslexie, un nombre croissant de gènes ont été identifiés et étiquetés comme étant impliqués dans leurs étiologies. Des études d'association cas-témoins sont mises en œuvre, mais souffrent d'un manque de puissance statistique et peinent à produire des résultats significatifs. Parallèlement à cela, des études de neuro-imagerie ont contribué de manière significative à la compréhension des aspects structurels et fonctionnels du langage dans le cerveau. La disponibilité récente des cohortes de grande taille comportant à la fois de la neuro-imagerie et de la génétique tels que "UK Biobank" et le "Human Connectome Project" a rendu possible l'étude du langage dans les populations générales via des phénotypes intermédiaires (endophénotypes). En effet, ceux-là donnent accès à des mesures biologiquement pertinentes de la variabilité individuelle des supports structurels et fonctionnels du langage et sont par conséquent adaptés à la recherche d'associations génétiques. Dans cette thèse, nous souhaitons d'une part, mettre en évidence la variation interindividuelle de la structure et/ou de la fonction cérébrale qui reflète des traits neurocognitifs spécifiques; d'autre part, tirer profit de cette variabilité pour mettre en évidence la base génétique de leur architecture cérébrale. Pour ce faire, nous avons étudié la neurobiologie du langage à l'aide de la connectivité fonctionnelle (CF) des régions corticales périsylvienne mesurée à l'état de repos, ainsi qu'à partir de la connectivité anatomique (CA) mesuré sur des faisceaux de matière blanche déterminé par l'IRM de diffusion et connus comme étant liés au langage. Tout d'abord, nous présentons une étude approfondie de l'héritabilité du langage en exploitant des endophénotypes du connectome du langage humain extraits des deux cohortes susmentionnées. Les résultats obtenus sont conformes aux estimations de l'héritabilité rapportées dans la littérature et confirment que l'organisation cérébrale liée au langage est en partie sous-tendue par la génétique. Deuxièmement, pour les endophénotypes de CF avec une héritabilité significative, nous avons réalisé une étude d'association multivariée pangénomique sur 32.186 participants de la cohorte UK Biobank. Vingt loci sont significativement associés aux CF du langage, dont trois sont répliqués dans un échantillon indépendant (ethnicité non britannique de UKB, N=4.754). L'annotation fonctionnelle approfondie pointe notamment vers le gène *EPHA3* avec un rôle dans le réseau sémantique fronto-pariéto-temporal, et le gène *THBS1* avec un rôle dans l'interaction perception-moteur,

requis dans le langage. Enfin, de la même manière, pour les endophénotypes de CA avec une héritabilité significative, nous avons effectué une étude d'association multivariée pangénomique sur 31.775 participants de la cohorte UK Biobank. 278 loci sont significativement associés aux CA du langage, dont ceux précédemment associés à des endophénotypes de CF. En conclusion, ce travail a permis d'étudier les bases génétiques du langage humain en utilisant des approches génomiques à l'état de l'art et ce, en développant des endophénotypes originaux associés au langage. Les gènes suggérés par l'analyse fournissent des pistes d'investigation dans des modèles animaux ou cellulaires du langage. Aussi, alors que de plus en plus de cohortes sont construites en population générale, ce travail renforce notre conviction que ce type de cohortes "multiphénotypées" permet d'augmenter la puissance statistique et de pallier les faibles tailles d'effet en neuro-imagerie afin d'étudier des questions ouvertes en neuroscience. De nombreux autres domaines des neurosciences pourraient probablement bénéficier de ce type de méthodologie.

Contents

List of Figures	xi
List of Tables	xxi
Abbreviations and Acronyms	xxv
1 State of the art	1
1.1 Genetic basis	3
1.1.1 The genome and inter-individual variability	4
1.1.2 From genotypes to phenotype	6
1.2 Neuroanatomy basics	6
1.2.1 Grey and white matter	7
1.2.2 Gyri and Sulci	7
1.2.3 Cells of the brain: From molecule to neurons and from neurons to brain	8
1.3 How to determine the brain architecture? A neuroimaging approach	9
1.3.1 Structural Neuroimaging	10
1.3.2 Functional Neuroimaging	13
1.4 A basic introduction to quantitative and molecular genetics methodology	17
1.4.1 Heritability	17
1.4.2 Endophenotypes	18
1.4.3 Pleiotropy	18
1.4.4 Candidate genes	18
1.4.5 Genome-wide association study	19
1.5 What are the bases of human language? A review of language approaches in human neurosciences	21
1.5.1 Understanding language from aphasia	21
1.5.2 Understanding language by brain imaging	22
1.6 Complex and multifactorial aetiology: the pursuit of brain–language–genetics relationships	29
1.6.1 Heritability of language	30

1.6.2	Developmental language disorder related loci	32
1.6.3	Genome-wide association studies of language abilities	33
1.6.4	Moving Forward: Emergence of imaging-genetics approach	35
1.7	Aim of this thesis	39
2	Materials and methods	41
2.1	Cohorts	42
2.1.1	UK Biobank	42
2.1.2	Human Connectome Project	45
2.2	Definition of brain regions and tracts of interest	48
2.2.1	Functional connectomes	48
2.2.2	Anatomical connectomes	52
2.3	Rank-based inverse-normal transformation	55
2.4	Heritability	55
2.5	Genetic correlation	56
2.6	Massive univariate and multivariate genome-wide association study	56
2.7	Fine-mapping: identification of genomic risk loci and functional annotations	59
2.7.1	FUMA: SNP ₂ GENE	59
2.7.2	FUMA: GENE ₂ FUNC	60
2.8	Protein networks	60
3	Human language endophenotypes	63
3.1	Psychometric measures	64
3.2	Language neural activation: a volume-based analysis	67
3.3	Resting-state functional connectome	69
3.4	Anatomical connectome	70
3.5	Conclusion	72
4	Heritability of human language	73
4.1	Pedigree-based heritability of behavioural and neuroimaging (endo)phenotypes: a twin study	75
4.1.1	Cognitive behavioural assessment	75
4.1.2	Brain neural activation	76
4.1.3	Functional connectome	76
4.2	SNP-based heritability of neuroimaging endophenotypes: a general population study	77
4.2.1	Functional connectome	77
4.2.2	Anatomical connectome	80
4.3	Discussion	80
4.4	Are heritability estimates inflated in twin studies?	80
4.5	Conclusion	84

5	Shared genetic variance between language-related performance, language-brain activations and language functional connectivity	85
5.1	Shared heritability between both resting-state functional connectivity and neural activations	87
5.2	Shared heritability between both resting-state functional connectivity and language-related scores	88
5.3	Shared heritability between both task activations and language-related scores	88
5.4	Discussion	89
6	The genetic architecture of human language functional connectome	91
6.1	Multivariate genome-wide association analysis	93
6.2	Unraveling the genome: Evaluation of results from mvGWAS study of human language	103
6.2.1	<i>Perceptual motor</i> interactions process driven by <i>THBS1</i> gene	103
6.2.2	<i>Fronto-temporal semantic network</i> driven by <i>EPHA3</i> gene	110
6.2.3	PrecR↔RolS and T1a↔T1a/HeschlR endophenotypes driven by <i>PLCE1</i> , <i>NOC3L</i> and <i>HELLS</i> genes	115
6.3	Limitations	117
6.4	Conclusions	118
7	The genetic architecture of anatomical connectivity in human language	119
7.1	Multivariate genome-wide association studies	120
7.2	Validation of lead SNPs associated with ACs.	124
7.3	Functional annotations of genomic loci associated with language	129
7.4	Discussion	135
	General Conclusions	137

Appendices	141
A	143
A.1 Population structure of Human connectome project data	143
B	145
B.1 Pedigree-based heritability of behavioural and neuroimaging (endo)phenotypes: a twin study	145
B.2 SNP-based heritability of neuroimaging endophenotypes: a general population study	153
C	169
C.1 (Endo)phenotypic correlation between both resting-state functional connectivity and neural activations	169
C.2 (Endo)phenotypic correlation between both resting-state functional connectivity and language-related scores	172
C.3 (Endo)phenotypic correlation between both task activations and language-related scores	177
D	180
D.1 Genomic loci associated with heritable language anatomical connectivities using MOSTest	180
E Résumé en français (Abstract in French)	209
Bibliography	215

* * *
* *
*

List of Figures

1.1	Description of the chromosome structure. (Source: https://www.genome.gov/).	4
1.2	Neuroanatomy of the brain ([Bui and Nguyen, 2019]). A) Brain, Encephalon, Connections of the several parts of the brain, Cerebrum, Cerebellum, Pons, Cerebral; Superior; Middle; Inferior Peduncle, Medulla oblongata. B) Principal fissures and lobes of the cerebrum viewed laterally, Frontal Lobe, Parietal Lobe, Temporal Lobe, Occipital Lobe.	7
1.3	Brain basics. Example of a) the grey and white matter of the central nervous system (source : https://difference.guru/difference-between-white-and-gray-matter/), b) sulcus and gyrus.	8
1.4	Schematic representation of the major cellular elements of the white matter (Source: [Johansen-Berg and Behrens, 2013]). In this drawing, a neuronal cell body in the grey matter extends several short dendritic processes and a single, long axonal process. The axon, in the white matter, is surrounded by an insulating myelin sheath. This sheath is produced by oligodendrocytes. Generally, oligodendrocytes extend many processes, each of which wraps around a segment of an axon to form an internode. Internodes are separated from each other by small amyelinated regions called nodes of Ranvier. This specialized organisation of the myelin sheaths promotes saltatory conduction whereby electrical impulses “jump” from node to node as they pass along the axon from the neuronal cell body to the axon terminus. The white matter also contains astrocytes and microglia. Astrocytes are star-shaped glial cells that help maintain homeostasis by regulating local ion concentrations and extracellular fluid volume within the CNS. In the whiter matter, they extend processes to nodes of Ranvier and to blood vessels, where they contribute to blood–brain barrier function. Microglial cells, which are considered the immune cells of the CNS, help “mop up” debris and dying cells.	9

1.5 **Isotropic and anisotropic diffusion in the brain (Source: [Pandit et al., 2013]).** Fractional anisotropy image of a preterm born child reveals the nature of water diffusion within different brain tissue compartments (a). In the white matter of the corpus callosum (red), diffusion occurs preferentially along the axonal fibres, resulting in anisotropic diffusion (b). In the ventricular cerebro-spinal fluid (CSF; green), diffusion is unhindered and can be described as isotropic (c). Diffusion tensor ellipsoids representing anisotropic and isotropic diffusion are shown in b and c, respectively. Each tensor is expressed by three eigenvectors with values λ_1 , λ_2 and λ_3 . In isotropic diffusion $\lambda_1 = \lambda_2 = \lambda_3$, whereas in anisotropic diffusion the long axis of the ellipsoid aligns with the underlying white matter, and λ_1 is greater than λ_2 and λ_3 11

1.6 **Principles of diffusion tensor imaging (DTI) tractography (Source: [Catani and Forkel, 2019]).** (A) Visualization of the diffusion tensor as an ellipsoid. The size and the shape of the tensor are defined by the three eigenvalues (1, 2, 3 in red), while the spatial orientation is described by the three eigenvectors (v_1 , v_2 , v_3 in blue). In biological tissues, the tensor can vary between three possible configurations: (1) isotropy or weak anisotropy (equal or similar diffusivity along the three eigenvalues) is commonly observed, for example, inside the ventricles (isotropy) or in the gray matter (weak anisotropy); (2) planar anisotropy (unequal diffusivity between of one eigenvalue and the other two) is common in voxels containing, for example, two groups of crossing or diverging fibers; and (3) axial anisotropy is typical of voxels containing parallel fibers (unequal diffusivity between the axial eigenvalue and the two perpendicular eigenvalues). (B) Exemplified indices extracted from diffusion data, such as mean diffusivity (MD), fractional anisotropy (FA), principle eigenvector, and color-coded FA maps. (C) Streamline tractography is based on the assumption that in each white matter voxel the principal eigenvector (black arrows) is tangent to the main trajectory of the underlying fibers (black lines). Starting from a region of interest (red circle), the tractography algorithm propagates, voxel by voxel, a streamline (red) by piecing together neighboring principal eigenvectors. An example is shown in the neighboring panel where the streamlines are tracked on a principal eigenvector map and the tractography reconstruction of the arcuate fasciculus is visualized as 3D streamtubes. 12

1.7 Figure extracted from [Glover, 2011]. Block design fMRI experiment. A neural response to the state change from A to B in the stimulus is accompanied by a hemodynamic response that is detected by the rapid and continuous acquisition of MR images sensitised to BOLD signal changes. Using single- or multi-variate time series analysis methods, the average signal difference between the two states is computed for the scan and a contrast map generated. A statistical activation map is finally obtained using a suitable threshold for the difference; the map depicts the probability that a voxel is activated given the uncertainty due to noise and the small BOLD signal differences. 14

- 1.8 **Figure extracted from [Lee et al., 2013]. Surface plots of resting-state networks.** A: Default mode network, B: Somatomotor network, C: Visual network, D: Language network, E: Dorsal attention network, F: ventral attention network, G: Fronto-parietal control network. 15
- 1.9 **Principles of GWA illustrated with an example of a continuous endophenotype: a functional measure of the brain.** On the top figure, the y-axis is the endophenotype measure, and the x-axis represents the targeted genetic variant. This variant has two possible alleles: C and T. Each individual in the cohort will have a combination of CC, CT, TT alleles coded as 0, 1 or 2. The lower figure represent a Manhattan plot and is a figure-of-merit of GWASes. The y-axis is the $-\log_{10}$ of the p-value of the association, and the x-axis represents the SNP location along the genome. The most significant associations will have the smallest p-values and the $-\log_{10}$ of these p-values will have the highest height in the graph. The red dashed line corresponds to the $5e - 8$ genomic threshold. The points above correspond to the genetic variants significantly associated with the studied phenotype. 20
- 1.10 **Neuroanatomy of language (Source: [Friederici, 2017]).** Anatomical and cytoarchitectonic details of the left hemisphere (LH). Top: The different lobes (frontal, temporal, parietal, occipital) are marked by colored borders. Major language-relevant gyri (inferior frontal gyrus (IFG), superior temporal gyrus (STG), middle temporal gyrus (MTG)) are color-coded, superior temporal sulcus (STS) located between STG and MTG is marked by an asterisk. Numbers indicate language-relevant Brodmann areas (BA) that Brodmann (1909) defined on the basis of cytoarchitectonic characteristics. The coordinate labels (see bottom left) superior (dorsal)/inferior (ventral) indicate the position of the gyrus within a lobe (e.g., superior temporal gyrus) or within a BA (e.g., superior BA 44). The horizontal coordinate labels anterior (rostral)/posterior (caudal) indicate the position within a gyrus (e.g., anterior superior temporal gyrus). Broca's area consists of the pars opercularis (BA 44) and the pars triangularis (BA 45). Located anterior to Broca's area is the pars orbitalis (BA 47). The frontal operculum (FOP) is located ventrally and medially to BA 44, BA 45. The premotor cortex (PMC) is located in BA 6. Wernicke's area is defined as BA 42 and BA 22. The primary auditory cortex (PAC) and Heschl's gyrus (HG) are located in a lateral to medial orientation in the temporal lobe. White matter fiber tracts, i.e., the dorsal and ventral pathways connecting the language-relevant brain regions are indicated by color-coded arrows (see bottom right). Reprinted from Friederici [2011]. The brain basis of language processing: From structure to function. *Physiological Reviews*, 91 (4): 1357–1392 . 23

- 1.11 **Overview of the meta-analysis (Source: [Vigneau et al., 2006, 2011]).** Each foci is color-coded according to its language components category: phonology (blue), semantic (red), and syntax (green). A) Distribution of the peaks in each hemisphere for each component. The 128 neuroimaging studies selected for our previous meta-analysis reported 728 activation peaks in the left hemisphere (left column). Among the 128, 59 studies reported 218 activation peaks in the right hemisphere (right column). Activation peaks detected during phonological tasks are indicated in blue (top row), lexico-semantic tasks in red (middle row) and sentence and text processing in green (bottom row). For this last component, peaks corresponding to contrast investigating comprehension are in green and those investigating syntax are in white. Peaks are projected onto a lateral view of the MNI single subject. B) Top: sagittal projection map of the 730 activation peaks. Bottom: clusters identified for phonological (blue), semantic (red), and syntactic (green) processing and obtained from the spatial clustering of the peaks. Clusters of different components that were not spatially distinct are circled (yellow). Error bars correspond to twice the standard error on the y and z stereotactic coordinates. The left hemisphere lateral surface rendering of the MNI single-subject brain template, together with the corresponding stereotactic grid, is displayed for anatomical reference. RolS, Rolandic sulcus; RolOp, Rolandic operculum; F3t, pars triangularis of the left inferior frontal gyrus; F3op, pars opercularis of the left inferior frontal gyrus; F3orb, pars orbitaris of the left inferior frontal gyrus; SMG, supramarginalis gyrus; PT, planum temporale; T1, superior temporal gyrus; T2, middle temporal gyrus; T3, inferior temporal gyrus; Prec, precentral gyrus; F2, middle frontal gyrus; PrF3op, precentral gyrus/F3op junction; STS, superior temporal sulcus; AG, angular gyrus; Fusa: anterior fusiform gyrus; a, anterior; p, posterior; l, lateral; m, middle; d, dorsal; v, ventral. 25
- 1.12 **Tractography-based reconstruction of the language networks from classical models to contemporary neurolinguistics (Source: [Catani and Forkel, 2019]).** (A) Classical language model with the arcuate fasciculus connecting Broca's region in the inferior frontal gyrus to Wernicke's region in the superior temporal gyrus. (B) Extension of the classical arcuate fasciculus *sense strictu* to include the anterior segment, connecting inferior frontal to inferior parietal lobe, and the posterior segment linking the inferior parietal to the temporal lobes. (C) Current model of an extended language network beyond the three segments of the arcuate fasciculus. The frontal aslant tract (FAT) connects the inferior frontal gyrus to the pre-supplementary motor cortex. The ventral network includes the uncinate fasciculus between the anterior temporal lobe and the orbital frontal and inferior frontal cortex, the temporal longitudinal fasciculus (TLF) between the posterior temporal lobe and the temporal pole, the inferior fronto-occipital fasciculus (IFOF) connecting the ventral frontal cortex to the occipital cortex, and the inferior longitudinal fasciculus (ILF) between the occipital and anterior temporal cortex. 27

- 1.13 **Parent-offspring correlations for factor scores for specific cognitive abilities for adoptive, biological, and control parents and their children at 3, 4, 7, 9, 10, 12, 14, and 16 years (source: [Stromswold, 2001]).** Correlations are shown for verbal ability (a), spatial ability (b), speed of processing(c), and recognition memory (d). Parent-offspring correlations are weighted averages for mothers and fathers. The number of individuals range from 33 to 44 for biological fathers, 159 to 180 for biological mothers, 153 to 197 for adoptive parents, and 136 to 217 for control parents. 31
- 2.1 **Overview of the regions obtained from [Pallier et al., 2011].** Each language seed is color-coded according to its language category: phonology (blue), semantic (red), and syntax (green). ROIs of different components that were not spatially distinct are color-coded as pink (semantic/syntax), cyan (phonology/semantic) and white for the three language component. Different gyri and sulcus, known to be relevant for language: the inferior frontal gyrus (IFG), middle temporal gyrus (MTG), superior temporal gyrus (STG), and superior temporal sulcus (STS), are color-coded. Numbers in the left hemisphere (LH) represents language-relevant Brodmann areas (BA) which were defined on the basis of cytoarchitectonic characteristics. The pars opercularis (BA 44), the pars triangularis (BA 45) represents Broca’s area. The pars orbitalis (BA 47) is located anterior to Broca’s area. BA 42 and BA 22 represents Wernicke’s area Friederici [2011]. Both supramarginal gyrus (BA40) and angular gyrus (BA39), also known as Geschwind’s territory, are represented by green and yellow colors respectively. The primary motor cortex (BA4), the premotor cortex and the supplementary motor area (B6) are colored in orange. Within the left hemisphere, dorsal and ventral long-range fiber bundles connect language areas and are indicated by color-coded arrows. 49
- 2.2 **Overview of the regions obtained from the meta-analysis.** Each language seed is colour-coded according to its language category: phonology (blue), semantic (red), and syntax (green). ROIs of different components that were not spatially distinct are colour-coded as pink (semantic/syntax), cyan (phonology/semantic) and white for the three language component. For the sake of ROIs figure visibility, the coordinates were modified. The exact coordinates for each ROI are available in Table 2.2. Different gyri and sulcus, known to be relevant for language: the inferior frontal gyrus (IFG), middle temporal gyrus (MTG), superior temporal gyrus (STG), and superior temporal sulcus (STS), are color-coded. Numbers in the left hemisphere (LH) represents language-relevant Brodmann areas (BA) which were defined on the basis of cytoarchitectonic characteristics. Numbers in the right hemisphere (RH) represents the language-relevant BA counterpart. The pars opercularis (BA 44), the pars triangularis (BA 45) represents Broca’s area. The pars orbitalis (BA 47) is located anterior to Broca’s area. BA 42 and BA 22 represents Wernicke’s area [Friederici, 2011]. Both supramarginal gyrus (BA40) and angular gyrus (BA39), also known as Geschwind’s territory, are represented by green and yellow colours respectively. The primary motor cortex (BA4), the premotor cortex and the supplementary motor area (B6) are coloured in orange. Within the left hemisphere, dorsal and ventral long-range fibre bundles connect language areas and are indicated by colour-coded arrows. 50

2.3	<p>Spherical deconvolution tractography of frontal white matter connections. (Source: [Rojkova et al., 2016]) For each tract a single participant map is supplied as a representative example of the individual anatomy. First (SLF I, light blue, no. 1), second (SLF II, dark blue, no. 2) and third (SLF III, pink, no. 3) branches of the superior longitudinal fasciculus; cingulum (yellow, no. 4); uncinate (pink, no. 5); long (LS, red, no. 6) and anterior (AS, green, no. 7) segments of the arcuate fasciculus; inferior fronto-occipital fasciculus (IFOF, no. 8); frontal corpus callosum (CC, no. 9); corticospinal tract (CST, no. 10); fronto-thalamic projections or anterior thalamic radiations (blue, no. 11); fronto-striatal projections (yellow, no. 12); fronto-pontine projections (no. 13); paracentral U tract (pink, no. 14); hand superior (green, no. 15), middle (yellow, no. 16) and inferior (red, no. 17) U tract; face U tract (blue, no. 18); frontal aslant tract (pink, no. 19); fronto-insular tract 1 (FIT 1, light blue, no. 20), 2 (FIT 2, dark blue, no. 21), 3 (FIT 3, red, no. 22), 4 (FIT 4, yellow, no. 23) and 5 (FIT 5, green, no. 24); frontal superior longitudinal (FSL, light blue, no. 25); frontal inferior longitudinal (FIL, pink, no. 26), frontal orbito-polar tract (FOP, yellow, no. 27) and fronto-marginal tract (FMT, red, no. 28). The overall visualization and screenshots were performed in Anatomist (http://brainvisa.info).</p>	54
3.1	<p>Distribution of the NIH measures collected by HCP team independently from the language fMRI task</p>	65
3.2	<p>Distribution of behavioral scores collected by HCP team during the language fMRI task, composed of MATH and STORY tasks</p>	66
3.3	<p>First level model. A) Design matrix used in the first level model. Predictors includes fMRI events and derivatives, twelve movement regressors and four drift features. B) Coefficients of the contrast, indexed by the names of the columns of the design matrix.</p>	68
3.4	<p>Group analysis language task activation maps. Group average activations for the HCP language task: story versus baseline is shown with a lower threshold of $z = \pm 10$.</p>	68
3.5	<p>Functional connectome extraction pipeline with two main steps: 1) Map of brain regions already defined on resting-state fMRI images. 2) Extraction of mean BOLD signal for each ROIs and quantifying functional interactions from these time series signal extracted. At the end of the second step, we obtain a functional connectome for each participants where each cell represents a resting-state functional connectivity.</p>	70
3.6	<p>Functional connectome visualisation Mean functional connectivity of the 300 and 15 endophenotypes, calculated using a shrunked estimate of partial correlation Marrelec et al. [2006] (estimated with a Ledoit-Wolf estimator Ledoit and Wolf [2004]) over 32,186 and 778 individuals using both UKB and HCP rsfMRI participants respectively. (a) Circos plot represents the mean FC estimated using SmPSS ROIs on UK Biobank. (b) Circos plot represents the mean FC estimated using SmPSS ROIs on HCP. (c) Circos plot represents the mean FC estimated using OcSS ROIs on UK Biobank.</p>	70

3.7	Anatomical connectome extraction pipeline with one main steps consisting on mapping brain tracts of interest already defined on skeletonised images and extracting mean value for each tract of interest. At the end, we obtain an anatomical connectome for each participants where each cell represents an anatomical connectivity.	72
4.1	Pedigree-based heritabilities of brain functional connectivities extracted using SmPSS ROIs. Circle plot where all 25 regions are represented (19 ROIs on the left represent the left hemisphere and 6 ROIs on the right represent the right hemisphere) illustrating the FCs with significant pedigree-based heritability after false discovery rate correction at $p < 0.05$. Each link between two regions represents an endophenotype. The intensity of the color between two regions represents the pedigree-based heritability of an endophenotype as defined by the color bar. Please refer to Table B.1 for all pedigree-based heritability values.	77
4.2	SNP-based heritabilities of brain functional connectivities extracted using OcSS ROIs. Circle plot where all 6 regions are represented (6 ROIs on the left represent the left hemisphere. No region in the right hemisphere are included in OcSS ROIs.) illustrating the FCs with significant SNP-based heritability after false discovery rate correction at $p < 0.05$. Each link between two regions represents an endophenotype. The intensity of the color between two regions represents the SNP-based heritability of an endophenotype as defined by the color bar. Please refer to Table B.2 for all SNP-based heritability values.	78
4.3	SNP-based heritabilities of brain functional connectivities extracted using SmPSS ROIs. Circle plot where all 25 regions are represented (19 ROIs on the left represent the left hemisphere and 6 ROIs on the right represent the right hemisphere) illustrating the FCs with significant SNP-based heritability after false discovery rate correction at $p < 0.05$. Each link between two regions represents an endophenotype. The intensity of the color between two regions represents the SNP-based heritability of an endophenotype as defined by the color bar. Please refer to Table B.3 for all SNP-based heritability values.	79
4.4	Comparison of heritabilities estimated according to different design parameters. The endophenotypes extracted using SmPSS ROIs. Boxplots 1 and 2 display the heritability distributions estimated on UKB (using GCTA) and HCP (using SOLAR) respectively. Boxplot 3 : HCP-FCs using GCTA. Boxplot 4 : HCP-FCs using GCTA and considering the household effect. Boxplots 5 to 7 : HCP (using SOLAR) on rsfMRI sessions with the same duration as in UKB (6 min).	81
4.5	Comparison between heritabilities obtained with SOLAR on HCP (concatenation of the 4 runs) and the ones obtained with : (a) GCTA on HCP, (b) SOLAR on HCP taking into account the shared environment, (c) GCTA on UKB, (d) SOLAR on HCP (6 first minutes of the signal), (e) SOLAR on HCP (6 minutes in the middle of the signal), (f) SOLAR on HCP (6 last minutes of the signal)	82
4.6	Different cut of the resting-state time serie (6 minutes) obtained from HCP cohort. 83	83
6.1	Multivariate GWAS analysis of the SmPSS resting state functional connectivity in 32,186 participants. Manhattan plot for multivariate GWAS across 142 FCs. The red dashed line indicates the genome-wide significance threshold $p = 5e-8$. The Quantile-quantile plot is also shown.	93

- 6.2 Locus Zoom of the significant loci identified by the multiariate GWAS for functional connectivity. 96
- 6.3 **Genomic loci, eQTL associations and chromatin interactions identified via multivariate GWAS for functional connectivity.** Circos plot representing the genomic risk loci, and the genes associated with the loci by chromatin interactions and eQTLs. From outer layer to inner layer: Manhattan plot. Genomic risk loci are in blue. Genes mapped by chromatin interaction are in orange. Genes mapped by eQTL are in green. Genes mapped by both are in red. Chromatin interaction and eQTLs links follows the same color coding presented above. 97
- 6.4 Multivariate GWAS analysis of the OcSS resting state functional connectivity in 32,186 participants. Manhattan plot for multivariate GWAS across 14 FCs. The red dashed line indicates the genome-wide significance threshold $p = 5e-8$. The Quantile-quantile plot is also shown. 98
- 6.5 **Regional effects.** Circle plot where all 25 regions are represented (19 ROIs on the left represent the left hemisphere and 6 ROIs on the right represent the right hemisphere) illustrating the lead SNPs identified from the multivariate GWAS for functional connectivity. Each link between two regions represents an endophenotype. The intensity of the color between two regions represents the Z-values from the univariate GWAS for each FCs as defined by the color bar. The absolute Z-values scaling is clipped at 8 ($p = 1.2e-15$). Positive effects of carrying the minor allele are shown in red, and negative in blue. 102
- 6.6 **Main results for the 15q14 locus.** A) The two pairs of ROIs that forms the endpoints of the associated FCs reported as black bold lines. B) Effect sizes of the SNP rs1440802 for the two connections: (Prec↔F3opd) FC in green and (PrecR↔Rols) FC in yellow. C) Locus Zoom of the genomic region identified by the mvGWAS. Chromatin state of the genomic region. Brain tissue name abbreviations are the following; E054:Ganglion Eminence derived primary cultured neurospheres, E053: Cortex derived primary cultured neurospheres, E071: Brain Hippocampus Middle, E074: Brain Substantia Nigra, E068: Brain Anterior Caudate, E069: Brain Cingulate Gyrus, E072: Brain Inferior Temporal Lobe, E067:Brain Angular Gyrus, E073: Brain Dorsolateral Prefrontal Cortex, E070: Brain Germinal Matrix, E082: Fetal Brain Female, E081: Fetal Brain Male, E125: NH-A Astrocytes Primary Cells. The state abbreviations are the following; TssA: active transcription start site (TSS), TssFlnk: Flanking Active TSS, TxFlnk: Transcription at gene 5' and 3', Tx: Strong transcription, TxWk: Weak transcription, EnhG: Genic enhancers, Enh: Enhancers, ZNF/Rpts: ZNF genes & repeats, Het: Heterochromatin, TssBiv: Bivalent/Poised TSS, BivFlnk: Flanking Bivalent TSS/Enh, EnhBiv: Bivalent Enhancer, ReprPC: Repressed PolyComb, ReprPCWk: Weak Repressed PolyComb, Quies: Quiescent/Low. Expression quantitative trait loci (eQTL) associations (data source: eQTLGen [Võsa et al., 2018], PsychENCODE[Wang et al., 2018], DICE [Schmiedel et al., 2018], BIOS QTL browser [Zhernakova et al., 2017], GTEx/v8 [Consortium et al., 2017], eQTLcatalogue). D) Overlap of the genomic region risk region identified from FUMA for MOSTest results, (Prec↔F3opd) and (PrecR↔Rols). E) Gene expression from BrainSpan for the interesting genes prioritised by FUMA. 110

- 6.7 **Main results for the 3p11.1 locus.** A) The pairs of ROIs that forms the endpoints of the associated FCs reported as black bold lines. B) Effect sizes of the SNP rs35124509 for the nine connections: (AG↔F3orb), (Pole↔STSp), (Pole↔T2ml), (T1a↔STSp), (STSp↔F3orb), (SMG↔T3p), (AG↔STSp), (F2p↔AG) and (T2ml↔SMG) FCs. C) Locus Zoom of the genomic region identified by the mvGWAS. Chromatin state of the genomic region. Brain tissue name abbreviations are the following; Eo54:Ganglion Eminence derived primary cultured neurospheres, Eo53: Cortex derived primary cultured neurospheres, Eo71: Brain Hippocampus Middle, Eo74: Brain Substantia Nigra, Eo68: Brain Anterior Caudate, Eo69: Brain Cingulate Gyrus, Eo72: Brain Inferior Temporal Lobe, Eo67:Brain Angular Gyrus, Eo73: Brain Dorsolateral Prefrontal Cortex, Eo70: Brain Germinal Matrix, Eo82: Fetal Brain Female, Eo81: Fetal Brain Male, E125: NH-A Astrocytes Primary Cells. The state abbreviations are the following; TssA: active transcription start site (TSS), TssFlnk: Flanking Active TSS, TxFlnk: Transcription at gene 5' and 3', Tx: Strong transcription, TxWk: Weak transcription, EnhG: Genic enhancers, Enh: Enhancers, ZNF/Rpts: ZNF genes & repeats, Het: Heterochromatin, TssBiv: Bivalent/Poised TSS, BivFlnk: Flanking Bivalent TSS/Enh, EnhBiv: Bivalent Enhancer, ReprPC: Repressed PolyComb, ReprPCWk: Weak Repressed PolyComb, Quies: Quiescent/Low. Expression quantitative trait loci (eQTL) associations (data source: eQTLGen [Võsa et al., 2018], PsychENCODE[Wang et al., 2018], DICE [Schmiedel et al., 2018], BIOS QTL browser [Zhernakova et al., 2017], GTEx/v8 [Consortium et al., 2017], eQTLcatalogue). D) Overlap of the genomic region risk region identified from FUMA for MOSTest results and the nine FCs mentioned above. E) Gene expression from BrainSpan for the interesting genes prioritised by FUMA. 115
- 7.1 **Multivariate GWAS analysis of the anatomical connectivity in 31,753 participants.** Manhattan plot for multivariate GWAS accross 313 FCs. The red dashed line indicates the genome-wide significance threshold $p = 5e-8$. The Quantile-quantile plot is also shown. 121
- 7.2 **Genomic loci, eQTL associations and chromatin interactions identified via multivariate GWAS for anatomical connectivity.** Circos plot representing the genomic risk loci, and the genes associated with the loci by chromatin interactions and eQTLs. From outer layer to inner layer: Manhattan plot. Genomic risk loci are in blue. Genes mapped by chromatin interaction are in orange. Genes mapped by eQTL are in green. Genes mapped by both are in red. Chromatin interaction and eQTLs links follows the same color coding presented above. Due to a high number of genes identified using eQTLs mapping (361) and chromatin interaction (874), the chromosome 17 circos plot represent only chromatin interaction.123
- 7.3 **Genome-wide gene-based association analysis of the anatomical connectivity in 31,753 participants.** genome-wide gene-based association analysis of 313 ACs using mvGWAS results as input. Input SNPs were mapped to 19,157 protein coding genes. The red dashed line indicates the Bonferroni-corrected significance threshold for gene-based analysis $p = 2.61e-6(0.05/19,157)$. The Quantile-quantile plot is also shown. 130

7.4	Human language AC-associated genes integrated into a protein–protein interaction network. Proteins are represented by nodes. Edges between nodes represent different types of protein–protein interactions according to the STRING [Szklarczyk et al., 2017] database (chapter 2.8), including known interactions (turquoise and dark purple represent interactions identified by curated databases and biological experiments, respectively) and interactions determined by co-expression. Coloured nodes represent the queried proteins. Only at least medium-confidence (> 0.4) links were retained, and disconnected proteins are not shown.	131
7.5	MAGMA gene-property analysis. A) Relation between gene-based association with human language AC and relatively higher mRNA expression in the human brain at particular ages, using BrainSpan data from 29 age groups. Asterisks indicate significant age groups meeting $p < 0.05$ with FDR correction. pcw: post-conceptual weeks. B) Relation between gene-based association with human language AC and relatively higher mRNA expression in the human brain at particular ages, using BrainSpan data from 11 developmental stages. Asterisks indicate significant groups meeting $p < 0.05$ with FDR correction.	134
A.1	Overview of the first five principal components of the principal component analysis of the genotype data. We can observe that the fifth principal component separate individual from the white ethnicity.	143

List of Tables

1.1	Summary of the linkage regions to reading disability (RD) and Specific language impairment (SLI) and the candidate genes within them (Source: Hagoort [2019]).	32
1.2	Summary of the top genome-wide association signals from studies of reading and language traits (Source: Hagoort [2019]).	34
1.3	Summary of imaging genetics findings related to reading/language. (Source: Landi and Perdue [2019].)	36
1.4	Summary of the genome-wide association signals from studies of left-right asymmetry (source: Sha et al. [2021]).	38
2.1	Overview of the regions obtained from the meta-analysis. Each ROI is characterised by its abbreviated anatomical label defined by Vigneau et al. [2006, 2011] and is labelled according to the language component it belongs to: phonology, semantic, and syntax.	51
2.2	Centre of mass of the language processing regions of interests retained in both left and right hemispheres. Each ROI is characterised by its abbreviated anatomical label defined by [Vigneau et al., 2006, 2011] and its centre of mass MNI stereotactic coordinates (x, y, z, in mm).	52
4.1	Heritabilities of language related behavioral scores. Endophenotypes in bold represent language related behavioral scores with significant heritability after false discovery rate correction at $p < 0.05$	75
4.2	Heritabilities of language task activation. Bold endophenotypes represents language task activation with significant heritability after false discovery rate correction at $p < 0.05$	76

5.1	Shared heritabilities between both resting-state functional connectivity and neural activations.	Bold rows represents pairs of endophenotypes with significant shared heritability after false discovery rate correction at $p < 0.05$. For the sake of readability, only significant results are displayed. Pearson Correlation between pairs of endophenotypes are also shown. The asterisk represent significant endophenotypic correlation after false discovery rate correction at $p < 0.05$. Table C.1 summarises the endophenotypic correlation and their associated p-values. . . .	87
5.2	Shared heritabilities between both task activations and language-related scores.	Bold rows represents pairs of endophenotypes with significant shared heritability after false discovery rate correction at $p < 0.05$. For the sake of readability, only significant results are displayed. Pearson Correlation between pairs of endophenotypes are also shown. The asterisk represent significant endophenotypic correlation after false discovery rate correction at $p < 0.05$. Table C.3 summarises the endophenotypic correlation and their associated p-values.	89
6.1	Genomic loci associated with heritable language functional connectivities (SmPSS ROIs) using MOSTest.	Lead SNP: ID of the lead SNPs within each locus. Position: position of the SNP in the hg19 human reference genome. mvgnwasP discovery -British-: MOSTest association P value obtained using the discovery sample. mvgnwasP replication -non British-: MOSTest association P value obtained using the independent replication sample. Functional category: Functional consequence of the SNP on the gene obtained from ANNOVAR. 'Central' phenotypes: the phenotypes that contributed most to the multivariate association considering the genome-wide association threshold ($p = 5e - 8$).	99
6.2	Genomic loci associated with heritable language functional connectivities (OcSS ROIs) using MOSTest.	Lead SNP: ID of the lead SNPs within each locus. Position: position of the SNP in the hg19 human reference genome. mvgnwasP discovery -British-: MOSTest association P value obtained using the discovery sample. mvgnwasP replication -non British-: MOSTest association P value obtained using the independent replication sample. Functional category: Functional consequence of the SNP on the gene obtained from ANNOVAR. 'Central' phenotypes: the phenotypes that contributed most to the multivariate association considering the genome-wide association threshold ($5e - 8$).	100
6.3	The SNP-based genetic correlation analysis was estimated using GCTA software, for each pair of central FCs associated to 15q14, 3p11.1 or 10q23.33 genetic loci.	Bold phenotypes represents pair of FCs with significant genetic correlation after false discovery rate correction at $p < 0.05$	104
6.4	Univariate associations of 2 lead SNPs with diffusion MRI indices.	Bold phenotypes represents tracts significantly associated with the lead SNP after Bonferroni correction ($p = 6.94e - 4(0.05/(3 * 9 + 5 * 9))$)	106

7.1	Genomic loci associated with heritable language anatomical connectivities using MOSTest. Lead SNP: ID of the lead SNPs within each locus. Position: position of the SNP in the hg19 human reference genome. mvgwasP discovery -British-: MOSTest association P value obtained using the discovery sample. mvgwasP replication -non British-: MOSTest association P value obtained using the independent replication sample. Functional category: Functional consequence of the SNP on the gene obtained from ANNOVAR. 'Central' endophenotypes: number of the endophenotypes that contributed most to the multivariate association considering the genome-wide association threshold ($5e - 8$). 'Central' endophenotypes can be appreciated in the Table D.1.	125
7.2	MAGMA gene-set analysis. Gene-set attaining statistical significance following Bonferroni control for multiple test.	132
B.1	Heritabilities of brain functional connectivities extracted using SmSSPAs ROIs. Bold phenotypes represents FCs with significant heritability after false discovery rate correction at $p < 0.05$	146
B.2	Heritabilities of brain functional connectivities extracted using OcSS ROIs. Bold phenotypes represents FCs with significant heritability after false discovery rate correction at $p < 0.05$	153
B.3	Heritabilities of brain functional connectivities extracted using SmPSS ROIs. Bold phenotypes represents FCs with significant heritability after false discovery rate correction at $p < 0.05$	154
B.4	Heritabilities of brain anatomical connectivities. Bold phenotypes represents ACs with significant heritability after false discovery rate correction at $p < 0.05$	161
C.1	Pearson Correlation between both resting-state functional connectivity and neural activations. Bold values represents pairs of endophenotypes with significant correlation after false discovery rate correction at $p < 0.05$	169
C.2	Pearson Correlation between both resting-state functional connectivity and language-related scores. No significant correlation was observed after false discovery rate correction at $p < 0.05$	172
C.3	Pearson Correlation between both task activations and language-related scores. Significant correlation was observed after false discovery rate correction at $p < 0.05$	177

D.1	Genomic loci associated with heritable language anatomical connectivities using MOSTest. Lead SNP: ID of the lead SNPs within each locus. Position: position of the SNP in the hg19 human reference genome. mvgwasP discovery -British-: MOSTest association P value obtained using the discovery sample. mvgwasP replication -non British-: MOSTest association P value obtained using the independent replication sample. Functionnal category: Functional consequence of the SNP on the gene obtained from ANNOVAR. 'Central' phenotypes: the phenotypes that contributed most to the multivariate association considering the genome-wide association threshold ($5e - 8$).	180
-----	--	-----

Abbreviations and Acronyms

Glossary of Acronyms and Abbreviations

AAL	Automated Anatomical Labeling
AC	Anatomical connectivity
AF	Arcuate fasciculus
AI	Assymetry index
AICHA	Atlas of Intrinsic Connectivity of Homotopic Areas
AMICO	Accelerated Microstructure Imaging via Convex Optimization
ANNOVAR	ANNOtate VARIation
BA	Brodmann areas
BGEN	Binary GEN
BOLD	Blood-oxygen level dependent
BrainSpan	Developmental stage
CADD	Combined Annotation Dependent Depletion
CAP	Colorado Adoption Project
CG	Candidate-gene
CNS	Central nervous system
CNV	Copy-number variants
CSF	Cerebro-spinal fluid
DEG	Differentially expressed gene
dbGaP	Database of Genotypes and Phenotypes
DLD	Developmental language disorders
dMRI	Diffusion magnetic resonance imaging
DNA	Deoxyribonucleic acid
DTI	Diffusion tensor imaging
DTIFIT	Diffusion Tensor Imaging Fitting
DW-MRI	Diffusion-weighted magnetic resonance imaging
DWI	Diffusion-weighted imaging
DZ	Dizygotic
EEG	Electroencephalography
EPI	Echo planar imaging

eQTL	Expression quantitative trait loci
FA	Fractional anisotropy
FAT	Frontal aslant tract
FC	Functional connectivity
FDR	False Discovery Rate
FIT	Fronto-insular tracts
fMRI	Functional Magnetic Resonance Imaging
FMRIB	Oxford Centre for Functional Magnetic Resonance Imaging of the Brain
FOP	Frontal operculum
FOV	Field-of-view
FSL	FMRIB Software Library
FWHM	Full width at half maximum
GCTA	Genome-wide complex trait analysis
GDC	Gradient distortion correction
GLM	General linear model
GO	Gene Ontology
GREML	Genome-based restricted maximum likelihood
GRM	Genetic relationship matrix
GTE_x	Genotype-Tissue Expression
GWAS	Genome-wide association studies
h^2	Heritability
HCP	Human Connectome Project
HG	Heschl's gyrus
HRF	Hemodynamic response function
ICA	Independent Component Analysis
ICVF	Intra-cellular volume fraction - an index of white matter neurite density
IDP	Image derived phenotypes
IFG	Inferior frontal gyrus
IFOF	Inferior fronto-occipital fasciculus
ILF	Inferior longitudinal fasciculus
INT	Inverse normal transformation
iSECA	Independent SNP Effect Concordance Analysis
ISOVF	Isotropic or free water volume fraction
LD	Linkage disequilibrium
LDSC	Linkage disequilibrium score regression
LH	Left hemisphere
MAF	Minor allele frequency
MD	Mean diffusivity
MEG	Magnetoencephalography
MEGA	Multi-Ethnic Genotyping Array
MLM	Mixed effect linear model
MNI	Montreal Neurological Institute-Hospital

MO	Tensor mode
MOSTest	Multivariate Omnibus Statistical Test
MREC	Multi-Centre Research Ethics Committee
MRI	Magnetic Resonance Imaging
MTG	Middle temporal gyrus
mvGWAS	Multivariate genome-wide association study
MZ	Monozygotic
MetaPhat	Meta-Phenotype Association Tracer
MsigDB	Molecular Signatures Database
NIH	National Institute of Health
NODDI	Neurite Orientation Dispersion and Density Imaging
OD	Orientation dispersion index
OcSS	One-contrast sentences structure
PAC	Primary auditory cortex
PET	Positron emission tomography
PMAT	Penn Matrix Test
PMC	Premotor cortex
PNS	Peripheral nervous system
PPI	Protein-protein interaction
Polyphen	Polymorphism phenotyping
RD	Reading disability
REML	Restricted maximum likelihood
RL	Right hemisphere
RNA	RiboNucleic Acids
ROI	Region of interest
RdN	Reading network
rsfMRI	Resting-state functional magnetic resonance imaging
SE	Standard error
SENSAAS	SENtence Supramodal Areas AtlaS
SLF	Superior longitudinal fasciculus
SLI	Specific language impairment
SNP	Single Nucleotide Polymorphisms
SOLAR	Sequential Oligogenic Linkage Analysis Routines
SPECT	Single photon emission computed tomography
STG	Superior temporal gyrus
STS	Superior temporal sulcus
SmPSS	Supramodal phonology semantics and sentence processing
TBSS	Tract-Based Spatial Statistics
TE	Echo time
tfMRI	Task-based functional magnetic resonance imaging
TLF	Temporal longitudinal fasciculus
TR	Repetition time

UKB	UK Biobank
VBM	Voxel-based morphometry
WM	White matter
WMCA	Word-list Multimodal Cortical Atlas

State of the art

Chapter Outline

1.1	Genetic basis	6
1.1.1	The genome and inter-individual variability	7
1.1.2	From genotypes to phenotype	8
1.2	Neuroanatomy basics	9
1.2.1	Grey and white matter	10
1.2.2	Gyri and Sulci	11
1.2.3	Cells of the brain: From molecule to neurons and from neurons to brain	11
1.3	How to determine the brain architecture? A neuroimaging approach	13
1.3.1	Structural Neuroimaging	14
1.3.2	Functional Neuroimaging	17
1.4	A basic introduction to quantitative and molecular genetics methodology	21
1.4.1	Heritability	21
1.4.2	Endophenotypes	22
1.4.3	Pleiotropy	22
1.4.4	Candidate genes	22
1.4.5	Genome-wide association study	23
1.5	What are the bases of human language? A review of language approaches in human neurosciences	25
1.5.1	Understanding language from aphasia	25
1.5.2	Understanding language by brain imaging	26
1.6	Complex and multifactorial aetiology: the pursuit of brain–language–genetics relationships	33
1.6.1	Heritability of language	34
1.6.2	Developmental language disorder related loci	36
1.6.3	Genome-wide association studies of language abilities	37
1.6.4	Moving Forward: Emergence of imaging–genetics approach	39
1.7	Aim of this thesis	43

Language is a central component of human life. We live immersed in language as it underlies most human interactions; we use it to structure our thoughts, to communicate and interact with each other. Because it is so ubiquitous, it can seem simple, yet language is insufficiently understood as a phenomenon and process. Across different contexts and forms, researchers emphasize the special character of language compared to other human capacities. Accordingly, many difficult questions surround the topic of language. One of these concerns its origin: nature versus nurture. Language has an inherent dimension of learning, but also a natural instinctive tendency that is proven by the babbling of infants, for instance. Interestingly, this is not observed for other cultural practices and is specific to language. Language therefore exists in a unique space, at a junction between nature and culture, between individual and society.

As language is a phenomenon that is both biological and social, it can be approached with a wide variety of methods that includes methods and explanatory models of the formal and natural sciences, as well as the tools offered by the human and social sciences. The first approach focuses on the cognitive bases of language and the different aspects of the biological endowment that allow human species to learn and use languages. It weaves privileged relationships with cognitive neuroscience, human biology and also formal sciences which provides precise instruments for its modeling. The second approach on the other hand, will focus on the relationship between language and society, history, philologies, and literatures. The very nature of language, its complexity, its hybrid status justifies a plurality of approaches and methods [Luigi, 2021]. The present thesis is part of the first aforementioned approach and its main goal is to modestly contribute to the neurobiology of language question.

Interest in mapping genes for language skills has largely been driven by the need to understand the biological causes of language disability. Indeed, until recently, the biological foundations for language were mainly sought in lesion studies, where patients showed specific language impairments such as developmental language disorders and dyslexia. Imaging-genetics is an emerging scientific field that benefits from the maturity of non-invasive imaging techniques such as magnetic resonance imaging and high-throughput genotyping by DNA chips or sequencing. These techniques allow the systematic acquisition of anatomical and functional characteristics of an individual's brain along with its genotype. The ability to acquire these data in general population allows studies to be conducted on open questions of clinical or basic neuroscience. Data science tools - from big data hardware to analysis algorithms - open up the possibility of studying how genetics and environment modulate human brain function or anatomy and behaviour. With the recent availability of large-scale cohorts, it has made it possible to study language in the general population. The specific objective of this thesis is to highlight genes that underlie human language, contributing thus to our understanding of different processes required for language.

1.1 Genetic basis

In this section, we will present basic principles of genetics useful for understanding imaging-genetic studies. It has no intention to cover the genetic complexity but intend to provide brief definition of genetic terms necessary for the comprehension of the present work.

1.1.1 The genome and inter-individual variability

Genetic information is encoded in a macro-molecule called deoxyribonucleic acid (DNA). It is present in almost all living organisms, present in every cell (except for enucleated cells) and identical throughout the body (except for gametes). A nucleotide is the basic building block of nucleic acids. DNA consists in two complementary strands of nucleotides. Its double helix structure was discovered in 1953 by Watson and Crick [Pray, 2008]. In human DNA, each strand comprises approximately 3,200,000,000 nucleotides. It is made up of four elementary bases called nucleosides: Adenine (A), Thymine (T), Guanine (G) and Cytosine (C) associated with a sugar (deoxyribose) and a triphosphate, the whole forming nucleotides. These bases work in pairs, one purine (Adenine or Guanine) paired with one pyrimidine (Thymine and Cytosine respectively) which allows physical stability of the molecule formed. The DNA macro-molecule is divided in humans into 22 pairs of autosomes (non-sexual chromosomes) and 1 pair of sexual chromosomes (XX or XY) (Fig. 1.1).

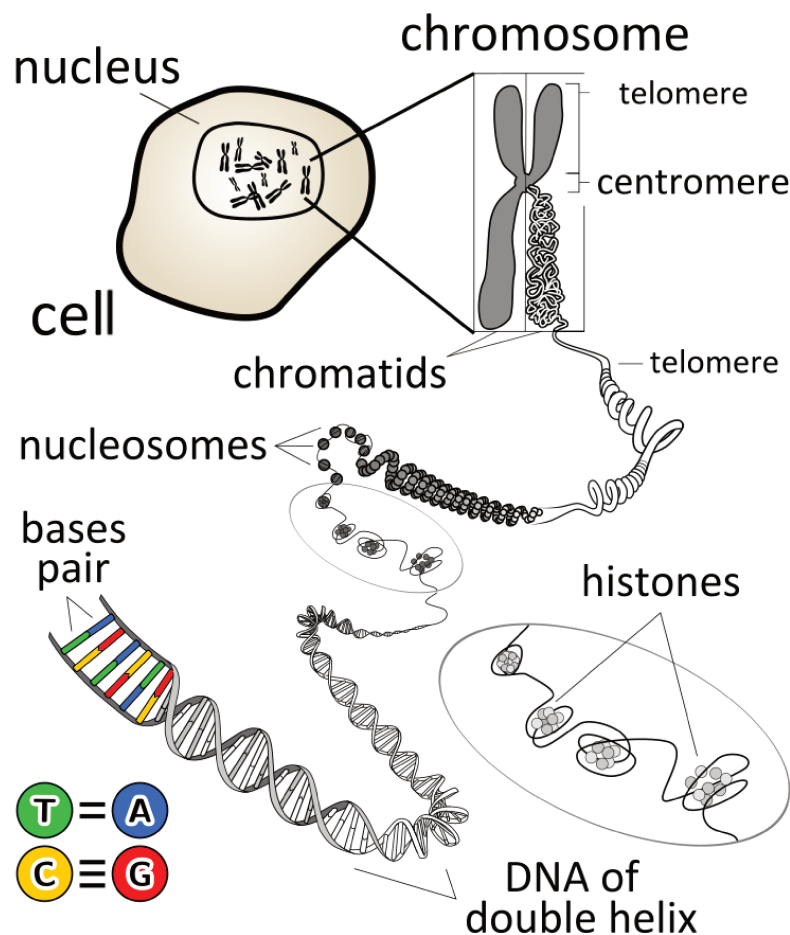


Figure 1.1 – Description of the chromosome structure. (Source: <https://www.genome.gov/>).

Along the DNA molecules, the genome has functional units called genes. These genes are DNA sequences made of exons and introns, transcribed into RiboNucleic Acids (RNA). Subsequently, only exons are translated into proteins. The human genome is estimated to contain approxim-

ately 23,000 genes. And these genes are separated by intergenic regions or non-coding DNA containing expression regulatory sequences such as promoters and enhancers and thus are involved in regulation of nearby genes expression. Nonetheless, the function of the major part of this non-coding DNA is still unknown. The sequence of nucleotide bases is important regarding the integrity of the genome. Its alteration can have consequences on the functional units produced. Genetic variations from a human reference sequence exist for each individual and are called genetic polymorphisms. Multiple mechanisms of polymorphisms are known. Relevant to our work, Single Nucleotide Polymorphisms (SNPs) are the smallest form of polymorphism and affects only one base pair. Except for rare cases, SNPs are bi-allelic polymorphisms. They represent markers of choice for researchers in order to establish a dense and precise mapping of the genome variations such as dbSNP [Sherry et al., 2001], HapMap [Gibbs et al., 2003], the 1000 genomes project [Consortium et al., 2015]. A SNP is categorised as common or rare variant and is characterised by its chromosomal position, its alleles and its minor allele frequency (See below for a definition).

More definitions to understand imaging-genetics

- Imputation is the process consisting in statistical inference of unobserved genotypes.
- Linkage disequilibrium (LD) is the non-random association of alleles from two or more polymorphic loci.
- Linkage analysis is a genetic approach aiming at establishing linkage between genetic markers and traits or diseases by examining patterns of inheritance among individuals within a family who exhibit the trait of interest and testing whether they co-segregate with the trait of interest, in a manner that is unlikely to have occurred by chance.
- Population stratification or genetic structure represents the presence of differences in allele frequencies between sub-populations in the studied population, due to their different genetic ancestries and can cause false positives in association studies.
- Hardy-Weinberg equilibrium models the behaviour of allelic and genotypic frequencies for a polymorphism. It states, under certain assumptions, that the allelic and genotypic frequencies of a polymorphism are stable within the population over generations.
- Minor allele frequency (MAF) is the frequency of the less frequent allele of a bi-allelic SNP, and it can vary depending on the considered population.
- epistasis is the phenomenon that describes an interaction between genes and/or genetic variations.
- pleiotropy is defined as the fact that one gene can have several different effects or functions, non necessarily independent from one another.

1.1.2 From genotypes to phenotype

Understanding the path from genotype to observable characteristics –phenotypes– is a major concern of molecular biology. A phenotype is almost always a multifactorial trait and results from the contribution of many proteins. The proteins are biological macro-molecules, made of a chain of amino acids and present in all living cells. A gene code for one protein, according to the genetic code: each amino acid in a protein is encoded by one or several triplets of nucleotides or codons. The process by which a gene is expressed into a protein is made of several steps: from each gene many messenger RNAs (ribonucleic acids) are produced (transcription), and are then edited (alternative splicing), translated into amino acid macro-molecules (translation) and finally matured in proteins (maturation). Interactions between proteins are important as they play a crucial role in cellular functions and biological processes, and produce phenotypes. Phenotype variability depends on the nature and/or the (relative) quantity of the potentially 23,000 concerned proteins.

Research on gene-phenotype association test for a non-random correlations between a phenotypic trait (qualitative or quantitative) and specific allelic variants. It can be carried out for just a single genetic marker (SNP), multiple sites within a gene, or across the entire genome. The research about mechanism underlying gene to phenotype manifestation is, on the other hand, beyond the scope of this study.

1.2 Neuroanatomy basics

The human brain is composed of three main parts: the cerebrum, cerebellum, and brainstem (Fig. 1.2a). The cerebrum is the largest part of the human brain and is made of two cerebral hemispheres separated by a deep *longitudinal fissure*, but remain connected via the corpus callosum. It has a folded appearance and is involved in higher functioning. It controls somatosensory, motor, language, cognitive thought, memory, emotions, hearing, and vision. Each hemisphere has a number of folds called **gyri** separated by grooves called **sulci**. They are divided into four main lobes separated by the *central sulcus*, *parieto-occipital sulcus*, and *lateral fissure (Sylvian fissure)*. The frontal lobe is situated anterior to the central sulcus while the parietal lobe is posterior to the central sulcus. The temporal lobe is separated from the frontal and parietal lobes by the lateral sulcus. Finally, the occipital lobe, located at the rear of the cerebrum, is separated from the parietal lobe by the parieto-occipital sulcus. These lobes, named after the bones of the skull that overlie them are distinguished by colors in Fig. 1.2b.

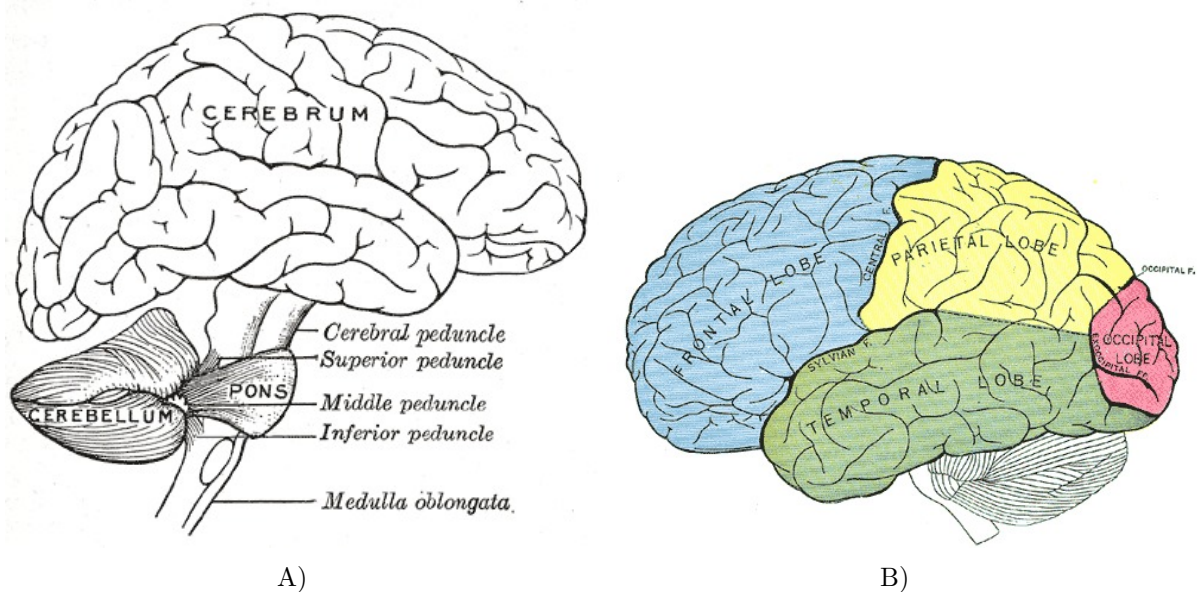


Figure 1.2 – Neuroanatomy of the brain ([Bui and Nguyen, 2019]). A) Brain, Encephalon, Connections of the several parts of the brain, Cerebrum, Cerebellum, Pons, Cerebral; Superior; Middle; Inferior Peduncle, Medulla oblongata. B) Principal fissures and lobes of the cerebrum viewed laterally, Frontal Lobe, Parietal Lobe, Temporal Lobe, Occipital Lobe.

1.2.1 Grey and white matter

At a macroscopic scale, the human brain appears to be composed of grey matter and white matter (See Fig. 1.3a). The grey matter contains the neuronal cell bodies of the central nervous system and serves to process information. It is located on the surface of the gyri and nuclei (collection of neurons in the central nervous system). A distinction is made between cortical grey matter located in the cerebral cortex and subcortical ones. The latter are a group of diverse neural formations that lie beneath the cerebral cortex which includes the Hippocampus, Amygdala, and the basal ganglia. The white matter contains the fibre tracts of the neuronal axons wrapped in myelin and is found in the inner layer of the cortex. It is essential for impulse conduction as it supports the communication between various grey matter areas. It is composed of three types of fibres: commissural fibres, projection fibres, and association fibres. Commissural fibres connect the left-right cerebral hemispheres and are composed of anterior forceps that connect the lateral and medial surfaces of the two frontal lobes, and posterior forceps that connect the two occipital lobes. Projection fibres consist of efferent and afferent fibres that connect higher and lower brain centres. Association fibres interconnect lobes on the same side. Two subtypes of these association fibres exist: short association fibres – also called 'U'-fibres– lie beneath the grey substance of the cortex and connect adjacent gyri, while long association fibres connect distant areas of the same hemisphere.

1.2.2 Gyri and Sulci

The surface of the cerebrum is folded with a large number of gyri and sulci that give it its wrinkled appearance (Fig. 1.3b). These cortical foldings are variable from one individual to

another. The folding of the cerebral cortex creates gyri and grooves (sulci) that separate brain regions and increase the cortical surface area. This increasing of the surface area of the brain allows more neurons to be packed into the cortex so that it can process more information.

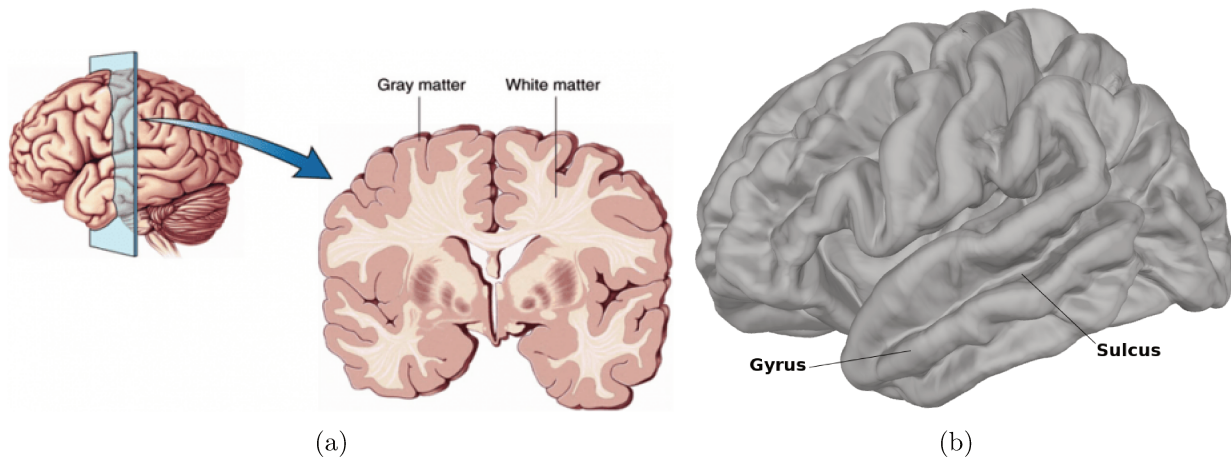


Figure 1.3 – **Brain basics.** Example of a) the grey and white matter of the central nervous system (source : <https://difference.guru/difference-between-white-and-gray-matter/>), b) sulcus and gyrus.

1.2.3 Cells of the brain: From molecule to neurons and from neurons to brain

The central nervous system is made up of neurons and glial cells. Neurons are excitable cells, they vary in shape and size and can be classified according to their morphology and function. They have three basic parts: a cell body and two extensions called an axon and a dendrite. A nucleus is within the cell body, which controls the cell's activities and contains the cell's genetic material. Synapse is a structure that permit a transmission of signals (electric or chemical) from one neuron to another. The axon carries the information to another nerve cell (post synaptic neurons), while the dendrite receive information from other nerve cells (presynaptic neurons). Glial cells are nerve cells. Different types of glial cells are present in the central nervous system (CNS) and the peripheral nervous system (PNS). There are four types of glial cells in the CNS including oligodendrocytes, astrocytes, Ependymal and microglia. The microglia is found in all regions of the brain and spinal cord. It is implicated in the immunity process and plays an important role in neuron pruning during the neurodevelopment process. The oligodendrocytes and Schwann cells are found in the CNS and PNS respectively. They provide myelination to axons in the central and peripheral nervous system, respectively. The astrocytes are the most abundant type of macroglial cells in the central nervous system. The major nervous cell types are depicted schematically in the Fig. 1.4a.

The network of neurons is organised in a hierarchical way at different scales. Although this network is highly structured, it is important to note that the details of the structure can vary, both between individuals of the same species and over time in a given individual due to synaptic plasticity. Our understanding of the organisation and function of the different levels is strongly influenced by the available experimental techniques that will be described in the next section.

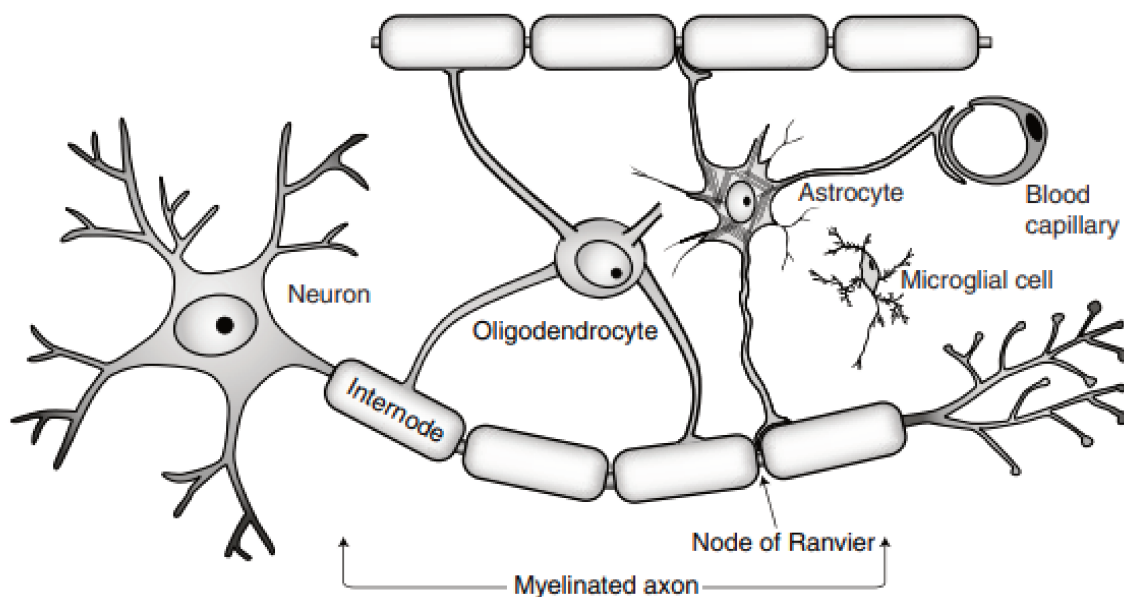


Figure 1.4 – Schematic representation of the major cellular elements of the white matter (Source: [Johansen-Berg and Behrens, 2013]). In this drawing, a neuronal cell body in the grey matter extends several short dendritic processes and a single, long axonal process. The axon, in the white matter, is surrounded by an insulating myelin sheath. This sheath is produced by oligodendrocytes. Generally, oligodendrocytes extend many processes, each of which wraps around a segment of an axon to form an internode. Internodes are separated from each other by small unmyelinated regions called nodes of Ranvier. This specialized organisation of the myelin sheaths promotes saltatory conduction whereby electrical impulses “jump” from node to node as they pass along the axon from the neuronal cell body to the axon terminus. The white matter also contains astrocytes and microglia. Astrocytes are star-shaped glial cells that help maintain homeostasis by regulating local ion concentrations and extracellular fluid volume within the CNS. In the white matter, they extend processes to nodes of Ranvier and to blood vessels, where they contribute to blood–brain barrier function. Microglial cells, which are considered the immune cells of the CNS, help “mop up” debris and dying cells.

1.3 How to determine the brain architecture? A neuroimaging approach

The last century has witnessed the development of brain investigation techniques that include the use of various systems such as the positron emission tomography (PET), single photon emission computed tomography (SPECT), magnetic resonance imaging (MRI), functional MRI (fMRI), electroencephalography (EEG), or magnetoencephalography (MEG). These *in-vivo* predominant approaches are usually coupled with behavioural and cognitive ones to study how the brain functions [Friston, 2009]. They make it possible to discover different maps in the brain and can be used to determine the cognitive brain architecture by understanding the brain organisation such as brain connectivity, or spatiotemporal dynamics underlying cognitive functions [Friston, 1994, Liang and Lauterbur, 2000]. Relevant to this study, we will focus in this section on the presentation of the structural and functional MRI techniques.

Voxel-wise approach was for a long time the dominant approach to analyse brain images [Friston et al., 1994]. Such an approach requires massive computational and memory resources. In order to make inference with this approach it is usually necessary to correct drastically for a huge number of multiple comparisons which leads to low statistical power. **Image derived phenotypes (IDP)** approach has gained popularity and is increasingly used in the context of population imaging. Indeed, this approach that consists on summarising distinct measures of brain structure and function such as volume of grey matter in distinct brain regions, structural and functional connectivity metrics [Alfaro-Almagro et al., 2018, Elliott et al., 2018, Miller et al., 2016], can easily be shared by the neuroscience research community and alleviates the limitations of the voxel-wise approach: it mitigates the loss of statistical power compared to voxel-wise approach. For example assessing the functional connectome between region of interests significantly reduces the number of variables making it computationally and statistically more within our reach; For predictive modelling, it certainly has statistical benefits. The high dimensionality of the signal is a challenge in machine learning and IDPs approach is one way to circumvent it.

1.3.1 Structural Neuroimaging

Structural magnetic resonance imaging. Structural magnetic resonance imaging provides anatomical images with high contrast between grey and white matter.

Morphometric analysis of the structural neuroimaging is widely used for the study of the cortex; the fine layer of gray matter. Multiple IDPs might be derived using different softwares like Freesurfer (gray matter volume, surface local gyrification index) [Fischl and Dale, 2000] (<https://surfer.nmr.mgh.harvard.edu/>), Morphologist (sulcal size, pits, depth or number of *plis de passage*) (<https://brainvisa.info/web/morphologist.html>) to neuropathologies or behaviour [Lerch et al., 2017]. These features are obtained at the voxel-level and can be analyzed using either a voxel-wise [Ashburner and Friston, 2000] or pre-defined region of interest [Desikan et al., 2006] approaches. Multiple applications are possible, such as: cognitive ageing [Ritchie et al., 2015], brain evolution across species [Evans et al., 2006], general intelligence [Cox et al., 2019], and brain disease [Thompson et al., 2007].

Diffusion-weighted magnetic resonance imaging. Diffusion-weighted imaging is a non-invasive, *in vivo* MRI technique that quantifies water diffusion in biological tissues. It is used to measure and map the local structures by tracking the movement of the water molecules along the white matter bundles. It is used to identify the different fiber tracts in the human brain.

In general, the diffusion in brain tissue differentiates between tissues based on the diffusion of water molecules within them. The diffusion process is different depending on which type of tissue it is observed. Regarding the grey matter, the diffusion is unrestricted -isotropic diffusion-; the water molecules move in any possible direction. Concerning the white matter, the diffusion is restricted -anisotropic diffusion-; the displacement of water molecules is not random due to the presence of biological structures such as cell membranes, filaments, and nuclei (See Fig. 1.5).

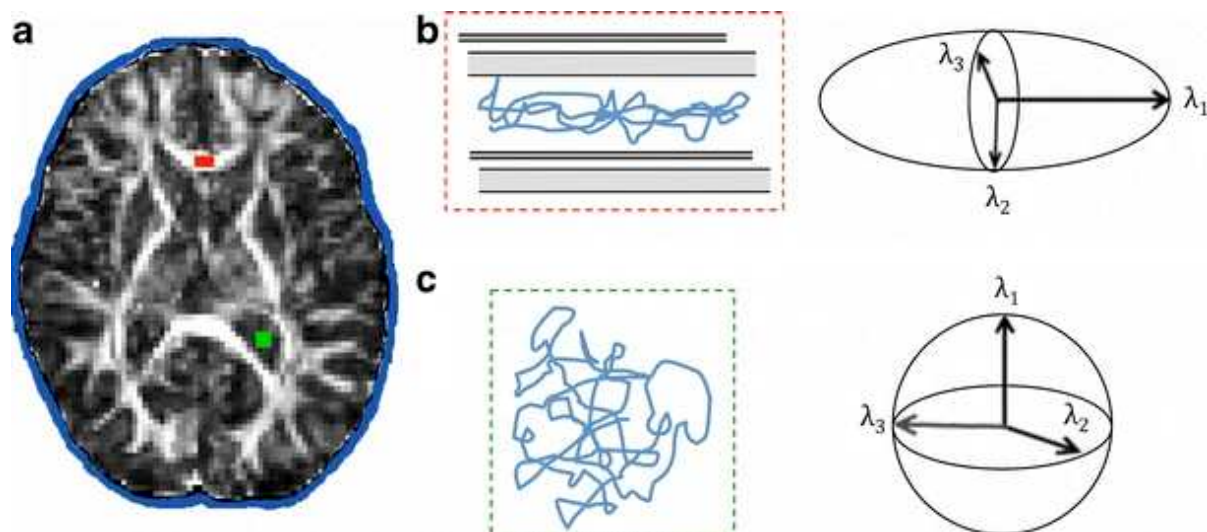


Figure 1.5 – **Isotropic and anisotropic diffusion in the brain** (Source: [Pandit et al., 2013]). Fractional anisotropy image of a preterm born child reveals the nature of water diffusion within different brain tissue compartments (a). In the white matter of the corpus callosum (red), diffusion occurs preferentially along the axonal fibres, resulting in anisotropic diffusion (b). In the ventricular cerebro-spinal fluid (CSF; green), diffusion is unhindered and can be described as isotropic (c). Diffusion tensor ellipsoids representing anisotropic and isotropic diffusion are shown in b and c, respectively. Each tensor is expressed by three eigenvectors with values λ_1 , λ_2 and λ_3 . In isotropic diffusion $\lambda_1 = \lambda_2 = \lambda_3$, whereas in anisotropic diffusion the long axis of the ellipsoid aligns with the underlying white matter, and λ_1 is greater than λ_2 and λ_3 .

Within a single voxel, the shape of the diffusion profile can be described geometrically as an ellipsoid (the tensor) calculated from the diffusion coefficient values (eigenvalues, $\lambda_1 - \lambda_3$) and orientations (eigenvectors, $v_1 - v_3$) of its three principal axes (Fig. 1.6) [Catani and Forkel, 2019]. The mean diffusivity which represents the average water molecular displacement within a voxel, is obtained by taking the mean of the three eigenvalues ($\lambda_1, \lambda_2, \text{ and } \lambda_3$). The degree to which diffusion is directionally constrained can be determined based on the ratio of the eigenvalues to one another (e.g. fractional anisotropy). Overall, these indices provides complementary information about white matter micro-structures [O'Donnell and Westin, 2011, Thomason and Thompson, 2011].

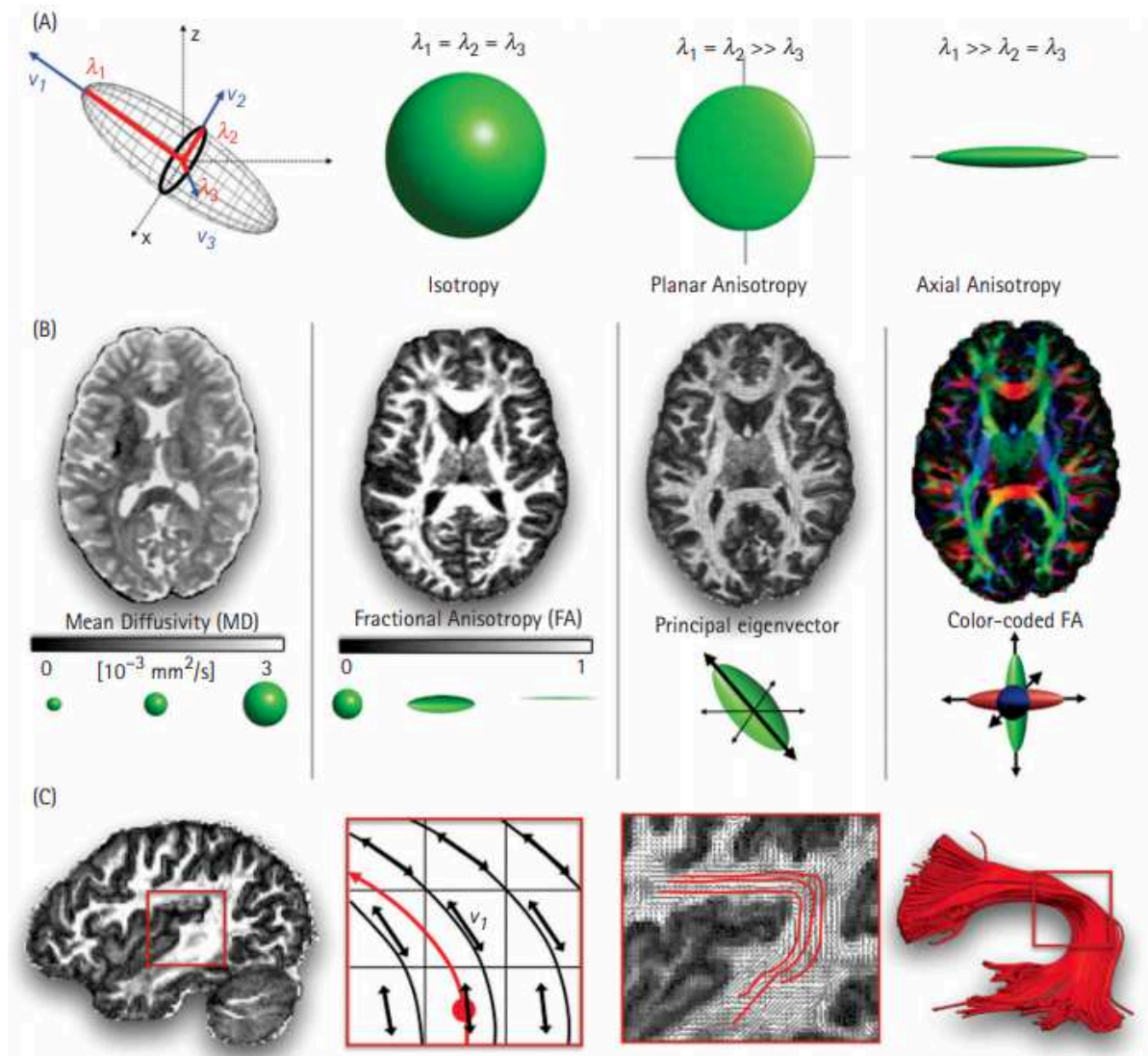


Figure 1.6 – **Principles of diffusion tensor imaging (DTI) tractography** (Source: [Catani and Forkel, 2019]). (A) Visualization of the diffusion tensor as an ellipsoid. The size and the shape of the tensor are defined by the three eigenvalues (1, 2, 3 in red), while the spatial orientation is described by the three eigenvectors (v_1 , v_2 , v_3 in blue). In biological tissues, the tensor can vary between three possible configurations: (1) isotropy or weak anisotropy (equal or similar diffusivity along the three eigenvalues) is commonly observed, for example, inside the ventricles (isotropy) or in the gray matter (weak anisotropy); (2) planar anisotropy (unequal diffusivity between of one eigenvalue and the other two) is common in voxels containing, for example, two groups of crossing or diverging fibers; and (3) axial anisotropy is typical of voxels containing parallel fibers (unequal diffusivity between the axial eigenvalue and the two perpendicular eigenvalues). (B) Exemplified indices extracted from diffusion data, such as mean diffusivity (MD), fractional anisotropy (FA), principle eigenvector, and color-coded FA maps. (C) Streamline tractography is based on the assumption that in each white matter voxel the principal eigenvector (black arrows) is tangent to the main trajectory of the underlying fibers (black lines). Starting from a region of interest (red circle), the tractography algorithm propagates, voxel by voxel, a streamline (red) by piecing together neighboring principal eigenvectors. An example is shown in the neighboring panel where the streamlines are tracked on a principal eigenvector map and the tractography reconstruction of the arcuate fasciculus is visualized as 3D streamtubes.

1.3.2 Functional Neuroimaging

Functional Magnetic Resonance Imaging (fMRI) is a non-invasive technique that measures brain activity over time by detecting changes in blood-oxygen level dependent (BOLD) signal. The local oxygen level in the brain is coupled to the metabolic needs of neurons while these metabolic changes can be consequent due to task-induced cognitive state changes such as language or visual task. BOLD signal variations can also be observed in the case of free processes in the brain of an individual at rest [Glover, 2011].

Task-oriented functional magnetic resonance imaging. The conventional task-evoked fMRI experiments use different stimuli such as visual or auditory to alternately induce two or more different cognitive states of the brain. For example in a two-condition design, one state is called the experimental condition, while the other is denoted the control condition. The goal is to test the hypothesis that the signals differ between the two states [Glover, 2011]. For example, the individual perform a task 'A' involving processes 'P₁', 'P₂', 'P₃', a task 'B' involving processes 'P₁', 'P₂'. The subtraction of the brain activation images obtained in tasks 'A' and 'B' makes it possible to isolate the region carrying out the 'P₃' process (See Fig.1.7). This typical fMRI experiments allowed great progress in the understanding of the brain organisation. Nevertheless, task-induced fMRI experiments has revealed limits. The association between a high level function and a brain region is much more complex than what the subtractive approach can reveal. A brain region can be characterised by its multi-functional profile. Meta-analysis of a large number of studies on cognitive traits such as language are crucial to determine the degree of involvement of brain regions and give us a more global view. Moreover, the manifestation of a cognitive trait is the result of interaction within a network rather than a single area.

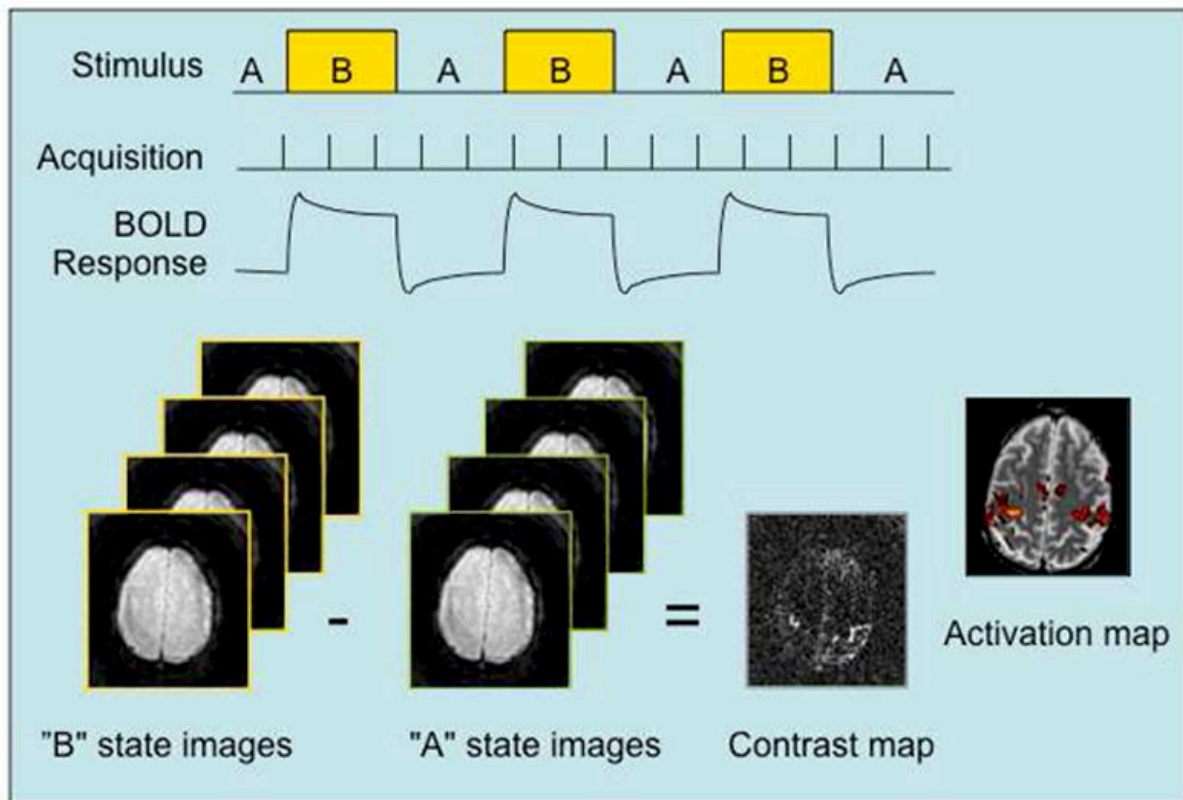


Figure 1.7 – Figure extracted from [Glover, 2011]. Block design fMRI experiment. A neural response to the state change from A to B in the stimulus is accompanied by a hemodynamic response that is detected by the rapid and continuous acquisition of MR images sensitised to BOLD signal changes. Using single- or multi-variate time series analysis methods, the average signal difference between the two states is computed for the scan and a contrast map generated. A statistical activation map is finally obtained using a suitable threshold for the difference; the map depicts the probability that a voxel is activated given the uncertainty due to noise and the small BOLD signal differences.

1.3.2.1 Resting-state functional magnetic resonance imaging

Resting-state functional MR images captures low-frequency fluctuation (< 0.1 Hz) in blood oxygenation and the significance of these fluctuation was initially presented by [Biswal et al., 1995]. It is paradigm free and as such it is easier to implement in very large cohorts like the UK Biobank. Indeed, during the resting-state scan, no explicit task is being performed and depending on the experience protocol, the participants are instructed to keep their eyes fixated on a crosshair, to relax, and to think of nothing in particular [Miller et al., 2016].

Functional images acquired during rest has recently emerged as a powerful tool and is a promising universal marker of brain function [Biswal et al., 2010]. Different network were identified using a paradigm free approach including the well known default mode network (See Fig. 1.8). This resting-state network was initially identified by [Raichle et al., 2001] using PET data and was further identified by fMRI by [Greicius et al., 2003]. In [Raichle et al., 2001] work, the researchers highlighted brain regions that were more active at rest than during task-oriented experience and observed that their activity decrease when cognitive tasks were performed. This

led to the hypothesis that a 'default mode' of the brain remained active in an organised fashion during the resting state.

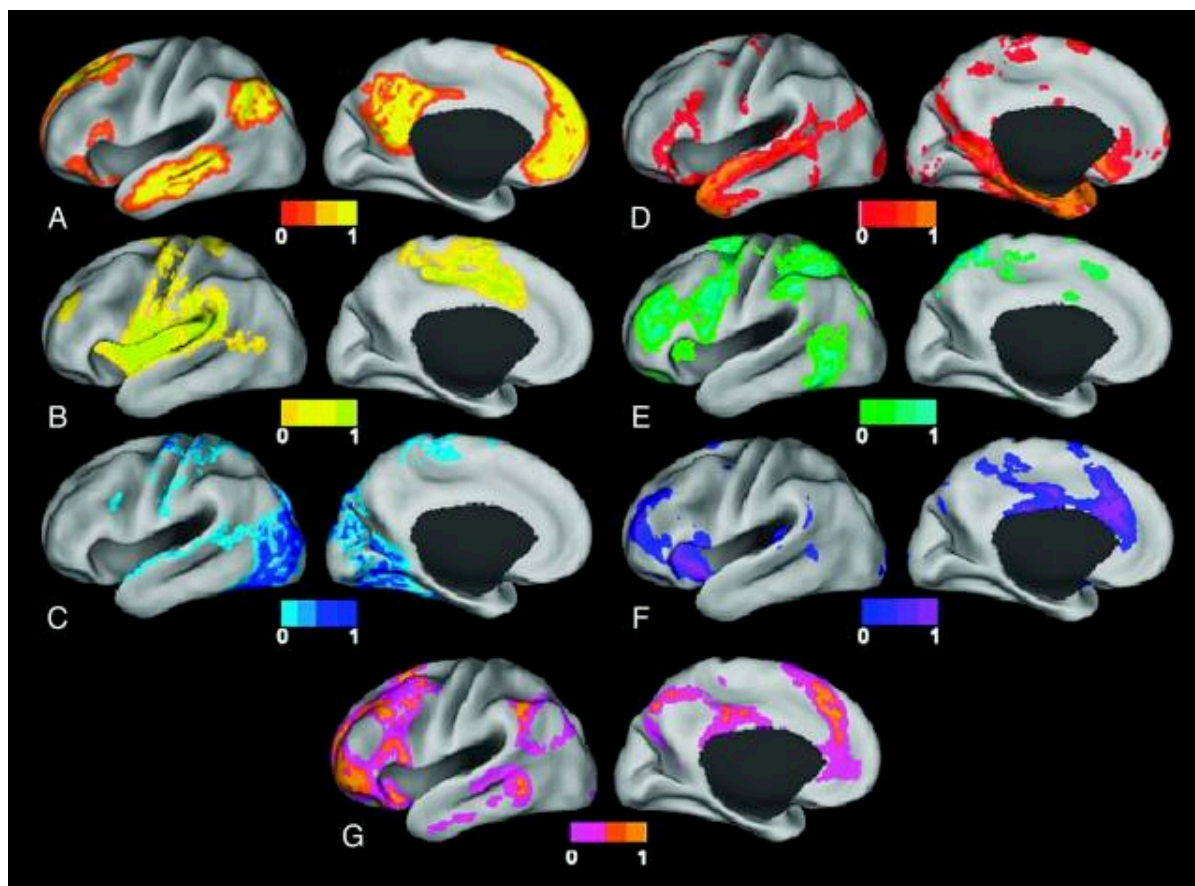


Figure 1.8 – Figure extracted from [Lee et al., 2013]. **Surface plots of resting-state networks.** A: Default mode network, B: Somatomotor network, C: Visual network, D: Language network, E: Dorsal attention network, F: ventral attention network, G: Fronto-parietal control network.

Despite low frequency blood-oxygen-level-dependent (BOLD) signals acquired in resting state condition and that rsfMRI does reflect large-scale circuit organisation [Biswal et al., 1995, Fox and Raichle, 2007], they are less used to study specific brain functions. However, a growing body of evidence suggests that there is a close correspondence between resting state networks and known cognitive task activation maps [Cole et al., 2014, Smith et al., 2009]. For example, [Tavor et al., 2016] showed that a prediction model relating FC estimated from rsfMRI to task-based activations can accurately predict individual differences in brain activity for unseen participants. In the [Smith et al., 2009] work, the authors analysed data from healthy volunteers comprising task maps from BrainMap database of functional imaging studies and resting-state fMRI maps. These maps were then associated with each other using (Pearson) spatial cross-correlation. They found close correspondence between the independent analyses of resting and activation brain dynamics. For example, the resting-state 'default mode network' were found to correspond to negative contrasts in cognitive paradigms. More specifically regarding language, several studies sought the brain dynamics and showed that resting-state FC map could map

language networks previously identified in task-based imaging studies [Hampson et al., 2006, Kelly et al., 2010, Koyama et al., 2010, Xiang et al., 2010]. Especially, [Xiang et al., 2010] used resting-state FC to infer the functional organisation of Broca’s area and the perisylvian language networks by investigating their functional correlations and reported a clear topographical FC pattern in the left middle frontal, parietal, and temporal areas. On the other side, several studies have examined the relationship between resting-state FC and behavioural scores. Relevant to our work, these studies have shown that language-related performances like basic reading abilities are correlated to the strength of resting-state FC [Cheema et al., 2021, Cross et al., 2021, Koyama et al., 2011, Stevens et al., 2017]. [Koyama et al., 2011] reported a positive correlation between the reading scores and the resting-state FC between the fusiform gyrus and the inferior frontal gyrus (IFG) -pars opercularis-, belonging to the reading network (RdN). [Cheema et al., 2021] reported a positive correlation between reading fluency and the strength of FC between the left IFG and left supplementary motor area in a group without reading impairment. They also found that the stronger the FC between the supramarginal and the angular gyri, the higher is the word recognition accuracy score. Moreover, [Cross et al., 2021] showed that the FC between the pars triangularis and the right fusiform gyrus was positively correlated with rapid automatized naming performance. Finally, [Stevens et al., 2017] reported a positive correlation between accuracy on word classification and resting-state FC strength of the left occipitotemporal sulcus and Wernicke’s area. Although the resting-state FC and behavioural (especially reading abilities) relationship are consistent through studies, the generalization of these findings should be taken with caution. First, the results might be sample dependant and not representative of the larger population of adult reader. Second, the reading tests represents relatively simple reading process and might not be generalized for more complex process (i.e. reading comprehension). Next, recent works suggest that mapping perisylvian language regions can be accomplished using either tbfMRI or rsfMRI [Branco et al., 2016, Jones et al., 2017, Lemée et al., 2019, Lu et al., 2017, Park et al., 2020, Tie et al., 2014]. Indeed, rsfMRI has been applied to mapping eloquent areas in surgical patients [Jones et al., 2017] and stroke patients [Klingbeil et al., 2019], where performing tbfMRI is not an option. Overall, these studies investigated the relevance of the rsfMRI approach in studying specific brain functions such as language, including its relationship with behavioral scores, task-based network at both individual and group levels. They also highlighted its advantages over task-based design, when certain population has difficulties to engage in specific tasks.

State-of-the-art analysis methods.

The preprocessing of resting-state functional MR images generally includes slice-timing correction, realignment, motion correction, high-pass temporal filtering, brain masking, spatial smoothing, co-registration to individual space using T1 template and registration to the MNI space to achieve spatial concordance between subjects.

Once these preprocessing steps are achieved, several methods can be used to analyse resting-state data. Each one has its own inherent advantages and disadvantages. In this section, we will provide a brief overview of some of the statistical and mathematical state-of-art approaches previously applied to rsfMRI data (For further discussion, see [Lee et al., 2013]).

- Data-driven approaches based on k -means or Ward clustering [Thirion et al., 2014], Independent Component Analysis (ICA) approaches [Beckmann and Smith, 2004, Calhoun et al., 2001], or dictionary learning [Mensch et al., 2016].
- Pre-defined reference anatomical atlases such as Automated Anatomical Labeling (AAL) [Tzourio-Mazoyer et al., 2002], or connectivity-based cortical landmarks such as SENTence Supramodal Areas Atlas (SENSAAS) [Labache et al., 2019], Atlas of Intrinsic Connectivity of Homotopic Areas (AICHA) [Joliot et al., 2015], Word-list Multimodal Cortical Atlas (WMCA) [Hesling et al., 2019].
- Spheres of radius centred at coordinates identified using probabilistic activation maps identified during task-oriented fMRI experiments [Power et al., 2011].

1.4 A basic introduction to quantitative and molecular genetics methodology

In this section, we will briefly introduce basic genetic concepts useful for understanding imaging-genetic studies. Its unique aim is to provide the reader with sufficient knowledge to well understood the different studies presented further and is therefore not intended to be exhaustive.

1.4.1 Heritability

A pivotal question in biology is the heritability of a particular trait. In quantitative genetics, neuroimaging endophenotypes variance (among other traits) can be attributed to either genetic or environmental factors or their interactions. Heritability is defined as the proportion of variance for a particular phenotype, measured at a particular age, that is attributable to genetic variation [Visscher et al., 2008]. An observed phenotype is usually represented as the sum of the unobserved genotype (G) and unobserved environmental factors (E): Phenotype (P) = Genotype (G) + Environment (E). Likewise, the phenotypic variance of the observable phenotypes (σ_P^2) can be expressed as the sum of effects: $\sigma_P^2 = \sigma_G^2 + \sigma_E^2$. The broad sense heritability of a trait (H^2) is the proportion of phenotypic variance attributable to all sources of genetic variance: $H^2 = \sigma_G^2 / \sigma_P^2$. The genetic variance can be partitioned into the variance of additive genetic effects (breeding values; σ_A^2), of dominance (interactions between alleles at the same locus) genetic effects (σ_D^2), and of epistatic (interactions between alleles at different loci) genetic effects (σ_I^2) [Visscher et al., 2008]: $\sigma_G^2 = \sigma_A^2 + \sigma_D^2 + \sigma_I^2$. The narrow-sens heritability of a trait (h^2) is the proportion of phenotypic variance attributable to additive genetic: $h^2 = \sigma_A^2 / \sigma_P^2$. It is the focus of most modern genetic investigation as a high proportion of genetic variance is additive [Hill et al., 2008]. Heritability estimates vary between 0 and 1 indicating either no genetic influence at all or complete genetic determination. In order to judge the value of endophenotypes and prioritise their study, it is common to rely on its heritability estimates. Indeed, a large and significant measure of heritability indicates that the trait of interest is significantly under genetic control and might be eligible for more extensive molecular genetic studies. Heritability is an important genetic concept however, its interpretation should be taken with caution.

First, heritability might be population-dependant in such, it summarises the strength of genetic influence on a specific trait measured in a specific population and may not be generalised to other cohorts where environmental influences and ancestry may differ [Thompson et al., 2013]. Second, heritability estimates is not much informative about which specific genes contribute to a trait, how many genes are involved, or the impact of any gene on it [Thompson et al., 2013]. For example, human height is a trait that is influenced by many genes with small effect [Yang et al., 2010].

1.4.2 Endophenotypes

Behaviour measures like language performance scores are not the direct expression of genetic effects. It is the manifestation of the combination of physiological characteristics as a result of molecular and cellular effects in the brain. Sometimes the physiological characteristic deficiencies may precede illness symptoms [Glahn et al., 2007, Gottesman and Gould, 2003] and they may be considered as biomarkers. The idea behind the concept of endophenotype is that these quantitative biomarkers considered as intermediate phenotype are more closely associated with gene effects than the complex phenotypic expression of a disease. The endophenotypes, like phenotypes, can be studied as regards heritability or genetic association. Such measures help our understanding of the individual neuroimaging variability and provide insight into genetic effects on the central nervous system. A certain number of putative endophenotypes have been proposed for functional and structural brain such as measures of structural and functional connectivity, and brain asymmetry.

1.4.3 Pleiotropy

Pleiotropy refers to the genetic effect of a single gene on two or more (endo)phenotypes. Neuroimaging allows us to generate numerous endophenotypes at a time regarding one individual. Pleiotropic effects have been observed in imaging-genetic studies when considering phenome-wide association studies. Indeed, these ones allow us to identify association of a locus with interrelated endophenotypes.

1.4.4 Candidate genes

The candidate-gene (CG) design was the first and is still a widely used approach in neuroimaging-genetics field. It involves genotyped markers in genes chosen *a priori* and hypothetically related to the trait of interest. For example and relevant to our study, the well-known language related gene *FOXP2* was identified in the KE family, in which multiple close relatives suffered from similar disruptions of speech and language skills [Fisher et al., 1998]. Further candidate-gene studies focused on this particular gene because of *a priori* hypotheses about its etiological role in developmental language disorders (DLD). The candidate-gene design consists in comparing between groups classified according to their alleles on a specific gene (AA, AB, BB). The presence of a specific allele can be associated quantitatively and linearly with the trait of interest and several models can be adjusted to this end. The CG design is simple and offers relatively

straightforward interpretability of results; the gene of interest is generally well-known and its effect is often well-characterised. Despite these strengths, CG studies have been subject to important criticism. The significant findings of association regarding candidate genes approach have not been replicated when followed up in subsequent association studies. Additionally, in most cases, the biological mechanism of these candidate genes is poorly understood and only few have direct effect on the trait of interest. Indeed, most genes have indirect effects by modulating others mainly at the level of transcription, (post)translation in such proteins can have altered function in the presence of absence of others.

1.4.5 Genome-wide association study

Genome-wide association studies (GWAS) aim to identify associations of genotypes with (intermediate) phenotypes. GWAS can consider different type of genetic information; copy-number variants (CNV), sequence variations, or single nucleotide polymorphisms (SNP). The predominant strategy for identifying genomic loci associated with complex traits is through a massive *univariate-imaging univariate-genetic association study*. Univariate approaches involve the association between a single phenotype and each of millions isolated genetic variants. According to the nature of the endophenotype of interest -binary or continuous value-, various models can be adjusted. An example with endophenotypes extracted using fMRI data can be appreciated in the Fig. 1.9.

This state-of-the-art massive univariate approach results in a total of more than one million statistical tests. To assess significance, a consensus threshold $p = 5e - 8$ is considered by the genomic community. It is comparable to a Bonferroni correction of the traditional 0.05 type 1 error rate for one million statistical tests. Furthermore, a massive univariate approach ignores the biological information shared across multiple traits -the phenomenon called pleiotropy (See section 1.4.3)-. When several (endo)phenotypes are of interest, the number of tests significantly increase and it forces the users to adopt a stringent multiple comparison correction of the p-values to control the family-wise error rate. The colossal correction for multiple tests often leaves no significant associations. This limitation has led to the emergence of new multivariate statistical methods, designed to take into account the correlations among the biologically related (endo)phenotypes, classically encountered in imaging-genetic studies.

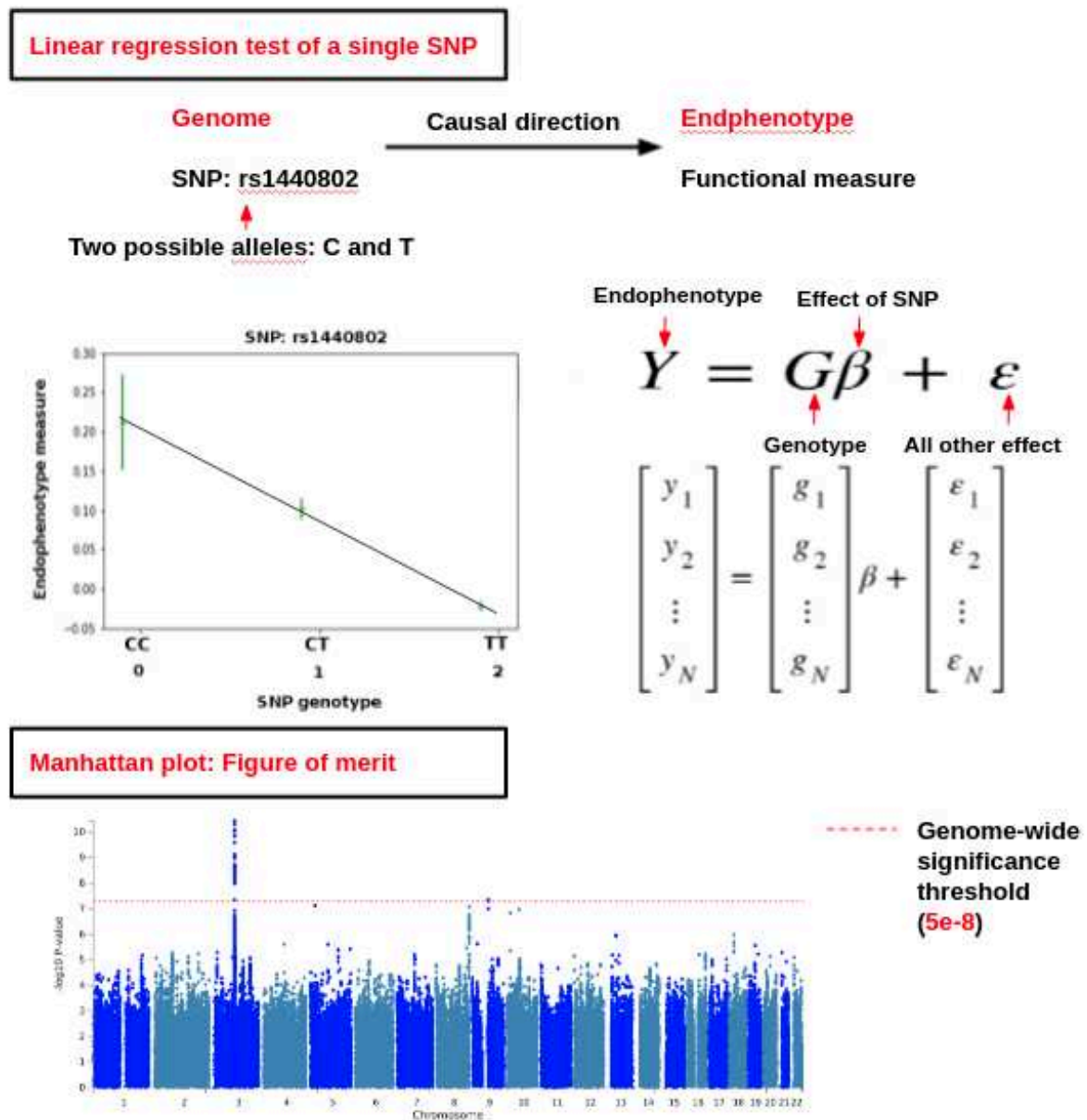


Figure 1.9 – Principles of GWA illustrated with an example of a continuous endophenotype: a functional measure of the brain. On the top figure, the y-axis is the endophenotype measure, and the x-axis represents the targeted genetic variant. This variant has two possible alleles: C and T. Each individual in the cohort will have a combination of CC, CT, TT alleles coded as 0, 1 or 2. The lower figure represent a Manhattan plot and is a figure-of-merit of GWASes. The y-axis is the $-\log_{10}$ of the p-value of the association, and the x-axis represents the SNP location along the genome. The most significant associations will have the smallest p-values and the $-\log_{10}$ of these p-values will have the highest height in the graph. The red dashed line corresponds to the $5e - 8$ genomic threshold. The points above correspond to the genetic variants significantly associated with the studied phenotype.

Multivariate approaches quantify the simultaneous contributions of loci to multiple traits by taking advantage of their genotypic-phenotypic structures. Importantly, in the case of multivariate-imaging univariate-genetic association frameworks, several advantages are observable. First, by taking into account the cross-trait covariance, multivariate analyses increase the statistical power

in case of the presence of genetic correlation between the different traits [Allison et al., 1998, Zhu and Zhang, 2009]. Second, by performing a single test for association with a set of traits, multivariate procedures alleviate the multiple comparison correction [Klei et al., 2008, Zhu and Zhang, 2009]. Third, it improves the ability to detect susceptible genetic variants whose effects are too small to be detected in univariate analysis. Fourth, contrary to the massive univariate approach, multivariate GWASes have the advantage of leveraging the distributed nature of genetic effects and the presence of pleiotropy across modalities and is thus, more consistent with biology [Chavali et al., 2010]. More details about multivariate methods that exists in the litterature are presented in the chapter 2.6

1.5 What are the bases of human language? A review of language approaches in human neurosciences

Language is complex and multi-faceted. Its study is conducted within many different disciplines and from different theoretical angles such as descriptive linguistics, theoretical linguistics, sociolinguistics, computational linguistics, historical linguistics, and neurolinguistics. The present thesis, following on the neurolinguistics approach, addresses the historical and relevant aspects of the language-brain relationship. Historically, neurolinguistics is rooted in the development of aphasiology, the study of linguistic deficits occurring as the result of brain damage. With the emergence of new brain imaging technologies, our understanding of the anatomical organisation of linguistic functions has greatly matured. In the following sections, we will attempt to produce a basic summary of our current knowledge about the language-brain relationship. These ones do not claim to be exhaustive and are only intended to provide the reader with the necessary background to understand the aspects that we addressed and studied in the rest of the present manuscript.

1.5.1 Understanding language from aphasia

Most of the early knowledge about the language-brain relationship comes from the examination of individuals with language impairments following brain injury caused by accidents or diseases. Several types of aphasia exist suggesting that language processes undergo several stages and take place in different brain regions. Based on these aphasiology discoveries, Broca and Wernicke (and Lichtheim and many others) offered fundamental knowledge about language-brain relationship. Accordingly, a language neurobiology model have been proposed; the well-known "Broca-Wernicke-Lichtheim-Geschwind model". This classical model, yet obsolete, will be furthermore, briefly described.

1.5.1.1 The Broca-Wernicke-Lichtheim-Geschwind classical model

During the 19th century, language-brain researches were mainly *post mortem* driven. Based on aphasiology, neurologists like Broca and Wernicke offered fundamental knowledge about language-brain relationship. Historically, Broca first described the association between motor aphasia and a lesion in the middle part of 'tan' patient's left frontal lobe (Broca's aphasia)

shortly followed by Wernicke who reported a damage in the left posterior superior temporal gyrus, resulting in deficits in language comprehension (Wernicke's aphasia). He proposed that a disconnection between these both language brain areas lead to aphasia. Afterward, Ludwig Lichtheim and Norman Geschwind expanded his work and proposed the "Broca-Wernicke-Lichtheim-Geschwind" model, widely known as the "Wernicke-Geschwind" classical model. Two main functions are modelised through this model, comprehension and responding to the written words as well as comprehension and responding to spoken language. According to this model, the sounds of the words are transferred through the auditory pathways to the primary auditory cortex and then to Wernicke's area, where the meaning of the words is extracted. In order for a person to produce speech, from Wernicke's area to Broca's area via the arcuate fasciculus white matter tract, the meanings of words are sent. Next, morphemes are formed and then passed on to the motor cortex. Concerning the written word, information is transferred through the primary visual cortex to the angular gyrus and from there to the Wernicke's area. The Broca-Wernicke-Lichtheim-Geschwind model is now considered to be too simplified and to contain a certain number of errors. For example, this model is based on both Broca's and Wernicke's areas connected by the arcuate fasciculus. The localisation of these cortical region is still a matter of ambiguity. Furthermore, studies showed evidence that there is no single area in our brain dedicated to language processing such as speech comprehension but a set of cortical sites [Mesulam et al., 2015]. Due to its simplicity and usefulness for understanding the classical categorization of aphasic syndromes, this classical model remains widely studied. However, researchers believe that it should be re-evaluated [Dronkers, 2000, Nasios et al., 2019, Tremblay and Dick, 2016, Vigneau et al., 2006].

1.5.2 Understanding language by brain imaging

Until recently, language-related brain processes were mainly investigated using association between language deficits and *post-mortem* analysis of brain lesions. Today, advances in structural and functional imaging techniques have provided new insights into our understanding of brain-language relationship. In many aspects, the data obtained by these imaging methods confirmed what was known from studies about language impairments. However, these also demonstrate that language processing may be even more complex than previously thought.

Modular paradigm has shaped our understanding of brain-language function. However, accumulating evidence suggest that cognition in the brain is supported by an overlapping and interactive large-scale cognitive networks rather than modules [Fuster, 2002, Mesulam, 1990]. Contrary to single-study analysis, meta analysis approaches provide empirical support for this theoretical shift and has the advantage of identifying previously neglected regions involved in language processing. Here, we will start on reporting the state-of-the-art language brain areas and large-scale cognitive networks proposed by [Vigneau et al., 2006, 2011] in terms of cortical regions. Next, we will briefly introduce the Hickok and Poeppel's dual stream model as a modern network-based model as well as the language-relevant white matter tracts.

1.5.2.1 From Broca and Wernicke to the neuromodulation era: a brief neuroanatomy of language

In order to map the language network in the human brain, we need a coordinate system that subdivides the brain into relevant parts and applies labels to these parts. In this section and throughout the manuscript, to refer to the neuroanatomical location of a given function, we use the different gyri and sulci in the cortex. The anatomical details of both left and right hemisphere are represented in the Fig. 1.10. Furthermore, an old, yet still valid, widely used parcellation provided by Korbian Brodmann in 1909 is the cytoarchitectonic description of the cortical structure [Brodmann, 1909].

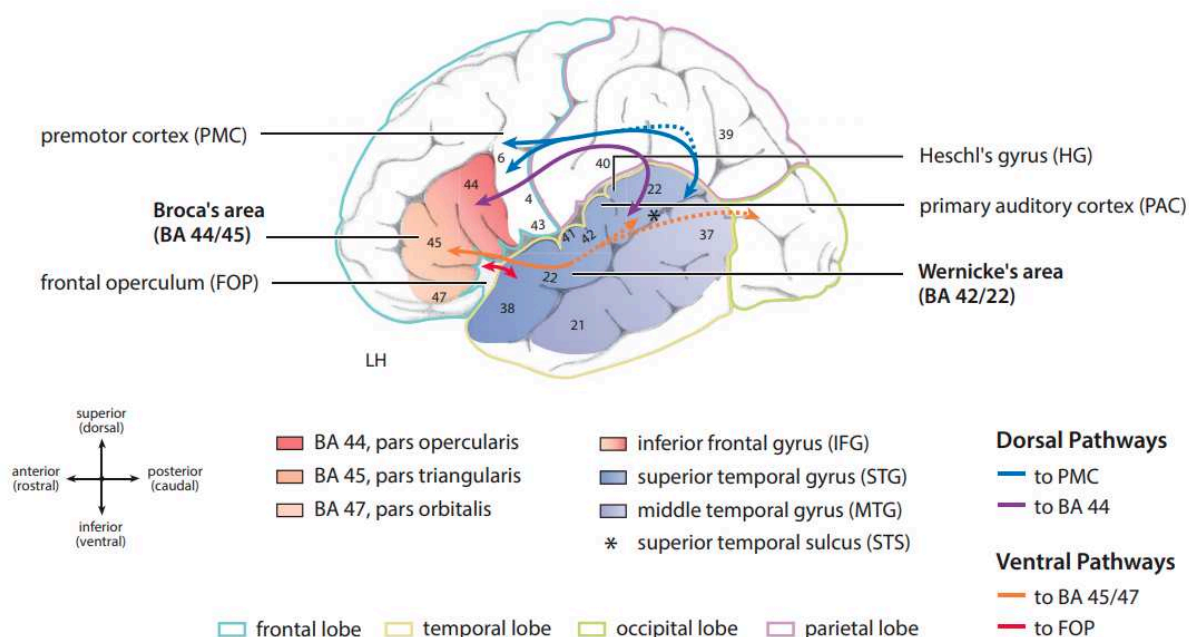


Figure 1.10 – **Neuroanatomy of language** (Source: [Friederici, 2017]). Anatomical and cytoarchitectonic details of the left hemisphere (LH). Top: The different lobes (frontal, temporal, parietal, occipital) are marked by colored borders. Major language-relevant gyri (inferior frontal gyrus (IFG), superior temporal gyrus (STG), middle temporal gyrus (MTG)) are color-coded, superior temporal sulcus (STS) located between STG and MTG is marked by an asterisk. Numbers indicate language-relevant Brodmann areas (BA) that Brodmann (1909) defined on the basis of cytoarchitectonic characteristics. The coordinate labels (see bottom left) superior (dorsal)/inferior (ventral) indicate the position of the gyrus within a lobe (e.g., superior temporal gyrus) or within a BA (e.g., superior BA 44). The horizontal coordinate labels anterior (rostral)/posterior (caudal) indicate the position within a gyrus (e.g., anterior superior temporal gyrus). Broca's area consists of the pars opercularis (BA 44) and the pars triangularis (BA 45). Located anterior to Broca's area is the pars orbitalis (BA 47). The frontal operculum (FOP) is located ventrally and medially to BA 44, BA 45. The premotor cortex (PMC) is located in BA 6. Wernicke's area is defined as BA 42 and BA 22. The primary auditory cortex (PAC) and Heschl's gyrus (HG) are located in a lateral to medial orientation in the temporal lobe. White matter fiber tracts, i.e., the dorsal and ventral pathways connecting the language-relevant brain regions are indicated by color-coded arrows (see bottom right). Reprinted from Friederici [2011]. The brain basis of language processing: From structure to function. *Physiological Reviews*, 91 (4): 1357–1392

Based on Brodmann's cytoarchitectonic parcellation, Broca's area for example, was subdivided into a posterior part (BA 44) and a anterior part (BA45) of the IFG. A subdivision that coincides with the neuroanatomical separation into pars opercularis and pars triangularis, respectively. A novel synthesis of old and new knowledge about the architecture of the language processing network is provided by the work of Vigneau et al. [2006, 2011]. In this study, the authors performed a large-scale meta-analysis of 946 brain language activation peaks collected from 129 task-based fMRI language studies from the literature (See Fig. 1.11a). This meta-analysis aimed at defining the composition of the phonological, semantic, and sentence processing networks in the fronto-parieto-temporal regions of the both cerebral hemispheres (See Fig. 1.11b). This work offers a global view of the language processing beyond the modular paradigm and accordingly, argues for large-scale architecture networks rather than modular organisation of language. The connections between different language regions distributed across the parieto-fronto-temporal cortex are ensured by the white matter bundles [Catani and Forkel, 2019, Catani et al., 2005]. These neuroanatomical white matter tracts can connect quite distant region and guarantee the information transfer between them. This short description show the complexity of the language brain relationship. More details about language-relevant fiber tracts will be given below.

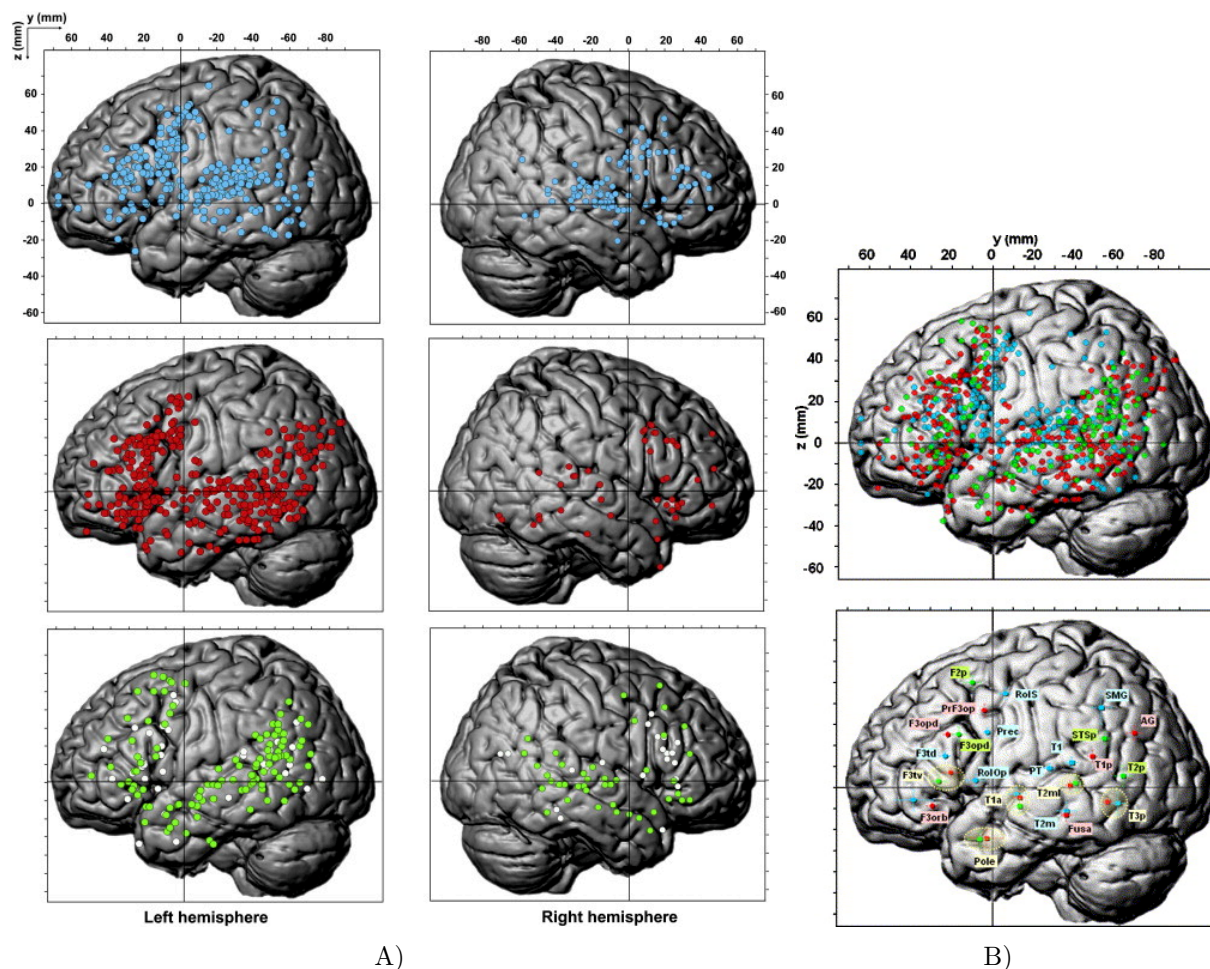


Figure 1.11 – **Overview of the meta-analysis** (Source: [Vigneau et al., 2006, 2011]). Each foci is color-coded according to its language components category: phonology (blue), semantic (red), and syntax (green). A) Distribution of the peaks in each hemisphere for each component. The 128 neuroimaging studies selected for our previous meta-analysis reported 728 activation peaks in the left hemisphere (left column). Among the 128, 59 studies reported 218 activation peaks in the right hemisphere (right column). Activation peaks detected during phonological tasks are indicated in blue (top row), lexico-semantic tasks in red (middle row) and sentence and text processing in green (bottom row). For this last component, peaks corresponding to contrast investigating comprehension are in green and those investigating syntax are in white. Peaks are projected onto a lateral view of the MNI single subject. B) Top: sagittal projection map of the 730 activation peaks. Bottom: clusters identified for phonological (blue), semantic (red), and syntactic (green) processing and obtained from the spatial clustering of the peaks. Clusters of different components that were not spatially distinct are circled (yellow). Error bars correspond to twice the standard error on the y and z stereotactic coordinates. The left hemisphere lateral surface rendering of the MNI single-subject brain template, together with the corresponding stereotactic grid, is displayed for anatomical reference. RolS, Rolandic sulcus; RolOp, Rolandic operculum; F3t, pars triangularis of the left inferior frontal gyrus; F3op, pars opercularis of the left inferior frontal gyrus; F3orb, pars orbitalis of the left inferior frontal gyrus; SMG, supramarginalis gyrus; PT, planum temporale; T1, superior temporal gyrus; T2, middle temporal gyrus; T3, inferior temporal gyrus; Prec, precentral gyrus; F2, middle frontal gyrus; PrF3op, precentral gyrus/F3op junction; STS, superior temporal sulcus; AG, angular gyrus; Fusa: anterior fusiform gyrus; a, anterior; p, posterior; l, lateral; m, middle; d, dorsal; v, ventral.

1.5.2.2 The neuroanatomical pathways of language

Despite the huge amount of knowledge accumulated the last past decades regarding the language-brain relationship, no large-scale models of the functional anatomy of language has been proposed beyond the Broca-Wernicke-Lichtheim-Geschwind classical model. To remedy this situation, Hickok and Poeppel's [Hickok and Poeppel, 2004] proposed a model, free of fiber tracts connecting the language-related brain regions but by thinking of aspects of language processing in terms of sensory-conceptual mapping of speech and sensory-motor mapping. This model consists on a dual stream model – analogous to the visual system [Hickok and Poeppel, 2004], which emphasize speech processing during perception and production in "dorsal and "ventral" pathways (streams). Each of the pathways consists of more than one fiber bundle terminating in different cortical regions. In the recent years, with the advent of new imaging techniques, such as diffusion-weighted magnetic resonance imaging, researchers addressed this question by tracking *in vivo* in human brain the white matter fiber bundles and the results supported Hickok and Poeppel's dual stream model [Hickok and Poeppel, 2007], as detailed below.

The dorsal pathway involves a left hemispheric structure, the **arcuate fasciculus (AF)**, that connects Broca's region in the posterior frontal lobe to Wernicke's region in the posterior dorsal temporal lobe as stipulated in the Broca-Wernicke-Lichtheim-Geschwind model. However, [Catani et al., 2005] tracked the arcuate fasciculus using dMRI and suggested two dorsal, direct and indirect, pathways. The direct pathway known as the **arcuate fasciculus -long segment-** connects Wernicke's region with Broca's region, and the indirect pathway is mediated by the inferior parietal cortex. It consists of the **arcuate fasciculus -anterior segment-**, linking Broca's region to Geschwind's region (encompassing the angular (BA 39) and supramarginal (BA 40) gyrus), and the **arcuate fasciculus -posterior segment-** between Geschwind's region and Wernicke's region. The **superior longitudinal fasciculus (SLF)** is also a dorsal tract composed of three parallel longitudinal branches connecting the parietal and the frontal lobes. Historically, this fasciculus has been referred to as (AF/SLF) because both the AF and the SLF run partly in parallel. The superior branch (SLF I) runs from the superior parietal lobule and the precuneus to the superior frontal gyrus. The middle branch (SLF II) connects the angular gyrus (BA39) to the posterior regions of the middle frontal gyrus (BA6). The inferior branch (SLF III) connects the supramarginal gyrus and Broca's area (the pars opercularis, the pars triangularis and the pars orbitalis) [Rojkova et al., 2016].

The ventral pathway involves bilaterally distributed hemispheric structures. It was originally viewed as a unifunctional tract but was later revisited and several ventral fiber tracts were assigned as subserving sound-to-meaning mapping, specifically the **inferior fronto-occipital fasciculus**, a long-ranged tract connecting the inferior frontal regions (BA45 et BA47) with the temporal and occipital cortex, the **uncinate fasciculus** which connects the anterior temporal lobe (BA38) to the frontal operculum, and the **inferior longitudinal fasciculus** that connects the occipital and anterior temporal cortex.

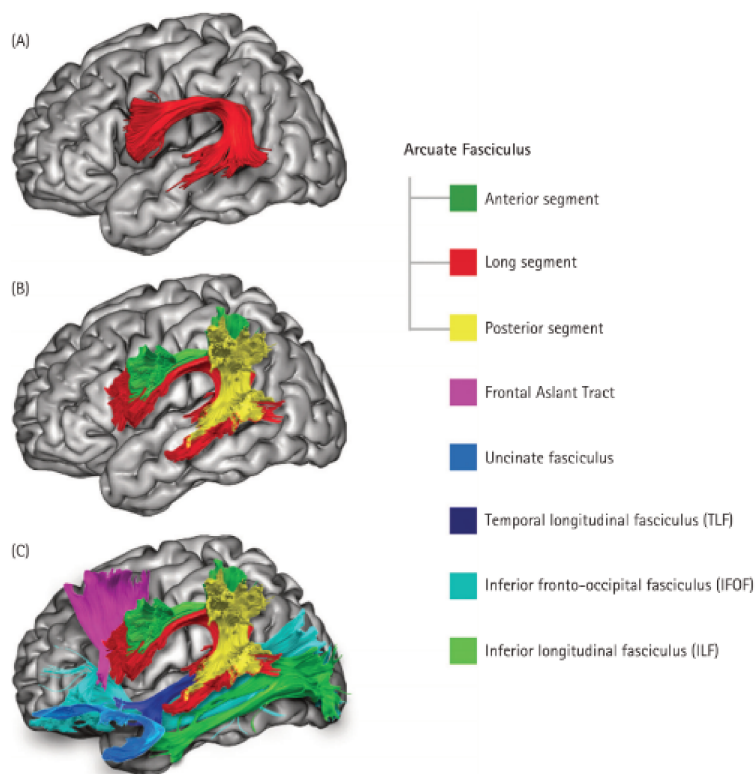


Figure 1.12 – **Tractography-based reconstruction of the language networks from classical models to contemporary neurolinguistics (Source: [Catani and Forkel, 2019]).** (A) Classical language model with the arcuate fasciculus connecting Broca’s region in the inferior frontal gyrus to Wernicke’s region in the superior temporal gyrus. (B) Extension of the classical arcuate fasciculus *sense strictu* to include the anterior segment, connecting inferior frontal to inferior parietal lobe, and the posterior segment linking the inferior parietal to the temporal lobes. (C) Current model of an extended language network beyond the three segments of the arcuate fasciculus. The frontal aslant tract (FAT) connects the inferior frontal gyrus to the pre-supplementary motor cortex. The ventral network includes the uncinete fasciculus between the anterior temporal lobe and the orbital frontal and inferior frontal cortex, the temporal longitudinal fasciculus (TLF) between the posterior temporal lobe and the temporal pole, the inferior fronto-occipital fasciculus (IFOF) connecting the ventral frontal cortex to the occipital cortex, and the inferior longitudinal fasciculus (ILF) between the occipital and anterior temporal cortex.

Additional structural networks relevant for communication are proposed [Catani and Bambini, 2014, Catani and Forkel, 2019, Rojkova et al., 2016]: *i*) the **frontal aslant tract (FAT)** is a recently described pathway that connects Broca’s region to the supplementary and pre-supplementary motor area (BA6); *ii*) the **temporal longitudinal fasciculus (TLF)** is a recently described fasciculus connecting Wernicke’s region to the anterior temporal pole; *iii*) The **fronto-insular tracts (FIT)** is a system composed of five U-shaped tracts that connects the frontal operculum to the insula. The four most anterior tracts link respectively the pars orbitalis -BA47- (FIT 1), the pars triangularis -BA45- (FIT 2), the pars opercularis -BA44- (FIT 3) and the precentral gyrus -BA6- (FIT 4) of the frontal lobe with the anterior insula. The most posterior tract (FIT 5) connects the subcentral gyrus -BA43- to the posterior insula. The Fig. 1.12 summarises these language-relevant white matter tracts.

What about the pathways in the right hemisphere and cross-hemispheric pathways?

The search of the implication of the right hemisphere in language have been, for a long time, neglected. Recent studies have demonstrate its involvement after aphasia resulting from left hemisphere but also its involvement alone or in interaction with the left hemisphere during auditory language comprehension in healthy individuals. These findings called for a structural description of the pathways within the right hemisphere and cross-hemispheric ones. Similar to the left hemisphere, the right hemisphere shows a dual-stream model with a lower density of fiber bundles compared to their left counterpart. The cross-hemispheric pathways is the corpus callosum. Its anterior and posterior parts have been discussed to be involved in language-related processes.

1.5.2.3 The neural bases of language processing networks

In [Vigneau et al., 2006], the authors identified and revisited different language processing networks. These are briefly described below.

Fronto-temporal auditory–motor network.

The motor theory of speech perception has been quite an old debate [Flinker et al., 2015, Galantucci et al., 2006, Liberman and Mattingly, 1985, Schwartz et al., 2008, Whalen, 2019]. [Vigneau et al., 2006] advocate for an audio–motor loop that includes an upper motor area for mouth motion control, a lower premotor area in the precentral gyrus that is dedicated to pharynx and tongue fine-movement coordination, and a sensory–motor integration region in the Rolandic operculum in the frontal lobe as well as the left planum temporale and Heschl’s gyrus in the temporal lobe. These fronto-temporal areas, support of a perception–action cycle, are supposed to be connected through the arcuate fasciculus fibers [Vigneau et al., 2006].

Fronto-parietal working memory loops.

The phonological working-memory loop connects the dorsal part of the pars triangularis of the inferior frontal gyrus in the frontal lobe and the supramarginal gyrus in the parietal lobe connected by both the arcuate fasciculus [Catani et al., 2005] and U-fibers [Duffau et al., 2003]. These regions constitute the neural basis of a perception–action cycle defined above for phonological working memory.

The semantics working-memory loop includes a frontal area at the junction of the precentral gyrus and opercular part of the inferior frontal gyrus and the angular gyrus in the parietal lobe connected by the arcuate fasciculus. These regions constitute the perception–action cycle for semantic working memory.

The sentence and text comprehension working-memory loop includes the posterior part of the middle frontal gyrus and the posterior part of the superior temporal sulcus connected by the arcuate fasciculus.

Fronto-parieto-temporal semantic network

Research studies described the general semantic network to typically include left parieto-frontal-temporal areas [Binder et al., 2009]. [Vigneau et al., 2006] and [Lau et al., 2008] proposed neuroanatomical model for semantic processing that includes the angular gyrus in the posterior

inferior parietal lobe, the lateral temporal lobe, including the middle temporal gyrus, the anterior temporal cortex and the fusiform gyrus, the anterior and posterior inferior frontal gyrus.

1.6 Complex and multifactorial aetiology: the pursuit of brain–language-genetics relationships

Multiple factors influence language process. Both environment and genetics plays a crucial role. Environmental factors include home literacy environment, socio-economic status, and parental education [Peterson and Pennington, 2015]. Genetic factors also affect variation as language disabilities such as dyslexia are familial. Genetic factors are not restricted to language disabilities but also affect normal variation that support language performance. The molecular genetic framework underlying language genetic component is complex and heterogenic. Many genes are likely to contribute to it. Indeed, Today, there is broad agreement that language is a complex and multifactorial trait and is unlikely to be associated to only few genes but rather to many related genes and their interactions, contributing to neural pathways that are important to normal brain maturation and possibly language development (i.e. neuronal differentiation, migration, and connectivity), together with experience-dependent contributions from the environment [Fisher and Vernes, 2015]. Here, we will briefly summarise the results found to date regarding both heritability of multiple language subcomponent and genes supporting this trait.

1.6.1 Heritability of language

A pivotal question in biology is the heritability of a particular trait. It provide a quantitative estimation of its association with a set of genetic markers and thus, informs us about the richness of its contribution in the trait of interest. How much of language is genetic and how much learned? The heritability of language has been extensively reported in the literature and the work of [Stromswold, 2001] provides a meta-analysis of the available evidence concerning a possible genetic basis for human language. These previous studies have mainly focused on twins, adoption, and linkage studies of language.

Twin studies are based on the difference between monozygotic (MZ) and dizygotic (DZ) twins. The assumption underlying these studies is that MZ twins share 100% of their genetic material, whereas DZ twins share only 50% of their genetic material. In both cases, the twins share a substantially similar environment as they were raised by the same parents in the same household. If language has a genetic component, then language abilities of MZ twins should be more similar to each other than DZ are to each others. If no difference is observed, this suggests that the heritability of language is negligible. Different linguistic ability are reviewed for both individuals with or without developmental language disorders such as dyslexia and specific language impairments. These includes spoken and non spoken vocabulary, phenoeme, phonology, articulation, auditory and phonological short-term memory, articulation, morphology, syntax, reading and spelling (syntactic comprehension and production). These studies strongly suggest that genetic factors account for some of the individual differences in linguistic ability for both people with or without developmental language disorders [Stromswold, 2001].

The adoption studies compare adopted children's linguistic abilities with those of their adopted and biological relatives. If language has a genetic component, then linguistics ability of adopted children should resemble more to their biological relatives than to their adopted ones. Database such as the Colorado Adoption Project (CAP) [Rhea et al., 2013] offer a unique opportunity to disentangle nature and nurture regarding cognitive abilities including language trait. Taken as a whole, the adoption studies also suggest that genetic factors affect language abilities and more specifically, verbal performance (See Fig. 1.13).

Overall these results advocate for a genetic predetermination and heritability of human language. Depending on the language score test, age of the volunteers, or the methodology adopted, these genetic studies of language acquisition estimate the heritability of language to be between 1 and 82% [Dale et al., 2000, Ganger et al., 2002, Reznick et al., 1997, Stromswold, 2001]. Nevertheless, these results should be taken with caution. One limitation of these studies is regarding the generalizability of their results. Indeed, most of the twin studies included in this review involved relatively small numbers of twins. Regarding adopted studies, The results are mainly obtained using data from the same group of adopted children (CAP cohort), which can or cannot be representative of the general population. Furthermore, most of these studies did not take into account corrections regarding the multiple tests performed and thus, call for replications.

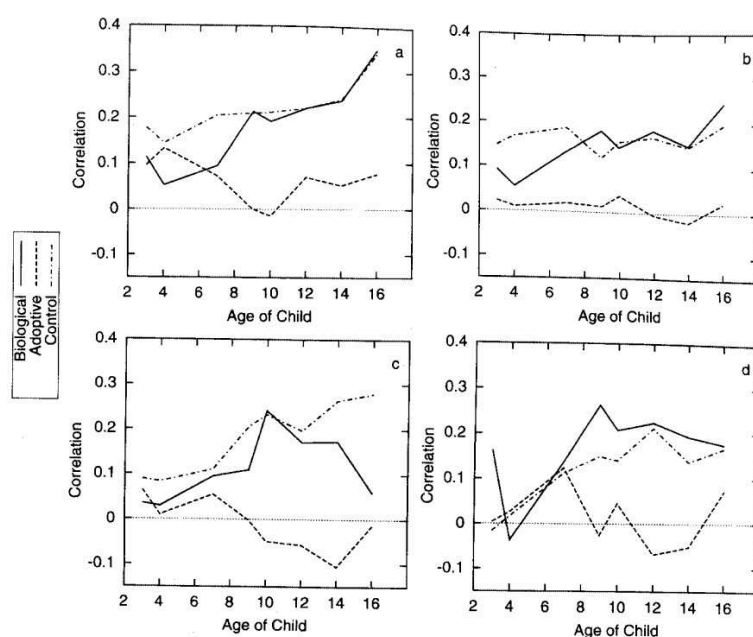


Figure 1.13 – Parent-offspring correlations for factor scores for specific cognitive abilities for adoptive, biological, and control parents and their children at 3, 4, 7, 9, 10, 12, 14, and 16 years (source: [Stromswold, 2001]). Correlations are shown for verbal ability (a), spatial ability (b), speed of processing(c), and recognition memory (d). Parent-offspring correlations are weighted averages for mothers and fathers. The number of individuals range from 33 to 44 for biological fathers, 159 to 180 for biological mothers, 153 to 197 for adoptive parents, and 136 to 217 for control parents.

To date, behavioural genetics investigations are the main type of studies that address this question. The recently available large-scale imaging-genetics cohorts provide an illuminating new tool to explore this question. Contrary to prior studies, recent studies propose to use image-derived phenotypes to explore this question by taking advantage of the UK Biobank cohort, a large-scale imaging-genetics resources constituted of unrelated volunteers introduced in the 2.1.1. The left–right hemispheric asymmetry is an important aspect of healthy brain organization for many functions including language. These studies focused on this aspect of the brain and reported the heritability of endophenotypes extracted from structural MRI [Carrion-Castillo et al., 2020, Le Guen et al., 2020, Sha et al., 2021]. Carrion-Castillo et al. [2020] evaluated the association between the planum temporale volume asymmetry with neurodevelopmental disorders including dyslexia, schizophrenia and autism and estimated its narrow-sens heritability to roughly 14%. Le Guen et al. [2020] evaluated the superior temporal assymetrical pit and estimated its narrow-sens heritability to 11%. The work of [Sha et al., 2021] offers a more general view of the left–right hemispheric asymmetry. They reported a significant narrow-sens heritability ranging from 2.2% for the assymetry index (AI) of entorhinal cortical thickness to 9.4% for the AI of superior temporal surface area. A discrepancy between heritabilities reported by both [Carrion-Castillo et al., 2020, Le Guen et al., 2020] and [Sha et al., 2021] can be explained by the difference of phenotypes studies as well as the sample size used. Indeed, while [Carrion-Castillo et al., 2020, Le Guen et al., 2020] used a sample size of 18,057 and 16,515

participants respectively, [Sha et al., 2021] take advantages of a more recent release of 32,256 participants from the UK Biobank. Regarding functional brain aspect and to our knowledge, there is no study that extensively explored the heritability of human language. That being that, the heritability of functional MRI derived phenotypes in a brain-wide context has been evaluated by [Elliott et al., 2018] and although not specific to language cognitive trait, it allow us to frame the orders of magnitude of the heritability estimates. According to Elliott et al. [2018], the resting-state fMRI functional connectivity show the lowest levels of SNP heritability, compared to structural IDP.

1.6.2 Developmental language disorder related loci

Historically, before the emergence of the genotyping technology, genetic linkage was the only affordable method able to interrogate genetics regarding trait of interest. Consequently, the candidate genes that has been targeted for reading and language impairments has arisen from linkage analysis. These findings are briefly presented in the Table 1.1.

Table 1.1 – Summary of the linkage regions to reading disability (RD) and Specific language impairment (SLI) and the candidate genes within them (Source: Hagoort [2019]).

Classification	Locus name	cytogenetic position	Candidate gene	Associated phenotypes	Reference
Dyslexia	DYX1	15q21	DYX1C1	RD	[Taipale et al., 2003]
	DYX2	6p23–21.3	DCDC2, KIAA0319	RD	[Kaplan et al., 2002], [Deffenbacher et al., 2004]
	DYX3	2p15–16	MRPL19, C2ORF3	RD	[Anthoni et al., 2007]
		2p11		RD	[Kaminen et al., 2003]
	DYX4	6q11.2–q12		RD	[Petryshen et al., 2001]
	DYX5	3p12–q13	ROBO1	RD	[Hannula-Jouppi et al., 2005]
	DYX6	18p11.2		RD	[Fisher et al., 2002]
	DYX7	11p15.5		RD	[Hsiung et al., 2004]
	DYX8	1p36	KIAA0319L	RD	[Couto et al., 2008]
DYX9	Xq27.2–q28		RD	[De Kovel et al., 2004]	
Specific language impairment (SLI)	SLI1	16q23.1–q24	CMIP, ATP2C2	Phonological memory in an SLI cohort	[Newbury et al., 2009]
	SLI2	19q13.13–q13.41		SLI	[Consortium et al., 2004]
	SLI3	13q14.3–q31.1		SLI	[Consortium et al., 2004]
				Reading-IQ discrepancy in an SLI cohort	[Bartlett et al., 2002]
	SLI4	7q31–q36	CNTNAP2	Phonological memory in an SLI cohort including nonaffected family members	[Vernes et al., 2008]
	SLI5	2q36	TM4SF20	Language delay; exon 3 deletion co-segregated with language delay in Southeast Asian families	[Wiszniewski et al., 2013]
	Other	2p22		SLI	[Bartlett et al., 2004]
	17q23		Reading-IQ discrepancy in an SLI cohort	[Bartlett et al., 2004]	
	7q31	FOXP2		[Fisher et al., 1998]	

1.6.3 Genome-wide association studies of language abilities

Genome-wide association studies of reading and language abilities have used quantitative traits from both general population and case-control designs. The results are summarised in Table 1.2.

Table 1.2 – Summary of the top genome-wide association signals from studies of reading and language traits (Source: Hagoort [2019]).

Cytogenetic position	SNP	Closest known gene	Trait	SNP test, N	Total N	Reference
1p13.1	rs4839516		Single word reading	2,373,249 imputed	6,135 individuals (children and adolescents)	[Luciano et al., 2013]
2q31	rs12474600	CCDC141	Language factor	1,724,317 imputed	2,329 children	[Harlaar et al., 2014]
3p12.3	rs7642482	ROBO2	Expressive vocabulary in 15 month olds	2,449,665 imputed/ genotyped	8,889 individuals (toddlers)	[St Pourcain et al., 2014]
4p15.3	rs11945798		Single word reading	2,373,249 imputed	6,135 individuals (children and adolescents)	[Luciano et al., 2013]
4q26	rs482700		SLI	500,527 genotyped	163 SLI (based on test performance) and 4,117 controls (children)	[Eicher et al., 2013]
5p13	rs1353835 (maternal effect)		SLI	614,937 genotyped	278 families (297 SLI children)	[Nudel et al., 2014]
5q35.1	rs9313548	FGF18	RD	125,047 autosomal SNPs	718 families (400 with RD)	[Field et al., 2013]
6p24	rs479526		Reading and spelling factors	2,373,249 imputed	6,127 individuals (children and adolescents)	[Luciano et al., 2013]
7p21	rs6954796		Phonological memory	2,373,249 imputed	6,583 individuals (children and adolescents)	[Luciano et al., 2013]
7p21	rs13307587		Reading and spelling factors; single word reading	2,373,249 imputed	6,127 individuals (children and adolescents)	[Luciano et al., 2013]
7q32	rs59197085	CCDC136	Reading and language factor	5,518,496 imputed	1,862 SLI/RD children and their nonaffected siblings	[Gialluisi et al., 2014]
8q24.1	rs2590673		RD	500,527 genotyped	353 SLI (based on test performance) and 4,117 controls (children)	[Eicher et al., 2013]
10p15	rs17135159		Reading and spelling factors	2,373,249 imputed	6,127 individuals (children and adolescents)	[Luciano et al., 2013]
10q23.3	rs1326167		Language factor	1,724,317 imputed	2,329 children	[Harlaar et al., 2014]
10q25.3	rs180950		RD	500,527 genotyped	353 SLI (based on test performance) and 4,117 controls (children)	[Eicher et al., 2013]
10q21.2	rs3213056	CDK1	Reading and spelling factor	2,373,249 imputed	6,127 individuals (children and adolescents)	[Luciano et al., 2013]
11p15.2	rs10734234	INSC	Expressive vocabulary in 15 month olds	2,449,665 imputed/ genotyped	8,889 individuals (toddlers)	[St Pourcain et al., 2014]
12q15	rs11176749	CAND1	Expressive vocabulary in 15 month olds	2,449,665 imputed/ genotyped	8,889 individuals (toddlers)	[St Pourcain et al., 2014]
13q21.3	rs1928007		Single word reading	2,373,249 imputed	6,135 individuals (children and adolescents)	[Luciano et al., 2013]
14q12	rs11158632 (paternal effect)		SLI	614,937 genotyped	278 families (297 SLI children)	[Nudel et al., 2014]
14q23	rs11158345	PRKCH	Single word reading	2,373,249 imputed	6,135 individuals (children and adolescents)	[Luciano et al., 2013]
16q22	rs764255	C16orf47	Single word reading	2,373,249 imputed	6,135 individuals (children and adolescents)	[Luciano et al., 2013]

Cytogenetic position	SNP	Closest known gene	Trait	SNP test, N	Total N	Reference
16q23	rs7187223		Single word reading	2,373,249 imputed	6,135 individuals (children and adolescents)	[Luciano et al., 2013]
19p13.3	rs1654584	DAPK3	Expressive vocabulary in 15 month olds	2,449,665 imputed/genotyped	8,889 individuals (toddlers)	[St Pourcain et al., 2014]
19p13.3	rs4807927	DAZAP1	Reading and spelling factors; single word reading	2,373,249 imputed	6,127 individuals (children and adolescents)	[Luciano et al., 2013]
21q11.2	rs2192161	ABCC13	Phonological memory	2,373,249 imputed	6,583 individuals (children and adolescents)	[Luciano et al., 2013]
22q12.3	rs5995177	RBFOX2	Reading and language factors	5,518,496 imputed	1,862 SLI/RD children and their non-affected siblings	[Gialluisi et al., 2014]

Bold entries achieved genome-wide significance. Note that a comprehensive summary of the effect allele effect size and P value is given in Carrion-Castillo et al. (2016) along with replication results for many of these SNPs. Located within the gene.

1.6.4 Moving Forward: Emergence of imaging-genetics approach

The emerging field of neuroimaging genetics seek to better understand the language-brain-genetic relationships by establishing intermediate phenotypes at the neural level. As for linkage mapping and candidate gene approaches, the predominant research were regarding language impairments such as reading disability and developmental language disorder. Going forward, and with the recent availability of large-scale cohorts such as UK Biobank 2.1.1, more and more research focus on normal population and novel candidate genes arised from large GWASes. Indeed, GWASes of normal variation have been successful in demonstrating the importance of common SNP variation (minor allele frequency > 0.01) [Davies et al., 2016]. Findings from these both approaches are detailed below.

1.6.4.1 Neuroimaging genetics studies in developmental disorders

A comprehensive summary of research employing a neuroimaging genetic approach to study neurodevelopmental disorders of spoken and written language is given in [Landi and Perdue, 2019]. In this work, the authors have limited their review to studies that includes population with specific reading disability (SRD) and developmental language disorder (DLD) and no additional neurological or psychiatric diagnosis. The table 1.3 reports these findings.

Table 1.3 – Summary of imaging genetics findings related to reading/language. (Source: Landi and Perdue [2019].)

Gene	Chr Location	Function	Design	Imaging	N	Age	Ethnic Background	Phenotype (language/reading)	References
BDNF	11p13	Neuronal survival; Neuronal proliferation; Synaptic growth	Continuous	fMRI	81	6–10	70 Caucasian 1 African-American 2 Hispanic 3 Asian 5 Mixed ethnicity	Reading	[Jasińska et al., 2016]
C2orf3/ GFCF2/ MRPL19	2p11-q11.2	Unknown	Continuous	MRI; DTI	332	3–20	European	Reading	[Eicher et al., 2016]
CMIP	16q23.2-q23.3	T-cell signaling Type IV	Case/control	MRI	76 54	6–25 5–12	98% European NR	Reading	[Scerri et al., 2012] [Skeide et al., 2016]
COL4A2	13q34	collagen subunit encoding	Case/control	MRI	54	5–12	NR	Reading	[Skeide et al., 2016]
COMT	22q11.21	Dopamine metabolism	Case/control	fMRI	86	6–10	83 Caucasian 3 African-American NR	Reading	[Landi et al., 2013]
NRSN1	6p22.3	Neurite growth	Case/control	MRI	54	5–12	NR	Reading	[Skeide et al., 2016]
ROBO1	3p12.3	Neuronal migration; Axon guidance	Case/control	MEG	10	19–51	Finnish	Reading	[Lamminmäki et al., 2012]
CCDC136/ FLNC	7q32.1	Unknown	Continuous	MRI	1275	18–35	NR	Reading/ Language	[Gialluisi et al., 2017]
DCDC2	6p22	Neuronal migration	Continuous	fMRI	82	7–12	European American	Reading/ Language	[Cope et al., 2012]
			SRD	EEG	200	8–19	German		[Czamara et al., 2011]
			Continuous	MRI DTI	76	6–25	98% European		[Darki et al., 2012]
			Continuous	MRI DTI	76	6–25	98% European		[Darki et al., 2014]
			Continuous	MRI DTI	332	3–20	European		[Eicher et al., 2016]
			Case/control	DTI	47	16–21	NR		[Marino et al., 2014]
			Continuous	MRI	56	19–85	NR		[Meda et al., 2008]
DYX1C1	15q21.3	Neuronal migration	Continuous	MRI DTI	76	6–25	98% European	Reading/ Language	[Darki et al., 2012]
			Continuous	MRI DTI	76	6–25	98% European		[Darki et al., 2014]
FOXP2	7q31	Transcriptional regulation Neurogenesis	Case/control	MRI	34	9–27	European		[Belton et al., 2003]
			Case/control	fMRI	20	19–56	European		[Liégeois et al., 2003]
			Continuous	fMRI	94	M = 24.7	European	Reading/ Language	[Pinel et al., 2012]
			Continuous	MRI	54	5–12	NR		[Skeide et al., 2016]
			Case/control	PET MRI	34	NR	European		[Vargha-Khadem et al., 1998]
			Case/control	MRI	34	9–27	European		[Watkins et al., 2002]
			Case/control	fMRI	33	M = 11	German		[Wilcke et al., 2012]
			Continuous	MRI DTI	76	6–25	98% European		[Darki et al., 2012]
KIAA0319	6p22	Neuronal migration	Continuous	MRI DTI	76	6–25	98% European	Reading/ Language	[Darki et al., 2014]
			Continuous	MRI DTI	332	3–20	European		[Eicher et al., 2016]
			Continuous	fMRI	94	M = 24.7	Primarily Caucasian	Reading/ Language	[Pinel et al., 2012]
RBFOX2	22q12.3	Alternative exon splicing regulation	Continuous	MRI	1275	18–35	NR	Reading/ Language	[Gialluisi et al., 2017]
SETBP1	18q12.3	DNA replication, apoptosis, transcription, nucleosome assembly	Continuous	fMRI	73	5–12	116 Caucasian 2 African American 3 Hispanic 4 Asian 8 mixed 2 unreported	Reading/ Language	[Perdue et al., 2019]
SLC2A3	4q32.1	Neural glucose transport regulation	SRD	EEG	200	8–19	German	Reading/ Language	[Roeske et al., 2011]
			Continuous	rsfMRI DTI	34	9–12	NR		[Skeide et al., 2015]
ACOT13/THEM2	6p22.3	Cell proliferation	Continuous	MRI; DTI	332	3–20	European	Reading/ Language	[Eicher et al., 2016]
			Continuous	fMRI	94	M = 24.7	Primarily Caucasian		[Pinel et al., 2012]
CNTNAP2	7q35	Cell adhesion Voltage-gated	Continuous	DTI	328	M = 23.4	Caucasian	Language	[Dennis et al., 2011]
			Continuous	fMRI	108	M = 26.3	Japanese		[Koeda et al., 2015]
			Continuous	MRI	54	5–12	NR		[Skeide et al., 2016]
			Continuous	MRI DTI	114	NR	NR		[Tan et al., 2010]
			Continuous	MRI	1717	M = 24.3	Primarily European Caucasian		[Uddén et al., 2017]
			Continuous	fMRI	66	M = 20.5	Caucasian		[Whalley et al., 2011]

1.6.4.2 Neuroimaging genetics studies in general population

Recent works investigate the language-brain-genetics relationships in general population and novel potential genes arising from GWASes have been proposed. These studies with a main focus on the brain asymmetry. [Carrion-Castillo et al., 2020] reported two loci (2q37 and 10p14) significantly associated with the planum temporale volumetric asymmetry including a coding polymorphism within the gene *ITIH5* that is predicted to affect the protein's function and to be deleterious and a locus that affects the expression of the genes *BOK* and *DTYMK*. Additionally, [Le Guen et al., 2020] reported a locus near *DACT1* to be associate with the superior temporal

sulcus depth. Both finding based on structural IDPs were emphasised by [Sha et al., 2021] who through a multivariate analysis, identified twenty-one loci involved in the left-right asymmetry, an important aspect of healthy brain organisation for many functions including language. These authors yield broader insights into biological pathways by proposing a role of the cytoskeleton¹ in the left–right axis determination. Although these results are particularly interesting, the relationship between brain asymmetry and language is still poorly understood. The table 1.4 summarise these findings.

¹Cytoskeleton have an essential role in many cellular processes, such as cellular division, migration, and intracellular transport.

Table 1.4 – Summary of the genome-wide association signals from studies of left-right asymmetry (source: Sha et al. [2021]).

Cytogenetic position	SNP	Functional category	Effect allele	Effect allele frequency	mvGWAS P value	Nearest gene	Trait	N SNP	Total N	Samples	Design	Reference
2p23.3	rs62130503	NcRNA intronic	T	0.05	1.2210e10	MAPRE3	Thalamus (SUB), parahippocampal (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
2p23.3	rs12617392	Intronic	A	0.44	4.0210e11	CGREF1	Inferior temporal (SA), caudal anterior cingulate (SA), isthmus of cingulate (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
2q37	rs7420166	intronic	A	0.26	7.5410e10	BOK	Planum temporale asymmetry	8,011,919	16,515		Univariate GWAS	[Carrion-Castillo et al., 2020]
3q24	rs2279829	3'UTR	T	0.22	1.2610e9	ZIC4	Isthmus of cingulate (CT), precuneus (SA), posterior cingulate (CT), fusiform (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
5q15	rs869219775	Intergenic	T	0.14	3.0610e9	NR2F1	Inferior parietal (SA), transverse temporal (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
6p21.33	rs7781	Downstream	G	0.24	1.6210e10	TUBB	Isthmus of cingulate (CT), rostral anterior cingulate (CT), pars triangularis (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
7p14.3	rs6947352	Intronic	A	0.31	4.3810e8	BBS9	Banks of the superior temporal sulcus (SA), inferior parietal (SA), isthmus of cingulate (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
9q22.33	rs911934	Intergenic	G	0.70	2.3910e15	GALNT12	precuneus (SA), paracentral (SA), supramarginal (SA), entorhinal (CT)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
10p14	rs41298373	Exonic	A	0.10	4.7510e38	ITIH5	Superior temporal (SA), parahippocampal (SA), fusiform (SA), inferior temporal (CT), transverse temporal (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
					2.0110e15		Planum temporale asymmetry	15,120,452	18,057		Univariate GWAS	[Carrion-Castillo et al., 2020]
12q13.12	rs10783306	Intergenic	C	0.33	9.9910e12	TUBA1B	Superior frontal (SA), entorhinal (SA), medial orbitofrontal (SA), pars triangularis (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
14q23.1	rs160459	Intergenic	C	0.46	4.9810e12	DACT1	Banks of the superior temporal sulcus (SA), transverse temporal (SA), pericalcarine (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
					4.2710e12		superior temporal asymmetrical pit	8,011,919	16,515		Univariate GWAS	[Le Guen et al., 2020]
16q24.3	rs72813426	Intronic	G	0.24	2.4510e14	SPIRE2	Paracentral (SA), isthmus of cingulate (SA), middle temporal (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
16q24.3	rs111398992	Intronic	T	0.13	5.9910e15	TUBB3	Isthmus of cingulate (CT), fusiform (SA), rostral anterior cingulate (CT), pericalcarine (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
17q21.31	rs55938136	NcRNA intronic	G	0.22	4.9110e15	CRHR1	Parahippocampal (SA), middle temporal (SA), pallidum (SUB), hippocampus (SUB), pars triangularis (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
17q21.31	rs35908989	Intronic	C	0.23	1.3410e8	MAPT	Supramarginal (SA), caudate (SUB)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
17q21.31	rs80103986	Intronic	T	0.20	5.1610e16	KANSL1	Parahippocampal (SA), middle temporal (SA), pallidum (SUB), hippocampus (SUB), pars triangularis (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
19p13.3	rs11672092	Intronic	T	0.22	5.6910e10	TUBB4A	Isthmus of cingulate (CT), lateral orbitofrontal (SA), middle temporal (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
20p12.1	rs6135555	Intronic	A	0.39	7.0010e9	MACROD2	Pericalcarine (SA), caudate (SUB)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
21q22.3	rs7283026	Intronic	C	0.27	8.4210e10	COL18A1	Supramarginal (SA), transverse temporal (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
22q13.31	rs9615351	Exonic	G	0.25	3.0210e8	CELSR1	Isthmus of cingulate (CT), transverse temporal (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]
Xp22.33	rs12400461	Intergenic	C	0.58	1.2410e8	ASS1P4	Inferior temporal (SA), pars opercularis (SA)	9,803,522	32,256	UKB -british sample-	Multivariate GWAS	[Sha et al., 2021]

Bold entries were replicated in an external sample

1.6.4.3 Discussion

Several approaches combined with a variety of (endo)phenotypes have been considered for the language-brain-genetic study. Through linkage analysis and after with genome-wide association studies, numerous findings were highlighted regarding language trait. These illustrate the complex neurobiological underpinnings of human language. Whether regarding targeted genetic association or exploratory genome-wide association study, replication of genetic associations with language (endo)phenotypes has emerged as an important limitation in the literature. Additional concerns regarding sample size and power, correction for multiple testing, unreported effect size have been arisen regarding neuroimaging field. Accordingly, language-brain studies should be interpreted with caution.

Notwithstanding the literature regarding genetic of human language is quite rich, we are still

quite far from understanding which part of the genome codes for language, or how genetic differences lead to different language abilities.

1.7 Aim of this thesis

The main aim of this thesis was to improve our understanding of the genetic underpinnings of human language. To this end, we investigated common genetic variants that might affect human language related quantitative trait. We determined the inter-individual variations in brain function and structure that reflect specific language as neurocognitive traits and studied the potential associations between the observed inter-individual variability of language trait and the genotype profile. This study took advantage of two well known large-scale cohorts suitable for imaging-genetic studies; the UK Biobank and the Human Connectome Project. These two large-scale cohorts comprise detailed genotyping and a wide variety of endophenotypes ranging from health/activity questionnaires, extended demographics to neuroimaging and clinical health records.

First, in chapter 3, we present an experimental design set up to highlight interesting endophenotypes representative of human language. Different imaging modalities were considered. Using resting-state fMRI, we built endophenotypes estimating rsfMRI functional connectivities between perisylvian cortical ROIs drawn from an exhaustive meta-analysis of task-based fMRI language studies. Using diffusion MRI, we determined endophenotypes from well-known language related white matter bundles. Using task-based fMRI, we extracted language-brain activation using the perisylvian cortical ROI aforementioned.

Second, in chapter 4, we provided a contribution to the study of human language heritability by presenting in-depth results on it using both psychometric and imaging derived endophenotypes as presented in the chapter 3 using the two aforementioned large-scale resources currently available. Third, in chapter 5, we performed bivariate genetic analyses to quantify the shared genetic influence between human language performance, and brain functional manifestations of language represented by both neural activation measured in language task fMRI experiments and functional connectivity estimated in resting-state fMRI. To this end, we took advantage of the Human Connectome Project cohort which comprises language-related scores, language task-based fMRI data as well as resting-state fMRI data.

Fourth, in chapter 6, we contributed to enlighten the genetic architecture of language functional connectivity by exhibiting potential key genes related to language processing. We performed a multivariate genome-wide association study on 32,186 participants from UK Biobank using significantly heritable resting-state FC endophenotypes as highlighted in the chapter 4. The results obtained were then subjected to a replication study in an independent sample ($N=4,754$). Next we performed an extensive functional annotations of each genomic risk locus which allowed us to suggest new genes with a role in different aspects of the language processing system.

Fifth, in Chapter 7, we adopted a similar experimental design than the one presented in the chapter 6 to study the genetic architecture of anatomical connectivity. We performed a mvG-WAS between genetic variants and diffusion features measured in well-known language-related

white matter bundles, on 31,775 participants from UK Biobank. The findings suggested new candidate genes involved in different brain organisation.

Finally, in the general conclusion, we have summarised and reviewed the main findings of the experimental chapters (i.e. chapter 4 to 7), and discussed future perspectives of the genetics of human language, in relation to the complementary approaches that were considered within these studies.

* * *
* *
*

Materials and methods

Chapter Outline

2.1	Cohorts	46
2.1.1	UK Biobank	46
2.1.2	Human Connectome Project	49
2.2	Definition of brain regions and tracts of interest	52
2.2.1	Functional connectomes	52
2.2.2	Anatomical connectomes	56
2.3	Rank-based inverse-normal transformation	59
2.4	Heritability.	59
2.5	Genetic correlation.	60
2.6	Massive univariate and multivariate genome-wide association study	60
2.7	Fine-mapping: identification of genomic risk loci and functional annotations. . .	63
2.7.1	FUMA: SNP2GENE	63
2.7.2	FUMA: GENE2FUNC	64
2.8	Protein networks	64

In this chapter, we present the two cohorts and the statistical methods that we used to answer the questions regarding the genes that underlie human language processing in order to see how genetic variation influences this system variation. In the first part of this chapter, we introduce the two imaging-genetics cohorts used in the manuscript. We provide as well, details on the demographics, the neuroimaging and the genetic data available in both UK Biobank (UKB) and the Human Connectome Project (HCP). The second part of this chapter concerns the definition of the regions of interest that underlie language in human brain. This step is crucial and difficult to determine as the adequate choice of ROIs may vary depending on different conditions such as the experimental design or pathologies. This choice constitutes the entry point of our study and will lay the foundations for the neuroimaging genetic studies performed. Image-derived phenotypes will be extracted using these ROIs. Details on the standard methods used in the literature to estimate these endophenotypes will be discussed further. The next sections of this chapter describe the state-of-the-art genetic methods that are today largely adopted by the imaging-genetics community. More precisely, the heritability and genetic correlation estimation, as well as genome-wide association studies (GWAS) using massive univariate methods will be described. Also, innovative multivariate approaches for GWAS will be presented. Finally, the last part is dedicated to presenting how to relate the candidate variants identified from GWAS to biological functions or processes. This step is performed through a statistical fine-mapping and proteomics interactions. It is essential in order to select and prioritise genetic variants linked to language.

2.1 Cohorts

To investigate the neurobiology of language in human brain, we took advantage of two well known large-scale cohorts; the UK Biobank and the Human Connectome Project. These open-access cohorts comprises detailed genotyping, neuroimaging data and cognitive-behavioural assessments which make them suitable for imaging-genetics studies.

2.1.1 UK Biobank

The UK Biobank is an open-access longitudinal population-wide cohort that includes 500k participants from all over the United Kingdom [Sudlow et al., 2015]. Data collection comprises detailed genotyping and a wide variety of endophenotypes ranging from health/activity questionnaires, extended demographics to neuroimaging and clinical health records. All participants provided informed consent and the study was approved by the North West Multi-Centre Research Ethics Committee (MREC).

This study used the February 2020 release (UK BioBank application number #64984). This release consisted of 41,670 participants, age range between 40 to 70 years (21,135 females, mean age=54.9 \pm 7.47 years), with genotyping and brain imaging data including resting-state functional MRI and/or diffusion MRI. Furthermore, to avoid any possible confounding effects related to ancestry, we restricted our analysis to individuals with British ancestry using the sample quality control information provided by UK Biobank [Bycroft et al., 2018]. A final

cohort of 35,366 volunteers (18,298 females, mean age = 55.1 ± 7.43 years) were included. The potential remaining cryptic population structure can be a severe confound to genetic association studies [Marchini et al., 2004]. It was accounted for by considering the first ten principal components (Data field 22009) of the principal component analysis of the genetic data and were included as covariates. We used the 5,286 non-British individuals drawn from the UK Biobank as an independent replication dataset. The age range of these participants was 40 to 70 (2,837 females, mean age = 53 ± 7.59 years). It is noteworthy that UK Biobank uses its 22 centres across the UK to collect brain imaging. Three scanning centres were represented in our sample (Cheadle, Reading, and Newcastle). This difference in scanning hardware may lead to a site-specific variability in the imaging. To overcome this limitation, we identified the imaging centre information (Data field 54) and considered the scanning site as a confound-regressor variable. Details about data acquisition parameters, data processing of multi-modalities of MRI images used in our studies as well as details about genetic data and the quality control applied will be summarised and presented below. For additional technical details, refer to the UK Biobank Brain Imaging Documentation (v.1.7, January 2020) as well as to [Alfaro-Almagro et al., 2018, Miller et al., 2016].

2.1.1.1 Data acquisition details on UKB samples

Two MRI modalities were considered throughout the manuscript as each of them potentially captures unique neurobiological details: resting state functional MRI (rsfMRI) and diffusion-weighted MRI (DW-MRI). The acquisition parameters used for each modality are described here.

Resting-state functional MRI data. Resting-state functional MRI (rsfMRI) captures low-frequency fluctuations in blood oxygenation from which functional connectivity is then estimated Biswal et al. [1995]. rsfMRI data were acquired on 3T Siemens Skyra scanner using the following parameters:

- Resolution: $2.4 \times 2.4 \times 2.4mm^3$,
- Field-of-view matrix: $88 \times 88 \times 64$,
- Repetition time (TR) = 0.735 s,
- Echo time (TE) = 39 ms,
- Acquisition time = 6 min.

As a result, 3D volumes of the volunteers' brain are acquired every 0.735s during 6 minutes resulting in 490 brain volumes in time represented by a 4D image. All images were acquired in Anterior-Posterior phase encoding direction. During the resting-state scan, participants were instructed to keep their eyes fixated on a crosshair, to relax, and to think of nothing in particular [Miller et al., 2016].

Diffusion-weighted MRI data. Diffusion-weighted imaging is a noninvasive, *in vivo* MRI technique that quantifies water diffusion in biological tissues. It is used to map structures and

measures their local organisation by tracking the movements of water molecules contained in the brain tissues. The Diffusion-weighted MRI (DW-MRI) images were acquired using a 3T Siemens Skyra scanner with the following parameters;

- Resolution: $2 \times 2 \times 2\text{mm}^3$,
- Five non diffusion-weighted image $b=0\text{ s/mm}^2$,
- Diffusion-weighting of $b=1000$, and 2000 s/mm^2 with 50 directions each,
- Acquisition time: 7 min,

2.1.1.2 Data processing details on UKB samples

All the MRI data preprocessing steps were carried out by UK Biobank brain imaging team using FSL software (<http://www.fmrib.ox.ac.uk/fsl>). The preprocessing applied to each modality is described here.

Resting-state functional MRI data. We applied the following preprocessing pipeline on rsfMRI data before we extracted the IDPs. It includes motion correction using MCFLIRT [Jenkinson et al., 2002], grand-mean intensity normalisation, high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with $\sigma = 50.0\text{ s}$) including the unwarping of EPI (echo planar imaging) and co-registration to T1 template and gradient non-linearity distortion correction (GDC), brain masking, and registration to the MNI (Montreal Neurological Institute-Hospital) space. The rsfMRI volumes were further cleaned using ICA-FIX (Independent Component Analysis followed by FMRIB’s ICA-based X-noiseifier [Beckmann and Smith, 2004, Griffanti et al., 2014, Salimi-Khorshidi et al., 2014]) for automatically identifying and removing artefacts.

Diffusion-weighted MRI data. We applied the following preprocessing pipeline on dMRI data before we extracted the IDPs. It includes: correction for eddy current distortions, head motion and has outlier-slices (individual slices in the 4D data) corrected, using Eddy (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/EDDY>) [Andersson and Sotiropoulos, 2015]. A gradient distortion correction (GDC) is applied to these corrected images, resulting in the 4D output file. The preprocessed images were then further fed into the software DTIFIT (Diffusion Tensor Imaging Fitting procedure from FSL) to model the 50 diffusion directions to generate DTI outputs, e.g. FA (fractional anisotropy), MO (tensor mode), MD (mean diffusivity). In addition to the DTI fitting, the preprocessed images were fed into NODDI (Neurite Orientation Dispersion and Density Imaging) [Zhang et al., 2012] estimates using AMICO (Accelerated Microstructure Imaging via Convex Optimization) [Daducci et al., 2015]. This enabled modeling the biological properties of fiber tracts visible in the form of ICVF (intra-cellular volume fraction - an index of white matter neurite density), ISOVF (isotropic or free water volume fraction) and OD (orientation dispersion index, a measure of within-voxel tract disorganisation). Furthermore, in order to facilitate cross-subject comparisons on fiber tract based IDPs, all the outputs were aligned onto a common space with TBSS (Tract-Based Spatial Statistics) [Smith et al., 2006].

2.1.1.3 Genetic data: genotyping and imputation

The UK Biobank genetic data contain genotypes for 488,377 participants. Genotyping was performed using the UK BiLEVE Axiom array by Affymetrix [Wain et al., 2015] on a subset of 49,950 participants (807,411 markers) and the UK Biobank Axiom array on 438,427 participants (825,927 markers). Both arrays are extremely similar and share 95% of common SNP probes. Further details are available in [Bycroft et al., 2018]. Genotype imputation is the process of statistically infer unobserved genotypes using a reference panel haplotypes. It densifies the number of SNPs that can be tested, and facilitates combining GWAS summary statistics across studies. The process of imputation is twofold: pre-phasing, and imputation. First, in the pre-phasing step, a statistical method is applied to genotype data to infer the underlying haplotypes of each individual. Second, the inferred haplotypes are combined with a reference panel (*UK10K haplotype* reference merged with *1000 Genomes Phase 3* reference panel) and the unobserved genotypes are predicted (imputed). The imputation process is extensively described here https://biobank.ctsu.ox.ac.uk/crystal/ukb/docs/impute_ukbv1.pdf. Imputation was carried out with the IMPUTE4 program (<https://jMarchini.org/software/>). The result of the imputation process is a dataset with 93,095,623 autosomal SNPs, short indels and large structural variants in 487,442 individuals. The imputed genotypes, obtained from the UK Biobank repository, were provided in BGEN file format (<https://enkre.net/cgi-bin/code/bgen/dir?ci=trunk>) and consisted of 93,095,623 autosomal SNPs, short indels, and large structural variants. See [Bycroft et al., 2018] for more details.

Quality control protocol. Before all analyses, the genetic data underwent a stringent quality control protocol, excluding participants with unusual heterozygosity, high missingness (Data field 22027), sex mismatches, such as discrepancy between genetically inferred sex (Data field 22001) and self-reported sex (Data field 31). Variants with minor allele frequency (MAF) < 0.01 were filtered out from the imputed genotyping data using PLINK 1.9 [Chang et al., 2015] to retain the common variants only.

2.1.2 Human Connectome Project

The Human Connectome Project (HCP) is an open-access population-wide cohort study that includes 1,206 participants with extensive MRI, behavioural measurements, and whole genome genotyping data, released in April 2018 [Van Essen et al., 2013]. All participants are healthy young adult twins and non-twin siblings ranged in age from 22 to 37 years old and free from current psychiatric or neurological illnesses. They provided informed consent and the study was approved by the Institutional Review Board of Washington University.

This study used the HCP S1200 release. This release consisted of 1,113 volunteers, age range between 22 to 37 years (606 females, mean age = 28.8 ± 3.7 years), who underwent an MRI experience. Furthermore, to avoid any possible confounding effects related to ancestry, we restricted our analysis to individuals identified as Caucasian using the sample quality control information provided by the Human Connectome Project [Van Essen et al., 2013]. The potential remaining cryptic population structure were accounted for by considering the first four principal compon-

ents of the principal component analysis of the genotype data covering 78.75% of the explained variance (See Fig. A.1). These were included as covariates. Details about data acquisition parameters, data processing of multi-modalities of MRI images used in our studies as well as details about genetic data and the quality control applied will be summarised and presented below. For additional technical details, refer to the the Human Connectome Project S1200 Subjects Data Release Reference Manual (10 April 2018) (<https://www.humanconnectome.org/>) and to [Van Essen et al., 2013].

2.1.2.1 Data acquisition details on HCP samples

Two MR imaging modalities were considered in the manuscript as each of them potentially captures unique neurobiological details: resting state functional MRI (rsfMRI) and task-based functional MRI (tfMRI). Briefly, the acquisition parameters used for each modality are described here.

Resting-state functional MRI data.

Resting-state functional MRI (rsfMRI) data were acquired on 3T Siemens "Connectom Skyra" scanner using the following parameters:

- Resolution: $2.0 \times 2.0 \times 2.0mm^3$,
- Field-of-view matrix: $208 \times 180mm$,
- Repetition time (TR) = $720ms$,
- Echo time (TE) = $33.1ms$,
- Acquisition time = 14 : 33 min.

As a result, 3D volumes of the volunteers' brain are acquired every 720 ms during 14:33 minutes resulting in 1200 brain volumes in time represented by a 4D image. All images were collected over two scanning sessions. Each session consists of two runs that were acquired in an alternate opposite phase encoding directions (right-to-left (RL) and left-to-right (LR)), for a total of approximately 1 h of resting-state data collected for each participant. During the resting-state scan, participants were instructed to keep their eyes open with relaxed fixation on a projected bright cross-hair on a dark background and presented in a darkened room [Van Essen et al., 2013].

Task-based functional MRI data.

Task-based functional MRI images captures changes in the activity of the neuronal populations and seek to induce different neural states in the brain [Logothetis, 2008]. The task-based functional MRI (tfMRI) images were acquired on 3T Siemens "Connectom Skyra" scanner with the same EPI pulse sequence parameters as for rsfMRI data:

- Resolution: $2.0 \times 2.0 \times 2.0mm^3$,
- Field-of-view matrix: $208 \times 180mm$,

- Repetition time (TR) = 720ms,
- Echo time (TE) = 33.1ms,
- Acquisition time = 14 : 33 min.

Seven tasks of two runs each were acquired in an alternate opposite phase encoding directions (L/R, R/L): Working Memory (Duration=5:01 min, Frames per run=405), Gambling (Duration=3:12 min, Frames per run=253), Motor (Duration=3:34 min, Frames per run=284), Language (Duration=3:57 min, Frames per run=316), Social Cognition (Duration=3:27 min, Frames per run=274), Relational Processing (Duration=2:56 min, Frames per run=232), Emotion Processing (Duration=2:16 min, Frames per run=176). These tasks are described in more detail in [Barch et al., 2013]. During the task-based experience, stimuli were projected onto a computer screen behind the volunteer’s head and the screen was viewed by a mirror positioned approximately 8 cm above the volunteer’s face. Relevant to our work, we will describe the language task experiment implemented in the HCP cohort.

The language task.

The language task implemented in the Human Connectome Project cohort was developed by Binder and colleagues [Binder et al., 2011]. It consists of two runs that each interleaves 4 blocks of a story task and 4 blocks of a mathematics task. The story task presents participants with brief auditory stories adapted from Aesop’s fables, followed by a 2-alternative forced choice question about the topic of the story. The example provided in the original Binder paper (p. 1466) is "For example, after a story about an eagle that saves a man who had done him a favour, participants were asked, “Was that about revenge or reciprocity?”. The math task aurally presents to participants a series of arithmetic operations (e.g., "fourteen plus twelve"), followed by "equals" and then two choices (e.g., "twenty-nine or twenty-six"). Participants requires to complete these arithmetic operations. The math task is adaptive in order to try to maintain a similar level of difficulty across participants. The length of each block varies (average of approximately 30 seconds), but the task was designed so that both task blocks length match. Refer to [Binder et al., 2011] and to HCP manual (<https://www.humanconnectome.org/>) for more details.

2.1.2.2 Data processing details on HCP samples

All the MR data preprocessing steps were carried out by the HCP brain imaging team using the minimal preprocessing pipeline described in [Glasser et al., 2013]. Briefly, the preprocessing of resting-state functional MR images includes correction for B0 distortion, registration to the participant’s structural scan, normalisation to the 4D mean, brain masking, and non linear warping to MNI space. The volumes were denoised as well by using 24 confound time series derived from the motion estimation (3 rotations, 3 translations, their first derivatives and all 12 resulting regressors squared) followed by ICA-FIX denoising. The spatial preprocessing steps for task-based functional MRI are identical to those used for rsfMRI.

2.1.2.3 Genetic data: genotyping

The HCP genetic data contain genotypes for 1,142 of the 1206 total released participants, including 149 pairs of genetically-confirmed monozygotic twins (298 participants) and 94 pairs of genetically-confirmed dizygotic twins (188 participants). Overall, there are 457 different families in the study, as determined by genetic analysis. These data were released on the database of Genotypes and Phenotypes (dbGaP) on March 1st, 2018. Genotyping data were obtained under dbGaP application number #17771 (dbGaP Study Accession: phs001364.v1.p1). Genotyping completed on Infinium Multi-Ethnic Genotyping Array (MEGA) chip with PsychChip and ImmunoChip content (2,052,643 markers). Further details are available in [Elam et al., 2021, Van Essen et al., 2012] and <https://www.ncbi.nlm.nih.gov/>.

2.2 Definition of brain regions and tracts of interest

In the chapter 1.3, we introduced state-of-art methods that are generally used to analyse MRI data. In this work, we have chosen to address the genetics of the human language connectome using the region of interest approach. The first step of this approach consists on identifying a set of functional regions of interest that fits our needs. This step is an important, albeit difficult, choice. In this section, we provide details about the choice of the ROIs and the tools used to generate them regarding both human language functional and anatomical connectivity.

2.2.1 Functional connectomes

Different methods to estimate the functional connectomes exists in the literature (see chapter 1.3.2 for more details). During my thesis, we focused on the ROIs-based approach. This choice was motivated by two complementary reasons. First, as participants during resting-state are not engaged in any language task-experiments, the use of regions identified through task paradigm, ensures the validity of the resting-state signal masked regarding language processing. Second, this approach provides a greater interpretability regarding the different subprocesses implied by these regions. This is achieved by taking advantage of the richness of the knowledge acquired during task experiences and the derived contrasts. In this work, two sets of regions of interest were employed for defining the language network. At the beginning of this thesis, we had chosen a set of regions of interest that includes the perisylvian regions which cover both Broca and Wernicke's areas. Later, we decided to look at a more extended language network and identified a meta-analysis regrouping and analysing a significant number of scientific reports. This meta-analysis offered a more general view of the language regions, a rich interpretation of the regions' classification regarding language components (phonology, semantic and syntax) and more importantly, it offers potentially a more consensual set of ROIs within the neuroscientific community. These sets of regions will be successively defined below. To refer to the regions defined under each set of ROIs, we use the following terms: Supramodal phonology, semantics and sentence processing (SmPSS) for the latter one, and one-contrast sentences structure (OcSS) for the former one. The core steps of the used pipeline for the functional connectome estimation are common to both UK Biobank and HCP imaging data. It is estimated from each set of brain

ROIs that form brain parcellations (nodes) by measuring their pairwise interactions termed as functional connectivity (edges). It is worth noting that all regions with less than 100 voxels were excluded from the derived analyses.

One-contrast sentences structure (OcSS). OcSS was defined based on the probabilistic activation map of a single fMRI contrast [Pallier et al., 2011]. This atlas of 7 supramodal areas, estimated on a group of 20 healthy participants, is based on activation during language comprehension. These regions of interest are located in the frontal and temporal lobe and refer to the inferior frontal activations centred on the pars triangularis (BA_{45/44}) and pars orbitalis (BA_{45/47}) and along the superior temporal sulcus from its anterior part to the temporal pole junction. These ROIs are summarised in Fig. 2.1.

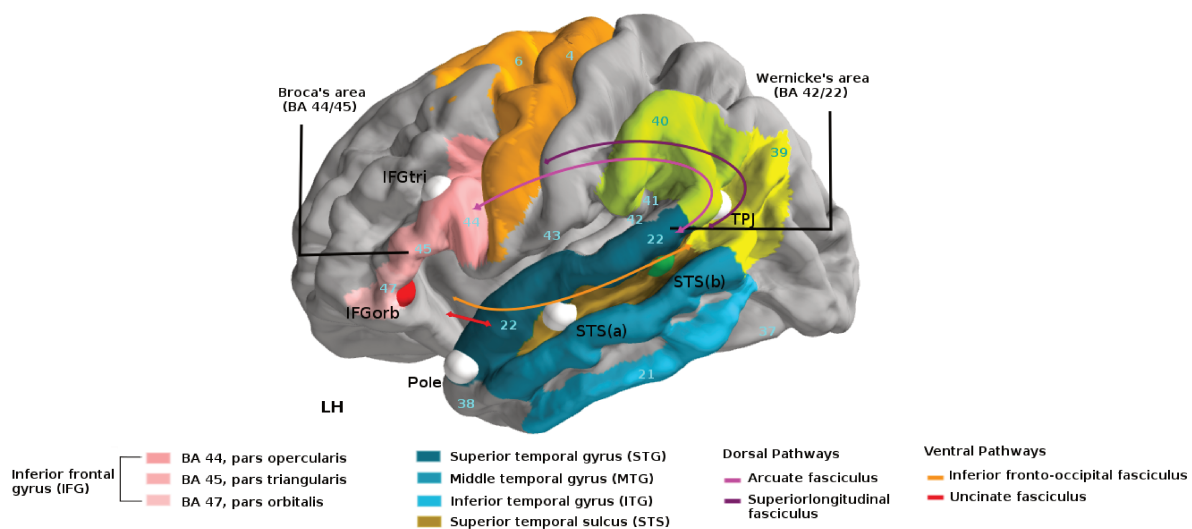


Figure 2.1 – Overview of the regions obtained from [Pallier et al., 2011]. Each language seed is color-coded according to its language category: phonology (blue), semantic (red), and syntax (green). ROIs of different components that were not spatially distinct are color-coded as pink (semantic/syntax), cyan (phonology/semantic) and white for the three language component. Different gyri and sulcus, known to be relevant for language: the inferior frontal gyrus (IFG), middle temporal gyrus (MTG), superior temporal gyrus (STG), and superior temporal sulcus (STS), are color-coded. Numbers in the left hemisphere (LH) represents language-relevant Brodmann areas (BA) which were defined on the basis of cytoarchitectonic characteristics. The pars opercularis (BA 44), the pars triangularis (BA 45) represents Broca's area. The pars orbitalis (BA 47) is located anterior to Broca's area. BA 42 and BA 22 represents Wernicke's area. Both supramarginal gyrus (BA40) and angular gyrus (BA39), also known as Geschwind's territory, are represented by green and yellow colors respectively. The primary motor cortex (BA4), the premotor cortex and the supplementary motor area (B6) are colored in orange. Within the left hemisphere, dorsal and ventral long-range fiber bundles connect language areas and are indicated by color-coded arrows.

Supramodal areas: phonology, semantics and sentence processing (SmpSS). We leverage a large-scale meta-analysis of 946 activation peaks (728 peaks in the left hemisphere, 218 peaks in the right hemisphere) obtained from a meta-analysis of 129 task-based fMRI language studies [Vigneau et al., 2006, 2011]. The identified fronto-parietal-temporal activation foci revealed via a hierarchical clustering analysis, 50 distinct, albeit partially overlapping, clusters of activation foci for phonology, semantics, and sentence processing: 30 clusters in the left hemisphere and 20 in the right hemisphere.

Because this overlap could unduly increase the co-activation between regions and to avoid a deconvolution bias in the estimation of the functional connectivity, we proceeded as follows. First,

because the clustering process was performed for each component independently, we checked whether pairs of clusters belonging to different language-component networks were spatially distinct considering the significance of their mean Euclidean distance with paired t-tests. We identified areas that are common to multiple language components; in the temporal lobe, the anterior part of the superior temporal gyrus (T1a) area appears to be common to all three language components, the anterior part of the superior temporal sulcus (Pole) and Lateral/middle part of the middle temporal gyrus (T2ml) are common to semantic and sentence’s clusters and the posterior part of the left inferior temporal gyrus (T3p) to semantic and phonology clusters. In the frontal lobe, the L|R dorsal part of the pars opercularis (F3opd) and the ventral part of the pars triangularis (F3tv) are common to semantic and syntactic clusters. In these cases, we retained the larger cluster and assigned it multiple labels. Second, ROIs were obtained for each cluster by building a 3D convex-hull of the peaks in the MNI space and were then subjected to a morphological opening operation. Third, in case of overlap between the convex-hull ROIs, the common region between the two ROIs was attributed to the most representative in terms of the number of peaks. This preprocessing resulted in 25 multilabelled ROIs: 19 in the left hemisphere and 6 in the right hemisphere which are summarised in Fig. 2.2 and Table 2.1.

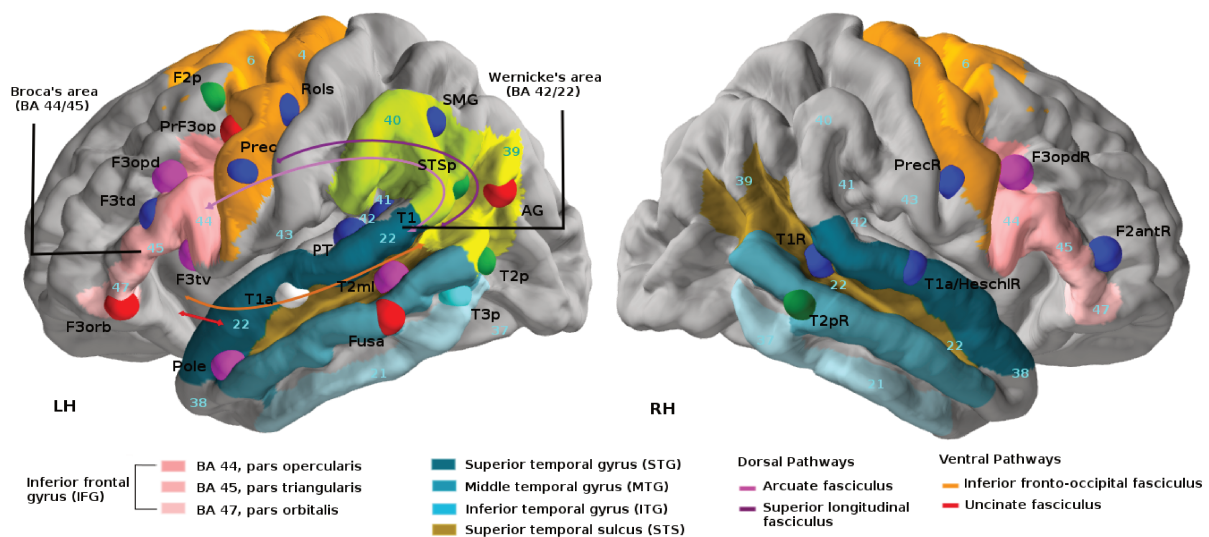


Figure 2.2 – Overview of the regions obtained from the meta-analysis. Each language seed is colour-coded according to its language category: phonology (blue), semantic (red), and syntax (green). ROIs of different components that were not spatially distinct are colour-coded as pink (semantic/syntax), cyan (phonology/semantic) and white for the three language component. For the sake of ROIs figure visibility, the coordinates were modified. The exact coordinates for each ROI are available in Table 2.2. Different gyri and sulcus, known to be relevant for language: the inferior frontal gyrus (IFG), middle temporal gyrus (MTG), superior temporal gyrus (STG), and superior temporal sulcus (STS), are color-coded. Numbers in the left hemisphere (LH) represents language-relevant Brodmann areas (BA) which were defined on the basis of cytoarchitectonic characteristics. Numbers in the right hemisphere (RH) represents the language-relevant BA counterpart. The pars opercularis (BA 44), the pars triangularis (BA 45) represents Broca’s area. The pars orbitalis (BA 47) is located anterior to Broca’s area. BA 42 and BA 22 represents Wernicke’s area [Friederici, 2011]. Both supramarginal gyrus (BA40) and angular gyrus (BA39), also known as Geschwind’s territory, are represented by green and yellow colours respectively. The primary motor cortex (BA4), the premotor cortex and the supplementary motor area (B6) are coloured in orange. Within the left hemisphere, dorsal and ventral long-range fibre bundles connect language areas and are indicated by colour-coded arrows.

Table 2.1 – Overview of the regions obtained from the meta-analysis. Each ROI is characterised by its abbreviated anatomical label defined by Vigneau et al. [2006, 2011] and is labelled according to the language component it belongs to: phonology, semantic, and syntax.

Hemisphere	Regions of interest	Phonology	Semantic	Sentence
Left hemisphere	Planum temporale (PT)	✓		
	Rolandic sulcus (RoIS)			
	Superior temporal gyrus (T1)	✓		
	Supramarginalis gyrus (SMG)	✓		
	Precentral gyrus (Prec)			
	Dorsal part of the pars triangularis of the inferior frontal gyrus (F3td)	✓		
	Anterior fusiform gyrus (Fusa)		✓	
	Opercular part of the inferior frontal gyrus (PrF3op)		✓	
	Angular gyrus (AG)			
	Pars orbitaris of the left inferior frontal gyrus (F3orb)			
	Posterior part of the middle frontal gyrus (F2p)			✓
	Posterior ending of the superior temporal gyrus (STSp)			
	Posterior part of the middle temporal gyrus (T2p)			✓
	Temporal pole (Pole)			
	Lateral/middle part of T2 (T2ml)		✓	✓
Posterior part of the left inferior temporal gyrus (T3p)	✓	✓		
Dorsal part of the pars opercularis (F3opd)				
Ventral part of the pars triangularis (F3tv)		✓	✓	
Anterior part of T1 (T1a)	✓	✓	✓	
Right hemisphere	Anterior part of the middle frontal gyrus (F2antR)	✓		
	Precentral gyrus (PrecR)			
	Anterior part of tT1 overlapping the Heschl's gyrus (T1a/HeschlR)	✓		
	Superior temporal gyrus (T1R)	✓		
	Middle part of the STS (T2pR)			✓
Upper part of the pars opercularis of F3 (F3opdR)		✓	✓	

Table 2.2 – Centre of mass of the language processing regions of interests retained in both left and right hemispheres. Each ROI is characterised by its abbreviated anatomical label defined by [Vigneau et al., 2006, 2011] and its centre of mass MNI stereotactic coordinates (x, y, z, in mm).

Hemisphere	ROIs	x	y	z
Left hemisphere	Prec	-46.76	2.01	25.85
	RoIS	-45.28	-9.40	44.17
	F3orb	-35.82	35.30	-10.77
	T3p	-45.90	-56.48	-7.35
	F3td	-42.15	25.10	14.54
	T2p	-39.30	-65.25	2.05
	SMG	-40.68	-50.25	40.86
	T1a	-54.07	-11.85	-5.85
	F3tv	-45.20	14.24	4.23
	STSp	-47.12	-56.92	24.07
	PrF3op	-39.61	2.90	39.34
	Fusa	-35.27	-37.17	-12.89
	T1	-50.46	-38.56	14.32
	T2ml	-54.99	-39.62	-2.84
	F2p	-34.72	10.94	48.80
	F3opd	-43.90	21.97	25.70
	PT	-59.89	-28.11	8.67
	Pole	-45.00	5.19	-26.79
	AG	-43.93	-70.85	22.20
Right hemisphere	T1R	51.84	-38.51	3.40
	F3opdR	49.67	14.93	27.48
	F2antR	46.64	39.80	5.06
	T2pR	54.76	-42.11	-7.73
	T1a/HeschlR	56.54	-13.38	0.81
	PrecR	52.68	-1.49	24.15

2.2.2 Anatomical connectomes

The core steps of a pipeline for the anatomical connectivity estimation are common to both UK Biobank and HCP imaging data. Although the UK Biobank team provides two types of meaningful diffusion IDPs, these do not include all key language white matter tracts making them not the most adapted to our study. Starting from this observation, we identified one set of tracts of interest that comprises well-known language white matter bundles and used it as a standard-space mask. Details about this atlas are provided below.

Atlas of Human-language Brain Connections Spherical deconvolution tractography of frontal white matter connections was performed by [Rojkova et al., 2016], resulting in a probabilistic atlas of the interlobar, commissural, projection and short U-shaped tracts in a population with a large age range. This atlas provides probabilistic maps of the well known white matter tracts that connects the different brain areas including language brain areas. The latter are represented by the three branches of the superior longitudinal fasciculus (SLF I, II, III), the uncinate fasciculus, the three segments of the arcuate fasciculus (anterior, long, posterior), the

inferior fronto-occipital fasciculus (IFOF), the frontal U-shaped tracts, the frontal aslant tract (FAT), five U-shaped fronto-insular tracts (FITs) and the frontal corpus callosum which connect both hemispheres. These tracts are summarised in Fig.2.3

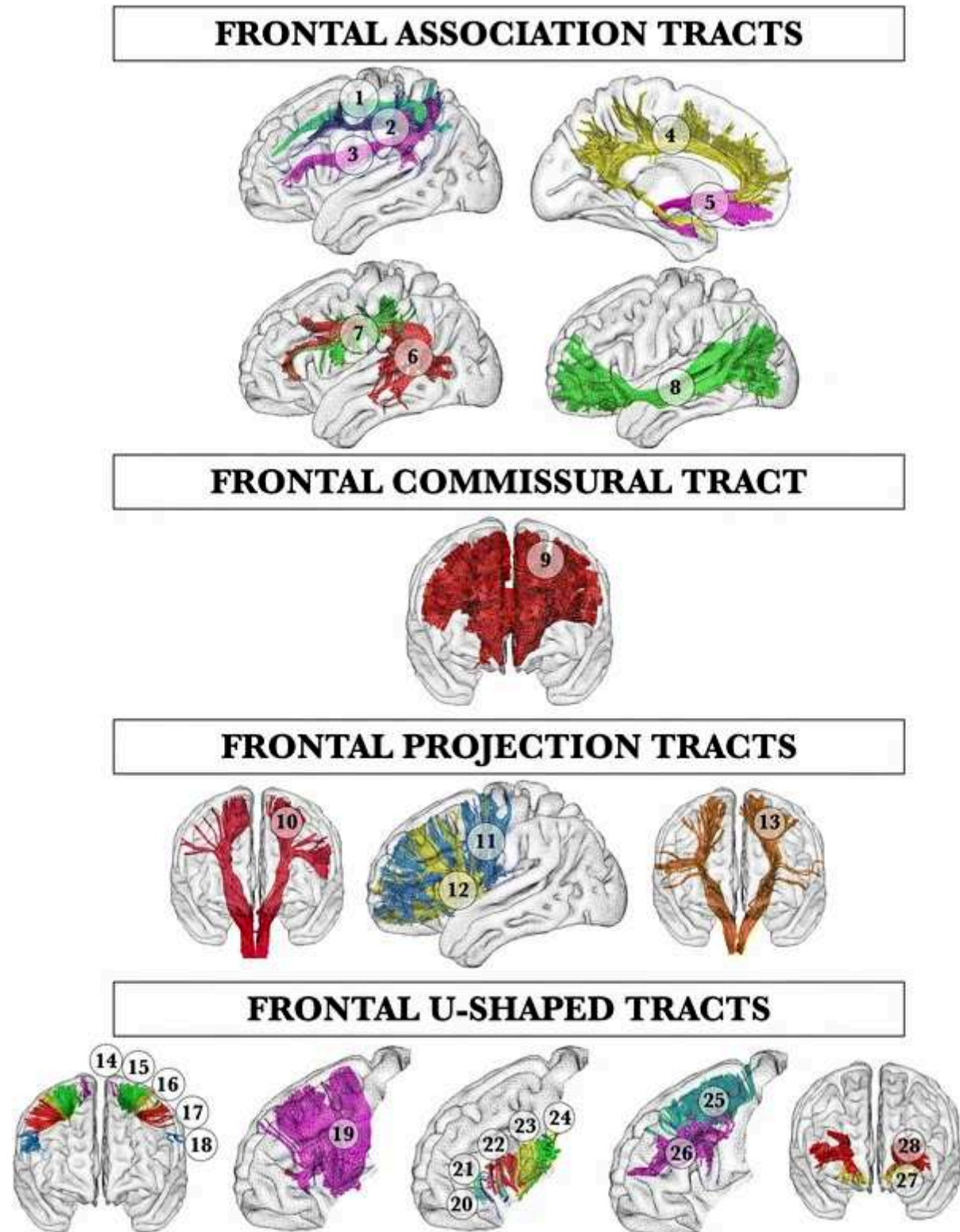


Figure 2.3 – Spherical deconvolution tractography of frontal white matter connections. (Source: [Rojkova et al., 2016]) For each tract a single participant map is supplied as a representative example of the individual anatomy. First (SLF I, light blue, no. 1), second (SLF II, dark blue, no. 2) and third (SLF III, pink, no. 3) branches of the superior longitudinal fasciculus; cingulum (yellow, no. 4); uncinate (pink, no. 5); long (LS, red, no. 6) and anterior (AS, green, no. 7) segments of the arcuate fasciculus; inferior fronto-occipital fasciculus (IFOF, no. 8); frontal corpus callosum (CC, no. 9); corticospinal tract (CST, no. 10); fronto-thalamic projections or anterior thalamic radiations (blue, no. 11); fronto-striatal projections (yellow, no. 12); fronto-pontine projections (no. 13); paracentral U tract (pink, no. 14); hand superior (green, no. 15), middle (yellow, no. 16) and inferior (red, no. 17) U tract; face U tract (blue, no. 18); frontal aslant tract (pink, no. 19); fronto-insular tract 1 (FIT 1, light blue, no. 20), 2 (FIT 2, dark blue, no. 21), 3 (FIT 3, red, no. 22), 4 (FIT 4, yellow, no. 23) and 5 (FIT 5, green, no. 24); frontal superior longitudinal (FSL, light blue, no. 25); frontal inferior longitudinal (FIL, pink, no. 26), frontal orbito-polar tract (FOP, yellow, no. 27) and fronto-marginal tract (FMT, red, no. 28). The overall visualization and screenshots were performed in Anatomist (<http://brainvisa.info>).

2.3 Rank-based inverse-normal transformation

The rank-based inverse normal transformation (INT) [Beasley et al., 2009] transforms the sample distribution of a continuous variable to make it appear more normally distributed. It is commonly used in genetic researches and applied during genome-wide association studies of non-normally distributed traits. It consists on creating a modified rank variable and then computing a new transformed value of the phenotype for the i th individual:

$$y'_i = \Phi^{-1}\left(\frac{r_i - c}{N - 2c + 1}\right), c \in \left[0, \frac{1}{2}\right] \quad (2.1)$$

where r_i is the ordinary rank of the measure for i th individual, N represents the sample size, Φ^{-1} denotes the standard normal quantile, y'_i is the value after transformation, and $c = \frac{3}{8}$ as recommended by [Blom, 1958]. Other such INTs are minor variations, with c replaced by other values. In MOSTest [van der Meer et al., 2020], the authors used $c = \frac{1}{2}$.

2.4 Heritability

Heritability is defined as the proportion of a variance for a particular phenotype, measured at a particular age, that is attributable to genetic variation [Visscher et al., 2008]. Please refer to chapter 1.4.1 for more details on its definition. This section is dedicated to the approaches and bioinformatic tools that are usually used by the genetic community for the heritability estimation. The proportion of additive genetic variance in the neuroimaging phenotypic variance, also called narrow-sense heritability, can be estimated using either the pedigree information using SOLAR (Sequential Oligogenic Linkage Analysis Routines) or the genotyped SNPs information from common genetic variant using genome-based restricted maximum likelihood (GREML) [Yang et al., 2011]. Briefly, these are detailed below.

Pedigree analysis.

The heritability estimation for imaging-genetic cohorts that include groups of related individuals is performed using extended pedigrees. The pedigree based variance components linkage method, as implemented in the SOLAR software [Almasy and Blangero, 1998], estimates the narrow-sense heritability defined as the proportion of the phenotypic variance explained by the additive genetic variance: $h^2 = \frac{\sigma_g^2}{\sigma_p^2}$. The variance of a quantitative phenotype is given by $\sigma_p^2 = \sigma_g^2 + I\sigma_e^2$, where σ_g^2 is the genetic variance due to the additive genetic factors and σ_e^2 is the variance due to individual environment.

Assuming that all environmental effects are uncorrelated among family members, the covariance matrix Ω for a pedigree of individuals is given by $\Omega = 2\Phi\sigma_g^2 + I\sigma_e^2$, where Φ is the kinship matrix representing the pairwise kinship coefficient among all individuals. It is noteworthy that it is a strong assumption that in many situation will not hold. Significance of the estimated heritability is determined by testing the null hypothesis $\sigma_g^2 = 0$ (no linkage). In order to do that, a comparison is done between the likelihood of the model in which σ_g^2 is constrained to zero with that of the model where σ_g^2 is estimated. Prior to heritability estimation, phenotype values were adjusted with respect to both *sex* and *age* covariates. In this thesis, we used SOLAR software

package [Almasy and Blangero, 1998] to estimate the heritability from the pedigree information. Outputs from SOLAR comprise the heritability value, the significant p-value and the standard error (SE) for a given phenotype.

SNP-based heritability estimation.

Contrary to the pedigree analysis, this approach computes the narrow-sense heritability based only on the genotyping data which is used to estimate the degree of relationship between unrelated individuals: the so-called genetic relationship matrix (GRM). It is noteworthy that the independence between participant is critical for the validity of the SNP-based heritability. In order to do that, the following steps are followed: first, a quantitative control consisting in filtering the SNPs is done via PLINK [Purcell et al., 2007]. The following parameters were used: missing genotype $< 5\%$, minor allele frequency $< 1\%$, Hardy-Weinberg equilibrium $< 10^{-6}$. The SNPs in moderate linkage disequilibrium (with a r^2 of 0.9) are kept. Next, the GRM is first estimated with the remaining SNPs and is then fitted by a mixed effect linear model (MLM) in order to identify the amount of variance attributed to the selected set of SNPs via the restricted maximum likelihood (REML) method. Prior to heritability estimation, phenotype values were adjusted for *sex*, *genotype array type*, *age*, *acquisition center* and the first 10 *genetic principal components* provided by UK Biobank to account for population structure. In this thesis, we used the Genome-wide complex trait analysis (GCTA) Yang et al. [2011] to estimate the heritability from the single nucleotide polymorphisms (SNPs) genotyped information. Similarly to SOLAR Almasy and Blangero [1998], outputs from GCTA comprise the heritability value, the significant p-value and the standard error (SE) for a given phenotype.

2.5 Genetic correlation

Genetic correlation is a key population parameter that describes the shared genetic architecture of complex traits and diseases [Ni et al., 2018]. It is the proportion of additive genetic covariance of pairs of endophenotypes. It can be estimated by current state-of-the-art methods, i.e., Bivariate genome-based restricted maximum likelihood (Bivariate GREML) [Lee et al., 2012] (<https://cnsgenomics.com/software/gcta/Overview>), linkage disequilibrium score regression (LDSC) [Bulik-Sullivan et al., 2015] (<https://github.com/bulik/ldsc>), or independent SNP Effect Concordance Analysis (iSECA) (<https://neurogenetics.qimrberghofer.edu.au/iSECA/>) [Nyholt, 2014]. The bivariate-GREML approach requires individual-level population-based SNPs genotype data while the two latter do not require individual-level genotype data but instead use GWAS summary statistics.

2.6 Massive univariate and multivariate genome-wide association study

As stated in chapter 1.4.5, the predominant strategy for identifying genomic loci associated with complex traits is through massive univariate genome-wide association studies (GWAS). Univariate approaches involve the association between a single phenotype and each of millions

isolated genetic variants. According to the nature of the endophenotype of interest -binary or continuous value-, various models can be adjusted. For example, for quantitative phenotypes, the following linear model is fitted:

$$y = G\beta_G + X\beta_X + e \quad (2.2)$$

where y is the phenotype vector, G is the genotype matrix for the current variant, X is the fixed-covariate matrix and contains an intercept column, and e is the error term subject to least-squares minimisation.

This state-of-the-art massive univariate approach results in a total of more than one million statistical tests. To assess significance, a consensus threshold $p = 5e - 8$ is considered by the genomic community. It is comparable to a Bonferroni correction of the traditional 0.05 type 1 error rate for one million statistical tests. However, this approach raise a certain number of limitations. A massive univariate approach ignores the biological information shared across multiple traits -the phenomenon called pleiotropy (See section 1.4.3)-. When several (endo)phenotypes are of interest, the number of tests significantly increase and it forces the users to adopt a stringent multiple comparison correction of the p-values to control the family-wise error rate. The colossal correction for multiple tests often leaves no significant associations. This limitation has led to the emergence of new multivariate statistical methods, designed to take into account the correlations among the biologically related (endo)phenotypes, classically encountered in imaging-genetic studies.

On the other hand, multivariate approaches quantify the simultaneous contributions of loci to multiple traits by taking advantage of their genotypic-phenotypic structures. Importantly, multivariate frameworks offer several advantages. These were listed in chapter 1.4.5 and we will take the liberty of listing them here as well for the sake of clarity. First, by taking into account the cross-trait covariance, multivariate analyses increase the statistical power in case of the presence of genetic correlation between the different traits [Allison et al., 1998, Zhu and Zhang, 2009]. Second, by performing a single test for association with a set of traits, multivariate procedures alleviate the multiple comparison correction [Klei et al., 2008, Zhu and Zhang, 2009]. Third, it improves the ability to detect susceptible genetic variants whose effects are too small to be detected in univariate analysis. Fourth, contrary to the massive univariate approach, multivariate GWASes have the advantage of leveraging the distributed nature of genetic effects and the presence of pleiotropy across modalities and is thus, more consistent with biology [Chavali et al., 2010].

To date, several multivariate approaches to genome-wide association study have been proposed; MQFAM implemented in the genetic association analysis software PLINK (MV-PLINK) [Ferreira and Purcell, 2009] is based on canonical correlation analysis, a Bayesian multiple phenotype test implemented in SNPTEST (MV-SNPTEST) [Marchini et al., 2007], the R package MultiPhen (MultiPhen) [O'Reilly et al., 2012] tests the linear combination of phenotypes for each SNP by performing ordinal regression, a Bayesian model comparison and model averaging for multivariate regression in BIMBAM (MV-BIMBAM) [Guan and Stephens, 2008, Stephens,

2013], the Principal Component of Heritability Association Test (PCHAT) [Klei et al., 2008], a Trait-based Association Test that uses Extended Simes procedure (TATES) [Van der Sluis et al., 2013], MultiABEL implements Pillai’s trace MANOVA [Pillai, 1955] to calculate multivariate p-value from summary statistics, adaptive multi-trait association test (aMAT) [Wu, 2020], Multi-Trait Analysis of GWAS (MTAG) [Turley et al., 2018], METAL [Willer et al., 2010], genome-wide association meta-analysis (GWAMA) [Mägi and Morris, 2010], Joint Analysis of GWAS Summary Statistics (JASS) [Julienne et al., 2020] and Multivariate Omnibus Statistical Test (MOSTest) [van der Meer et al., 2020].

Comparative studies have shown the utility of multivariate methods listed as it substantially boost statistical power and gene discovery over univariate tests [Galesloot et al., 2014, Porter and O’Reilly, 2017]. However, imaging-genetic community struggles to adopt these methods due to the lack of interpretation. Indeed, an important challenge is to identify "central" phenotypes driving the multivariate association from ones with minor contributions to the genotype-phenotypes association statistic. To date and to our knowledge, there is one bioinformatic tool; Meta-Phenotype Association Tracer (MetaPhat) [Lin et al., 2020] that address this issue by performing decomposition and tracing the traits of highest and lowest statistical importance to identify subsets of central traits at each associated variant.

In our work, we used the Multivariate Omnibus Statistical Test (MOSTest) [van der Meer et al., 2020] to conduct our multivariate analysis. Briefly, MOSTest is a twofold method. First, a univariate GWAS on each of pre-residualised and normalised (rank-based inverse-normal transformation) endophenotypes is carried out using a standard additive model of linear association between genotypes and each phenotype. Statistical significance was assessed from Pearson’s correlation coefficient. Second, the MOSTest statistic χ_j^2 , for the j th SNP is calculated as a Mahalanobis norm $\chi_j^2 = \mathbf{z}_j^T \tilde{\mathbf{R}}^{-1} \mathbf{z}_j$, where $\mathbf{z}_j = (z_{1j}, \dots, z_{Kj})$ is the vector of z -scores of j th SNP across K phenotypes and $\tilde{\mathbf{R}}$ is the K -by- K correlation matrix of $\tilde{\mathbf{Z}}$; matrix of z -scores, calculated from association tests on a randomly permuted genotype vector of each SNP.

In this thesis, we performed a multivariate genome-wide association study (mvGWAS) between the filtered imputed genotypes described in the genetic quality control section 2.1.1.3 and the significantly heritable endophenotypes (see chapter 4), using MOSTest [van der Meer et al., 2020] described above. All endophenotypes were pre-residualised controlling for covariates including sex, genotype array type, age, recruitment site, and first ten genetic principal components provided by UK Biobank. In addition, as mentioned above, MOSTest performs a rank-based inverse-normal transformation (see section 2.3) of the residualised endophenotypes to ensure that the inputs are normally distributed. The distributions across the participants of all endophenotypes were visually inspected before and after covariate adjustment. MOSTest generated summary statistics that capture the significance of the association across language endophenotypes. To account for multiple testing over the whole genome, statistically significant SNPs were considered as those reaching the genome-wide threshold $p = 5e-8$.

2.7 Fine-mapping: identification of genomic risk loci and functional annotations

To further process the output from the multivariate genome-wide association study, the FUMA online platform v1.3.6a [Watanabe et al., 2017] with default parameters was used. FUMA aims to use functional, biological information to prioritise genes based on GWAS outcomes. It is organised as two separate processes; SNP2GENE and GENE2FUNC. Details of these two processes are presented below.

2.7.1 FUMA: SNP2GENE

SNP2GENE consists in a multi-step process which characterises genomic loci, annotates candidate SNPs in genomic loci and performs a SNP-to-gene mapping using three different strategies. It also perform gene-based and gene-set genome-wide association analyses using MAGMA [de Leeuw et al., 2015]

Characterization of significant hits.

The SNP2GENE process allows us to identify distinct, significantly associated genomic loci, and independent lead SNPs within those loci.

The characterization of significant hits process allows us to identify distinct, significantly associated genomic loci, and independent lead SNPs within those loci. *UKB release2b 10k white British* is used as a reference panel to calculate LD across SNPs and genes. SNPs with genome-wide significant mvGWAS p-values $p < 5e-8$ and independent from each other at $r^2 < 0.6$ were identified. For each of these SNPs, all known SNPs, regardless of being available in the input, that had $r^2 \geq 0.6$ with them were included for further annotation (candidate SNPs). Based on the identified independent significant SNPs, independent lead SNPs were also defined among them as having low LD ($r^2 < 0.1$) with any others. Additionally, each genomic locus can contain multiple independent significant SNPs and lead SNPs. Indeed, if LD blocks of independent significant SNPs are closely located to each other ($< 250kb$, default parameter), they are merged into one genomic locus.

SNPs were annotated for functional consequences on gene functions using ANNOVAR [Wang et al., 2010], Combined Annotation Dependent Depletion (CADD) scores [Kircher et al., 2014], and 15-core chromatin state prediction by ChromHMM [Ernst and Kellis, 2012] by matching chromosome, position, reference, and alternative alleles.

Additionally, RegulomeDB v2.0 [Boyle et al., 2012] was queried externally to annotate SNP. Coding hit SNPs are also annotated with polymorphism phenotyping v2 (Polyphen-2) [Ramensky et al., 2002].

SNPs were mapped to genes at significant mvGWAS loci using three different strategies; positional mapping, expression quantitative trait loci (eQTL) mapping, and chromatin interaction mapping.

Gene-level analyses.

FUMA implements different methods that enable to summarise SNP associations at the gene level and to associate the set of genes to biological pathways.

Gene-based association analysis was performed using mvGWAS summary statistics as input into MAGMA [de Leeuw et al., 2015]. This process examines the joint association signals of all SNPs within a given gene (including 50 kb upstream to 50 kb downstream of the gene), while considering the LD between the SNPs. FUMA applies a Bonferroni correction for the number of tested genes to identify significant gene associations.

Gene-set enrichment analysis was performed using MAGMA [de Leeuw et al., 2015], to test for enrichment of association of predefined gene sets. This process is performed for curated gene sets and the Gene Ontology (GO) terms [gen, 2021, Ashburner et al., 2000] obtained from the Molecular Signatures Database (MsigDB) [Subramanian et al., 2005]. A Bonferroni correction is applied considering all tested gene sets.

Gene-property analysis was performed using MAGMA [de Leeuw et al., 2015] to test if genes are over-represented in developmental stages and considering gene expression from BrainSpan data source from 11 developmental stages (from early prenatal to middle adulthood) and 29 different age groups (from 8 post-conceptual weeks to 40 years old). An FDR correction was used to correct for multiple testing.

2.7.2 FUMA: GENE2FUNC

GENE2FUNC provides information on expression of prioritised genes and tests for enrichment of the set of genes in pre-defined pathways. It offers four main views regarding gene expression per tissue types (GTEx) or developmental stage (BrainSpan), enrichment of differentially expressed gene (DEG) sets in a certain tissue compared to all other tissue types, enrichment of input genes in gene sets using hypergeometric tests, and gene table with input genes mapped to OMIM ID, UniProt ID, Drug ID of DrugBank and linked to GeneCards.

2.8 Protein networks

Interactions between proteins are important as they play a crucial role in cellular functions and biological processes. Multiple databases track the protein-protein interaction network using both physical interaction and functional associations. For example, the STRING database [Szklarczyk et al., 2017] gathers interaction experiments where physical interactions are identified and take into account experimental evidence that includes:

- Conserved neighbourhood: genes that occur repeatedly in close neighbourhood in (prokaryotic) genomes
- Co-occurrence: the presence or absence of linked proteins across species.
- Fusion: individual gene fusion events per species.
- Co-expression: identify genes that show similar expression pattern across many different conditions.

- Experiments: significant protein interaction datasets, gathered from other protein-protein interaction databases.
- Databases: curated knowledge from manually annotated databases of protein complexes as well as molecular pathways.
- Automatic text mining: process the vast biomedical literature and extract from the abstracts of scientific literature.

In this work, we used curated experimental protein-protein interaction (PPI) observations from the STRING [Szklarczyk et al., 2017] database (<https://string-db.org/>, version 11.5), and sought for a functional enrichment through this analysis.

* * *
* *
*

Human language endophenotypes

Chapter Outline

3.1	Psychometric measures	68
3.2	Language neural activation: a volume-based analysis	71
3.3	Resting-state functional connectome	73
3.4	Anatomical connectome	75
3.5	Conclusion	76

In chapters 1 and 2, we presented basic principles of imaging and genetics useful for understanding imaging-genetic studies as well as the data and statistical methods explored. Large-scale neuroimaging studies usually summarise the imaging data into image-derived phenotypes (IDPs), which reduces raw data into a compact set of biologically meaningful measures. Compared to the traditional voxel level approach, the IDPs approach is simple, and effectively reduces the high-dimensional data into interpretable, compact, convenient features. However, IDPs remain indirect measures of the biological process of interest and meaningful interpretation requires care. Here, we provide the standard pipelines used to extract novel endophenotypes covering both psychometric and image derived phenotypes that encompass the human language connectome. To this end, the two well-known large-scale UK Biobank and HCP cohorts were used. These endophenotypes will be subjects to heritability analysis and genetic association studies as described further in the chapters 4 to 7. Genetic association analysis with complex cognitive traits can be enriched through the use of different information on trait development.

3.1 Psychometric measures

Two types of behavioural scores are available in the HCP cohort. Cognitive performance scores measured during the MRI session and psychometric measures acquired using standardized tests from the National Institute of Health (NIH), independently from the task performed inside the MRI scanner. These latter ones includes fluid intelligence, working memory, and language assessments such as vocabulary comprehension and oral reading decoding. A brief description taken as it is from (<https://nihtoolbox.force.com/s/article/nih-toolbox-scoring-and-interpretation-guide>) is presented below. The NIH measure consists of:

- Oral reading recognition: measured by the oral reading recognition Test. The participants are asked to read and pronounce letters and words as accurately as possible. Higher scores indicate better reading ability.
- Vocabulary Comprehension: measured by the Picture Vocabulary Test. The participants are presented with an audio recording of a word and four photographic images on the computer screen and is asked to select the picture that most closely matches the meaning of the word. Higher scores indicate higher vocabulary ability.
- Fluid Intelligence: measured by the Penn Matrix Test (PMAT) which measures fluid intelligence via non-verbal reasoning. Participants are presented with patterns made up of 2x2, 3x3 or 1x5 arrangements of squares, with one of the squares missing. They must pick one of five response choices that best fits the missing square on the pattern. Two features are declined from this test: the number of correct answers and the median reaction time (RT) for correct answers.
- Working Memory: measured by the List Sorting Working Memory Test. The participants sequence different visually and orally presented stimuli. Pictures of different foods and animals are displayed with both a sound clip and written text that name the item. The task has two different conditions: 1-List and 2-List. In the 1-List condition, participants

are required to order a series of objects (either food or animals) in size order from smallest to largest. In the 2-List condition, participants are presented both food and animals and are asked to report the food in size order, followed by the animals in size order.

The distribution of these psychometric measures can be appreciated in Fig.3.1

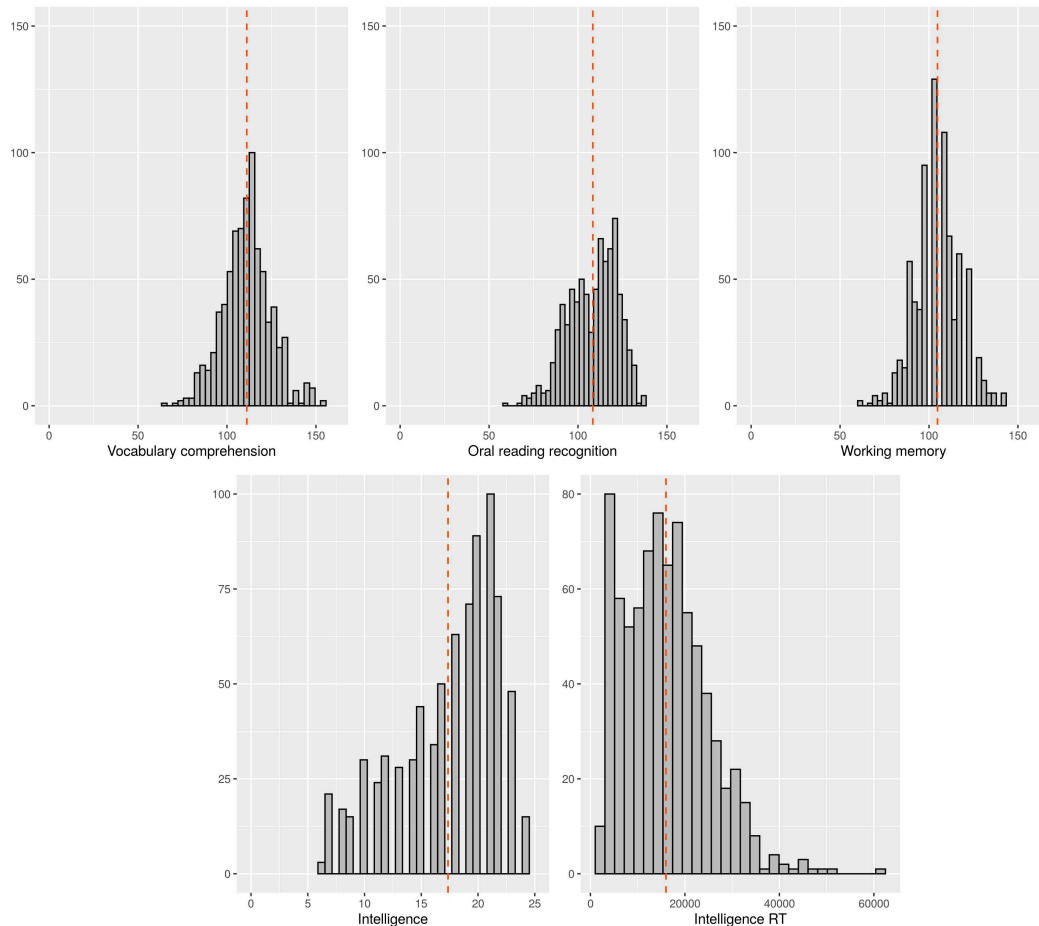


Figure 3.1 – Distribution of the NIH measures collected by HCP team independently from the language fMRI task

The cognitive performance scores measured during the MRI session consist of:

- Language task math: three features are measured, namely the accuracy percentage during math condition in language task, the median reaction time for correct trials and the average difficulty level of the presented stimuli.
- Language task overall: three features are measured, namely the average of accuracy from each condition in language task, the average of median reaction time from each condition in language task.
- Language task story: three features are measured, namely the accuracy percentage during story condition in language task, the median reaction time for correct trials and the average difficulty level of the presented stimuli.

The distribution of these psychometric measures can be appreciated in Fig.3.2

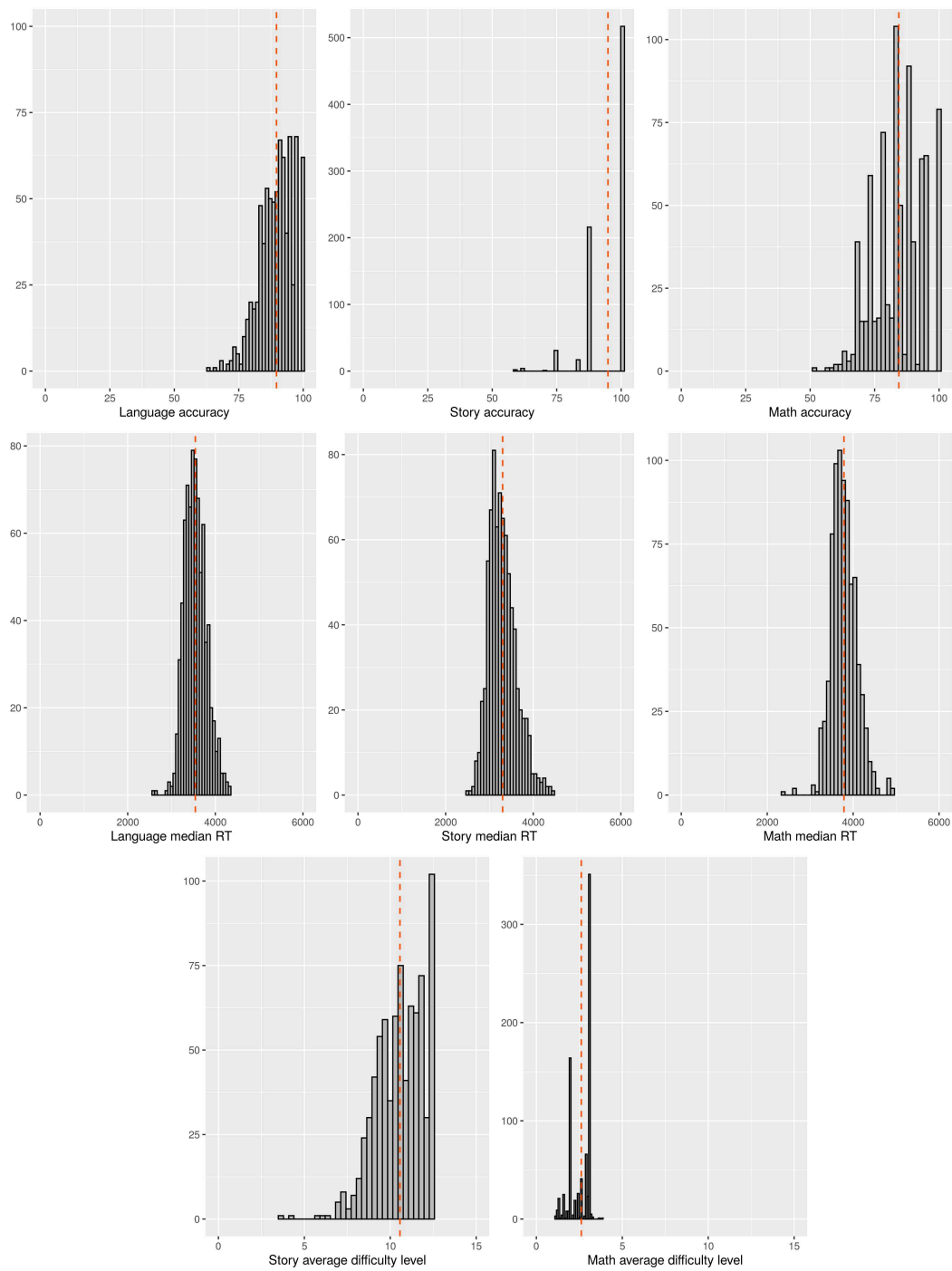


Figure 3.2 – Distribution of behavioral scores collected by HCP team during the language fMRI task, composed of MATH and STORY tasks

3.2 Language neural activation: a volume-based analysis

As described in the chapter 1.3.2, task-based fMRI (tfMRI) measures BOLD signal changes between task-stimulated states such as visual, auditory, or language and control states (resting state). These data are presented as a succession of 3D images that vary in time (4D image) and which reflect the state of the brain at a given moment regarding the cerebral stimulation of the experiment set up. The analysis of task-based fMRI data consist of participant-level and group-level analyses. The participant-level known as the first level consists of fitting a general linear model (GLM) using the task-based fMRI data and performing a t-test or an F-test according to the condition one wants to highlight; either a condition versus baseline or a combination of conditions (possibly two-, three- or higher-dimensional) that explains a significant proportion of the signal. The group-level, also known as the second level consists on averaging the activation across participants. Both first-level and second-level analysis are described in the chapter 1.3.2. In this section, we will describe the methodology applied to analyse the language task-based fMRI data in the HCP cohort and the pipeline used to extract the endophenotypes related to language.

Concerning the first level, activity estimates were computed for the preprocessed functional time series (4D image) using a GLM implemented in Nilearn. Two predictors of interest were included in the GLM model design: the story predictor covered the variable duration of a short story, question, and response period ($\sim 30s$), and the math predictor covered the duration of a set of math questions designed to roughly match the duration of the story blocks. These predictors were convolved with a double gamma “canonical” hemodynamic response function (HRF) to generate the main model regressors. To compensate for slice-timing differences and variability in the HRF delay across regions, temporal derivative terms derived from each predictor were added and were treated as confounds of no interest. A cosine drift model was used. To correct for autocorrelations in the fMRI data, an autoregressive AR₁ noise model was used. Finally, we applied to the task-volume, a spatial smoothing using an unconstrained 3D Gaussian kernel of full width at half maximum ($FWHM$) = 4mm. The Fig.3.3 represents the design matrix used to fit the GLM model. At the end, we obtain a β -map for each participants. To access the estimated β coefficients of the GLM model, we used the contrast represented in the Fig.3.3 to perform a story versus baseline contrast. The role of the contrast is to select columns of the model –and potentially weight them– to study the associated statistics. In order to get statistical significance, we form a t-statistic, and convert it into z-scale. The values are scaled across voxels to match a standard Gaussian distribution (mean=0, variance=1). At the end, we obtain a z-map for each individuals.

Regarding the second level, activity estimates were computed for the individual β -maps using a GLM implemented in Nilearn followed by a t contrast. The Fig3.4 shows the activations for the story versus baseline (The intercept of the general linear model being considered as baseline). The intercept reflects the mean of the residual BOLD time series after removing variance explained by all other regressors.

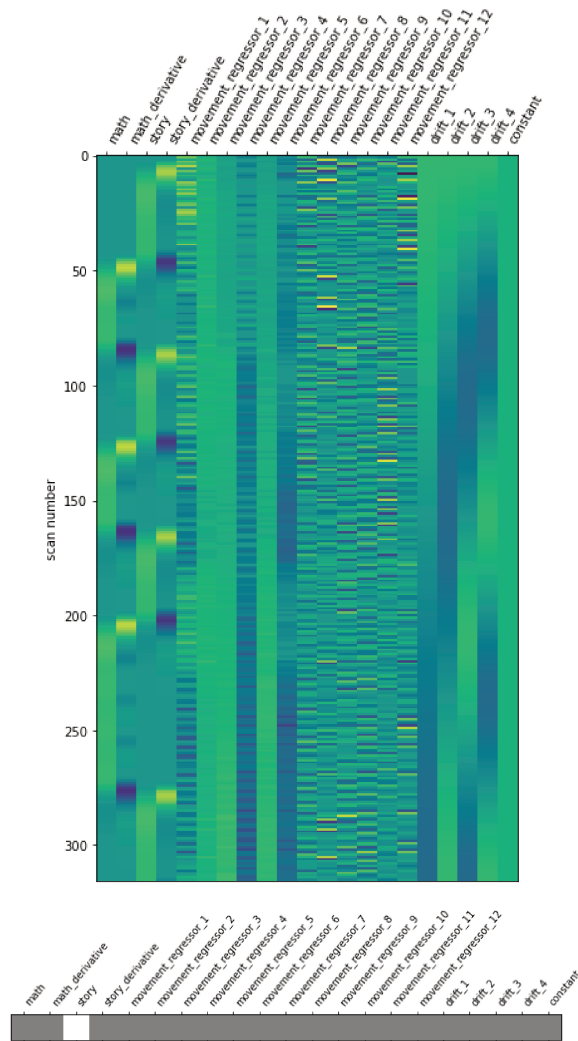


Figure 3.3 – **First level model.** A) Design matrix used in the first level model. Predictors includes fMRI events and derivatives, twelve movement regressors and four drift features. B) Coefficients of the contrast, indexed by the names of the columns of the design matrix.

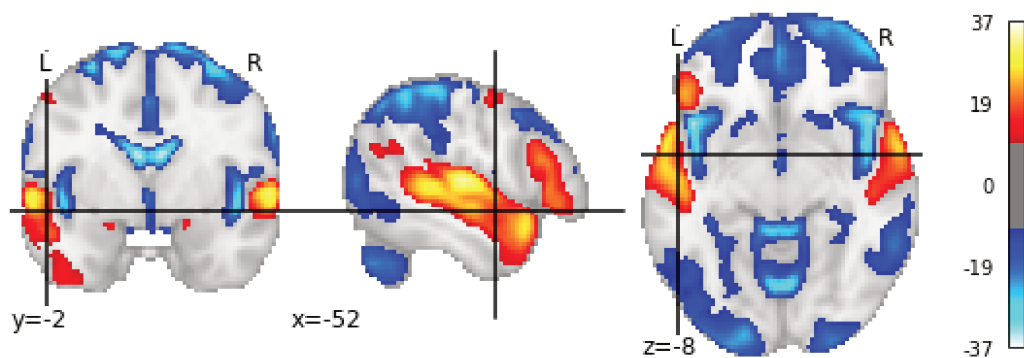


Figure 3.4 – **Group analysis language task activation maps.** Group average activations for the HCP language task: story versus baseline is shown with a lower threshold of $z = \pm 10$.

Estimating neural activation endophenotypes.

In this work, we estimated task-based neural activation in HCP, using perisylvian cortical areas as regions of interest (ROI) introduced in chapter 2.2.1. The obtained first-level z maps were masked using the ROIs and averaged for each individual. This resulted in 25 language neural activation regarding the SmPSS set of regions.

3.3 Resting-state functional connectome

As discussed in the chapter 1.3.2, intrinsic brain activity assessed while at rest can recover specific brain functional activations making it a good surrogate of functional neuronal processes. The identification of brain networks from fMRI data whether from resting-state or task-based, is often attempted by first, identifying a set of functional "nodes" such as ICAs maps or spatial ROIs. Next, conducting a connectivity analysis between the regions, based on the fMRI time-course signal averaged in each region. In the next section, we will summarise the pipeline undertaken in this work for the functional connectivity estimation.

Estimating functional connectome endophenotypes.

In this work, we estimated task-free FC in UK Biobank [Bycroft et al., 2018], using perisylvian cortical areas as regions of interest (ROI) introduced in chapter 2.2.1. Different measures can be considered regarding the estimation of the connectivity between nodes as highlighted in the chapter 1.3.2. General correlation-based approaches are known to be quite successful [Smith et al., 2011]. The simplest measure of pairwise similarity between two timeseries is covariance and in case of normalisation of the timeseries, we refer to it as *full correlation*. The correlation between two timeseries does not imply that the functional connection between two nodes is direct. Indeed, the correlation coefficient might give misleading results if there is a third node that is numerically related to both initial nodes. *Partial correlation* overcomes this limitation by ruling out indirect effects in the correlation structure. In this thesis, we chose to use the *Partial correlation* as it offers a more accurate interpretation of our results.

Overall, the functional connectivity estimation was performed by applying the following steps: the preprocessed resting-state BOLD signal was masked using the ROIs and averaged at each time volume. A connectome matrix was computed using Nilearn [Abraham et al., 2014] for each participant using a shrunk [Ledoit and Wolf, 2004] estimate of partial correlation [Marrelec et al., 2006] (See Fig. 3.5). This resulted in 300 ($= 25 \times 24/2$) and 15 ($= 6 \times 5/2$) edges connecting language ROIs for each individual for both the SmPSS and OcSS sets of regions, respectively. Each edge —also denoted functional connectivity (FC)— is further considered as a candidate endophenotype.

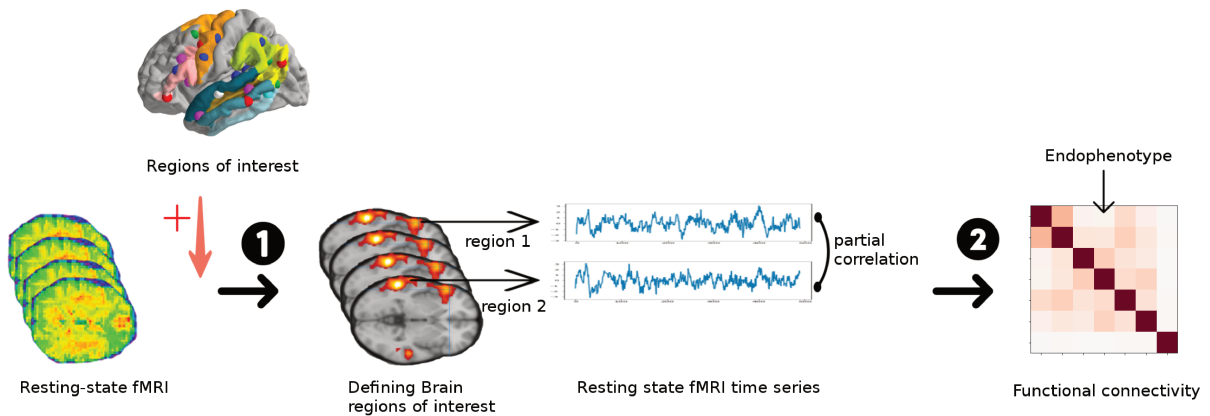


Figure 3.5 – **Functional connectome extraction pipeline** with two main steps: 1) Map of brain regions already defined on resting-state fMRI images. 2) Extraction of mean BOLD signal for each ROIs and quantifying functional interactions from these time series signal extracted. At the end of the second step, we obtain a functional connectome for each participants where each cell represents a resting-state functional connectivity.

This pipeline was applied to both UKB and HCP cohorts and the endophenotypes extracted can be appreciated in the Fig.3.6

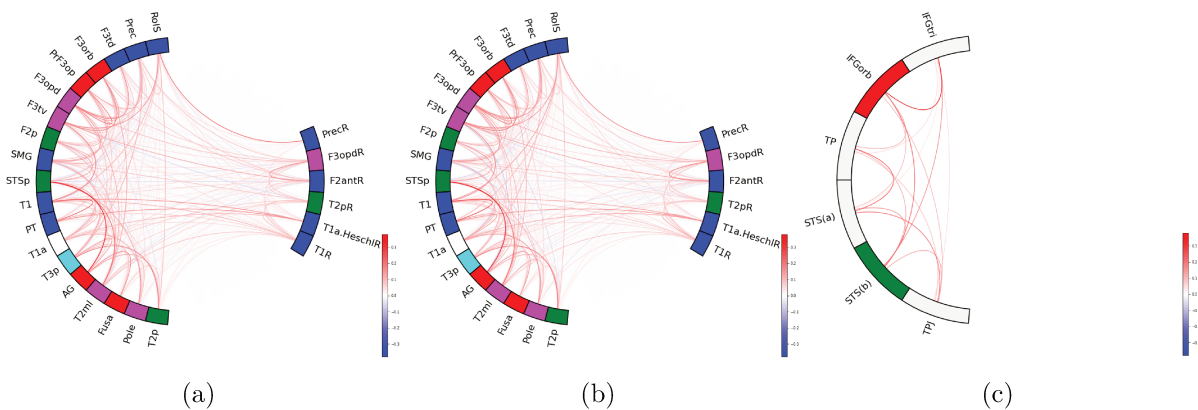


Figure 3.6 – **Functional connectome visualisation** Mean functional connectivity of the 300 and 15 endophenotypes, calculated using a shrunk estimate of partial correlation Marrelec et al. [2006] (estimated with a Ledoit-Wolf estimator Ledoit and Wolf [2004]) over 32,186 and 778 individuals using both UKB and HCP rsfMRI participants respectively. (a) Circos plot represents the mean FC estimated using SmpSS ROIs on UK Biobank. (b) Circos plot represents the mean FC estimated using SmpSS ROIs on HCP. (c) Circos plot represents the mean FC estimated using OcSS ROIs on UK Biobank.

3.4 Anatomical connectome

Diffusion-weighted imaging quantifies water diffusion in biological tissues and is used to identify the different fiber tracts in the human brain by tracking the directional diffusion of water in the brain. It provides a number of measures that can be used for statistical and genetic analyses. Regarding the heritability of structural connectivity, several genetic studies showed that DTI measures are heritable and therefore worthy targets for more in-depth genetic analysis.

More details are provided in the chapter 1.3.1. In the next section, we will summarize the methodological steps undertaken for the endophenotypes estimation.

Estimating anatomical connectome endophenotypes.

The anatomical connectivity endophenotypes are estimated using a set of 35 white matter tract masks defined from the probabilistic atlas [Rojkova et al., 2016] introduced in the chapter 1.5.2.2 and that are enumerated below: The arcuate fasciculus (long, anterior, and posterior segments), the superior longitudinal fasciculus (I, II, III), the inferior fronto-occipital fasciculus, the uncinate fasciculus, the frontal inferior longitudinal, the frontal superior longitudinal, the inferior longitudinal fasciculus, the frontal aslant tract, the fronto-insular tracts (1-5) and the corpus callosum.

In this thesis, we mainly conduct phenotypic multivariate analysis. Overlapping white matter tracts might be a serious limitation regarding the interpretation of the mvGWAS results. Indeed, the identification of the set of tracts that weight the most in the multivariate association becomes difficult. In light of this observation, we apply a stringent threshold at 90% of probabilities, defined after a visual inspection, in order to reduce significantly the overlap and allow us a fine interpretation of our results. Finally, the images derived phenotypes (IDPs) are estimated as follow: first, a binarisation of the set of 35 white matter tracts is performed to generate the tract-masks. Next, for each individuals, and for each DTI/NODDI output image type, an IDPs is generated by averaging the skeletonised images across the set of the tracts of interest. As these structural white matter tracts are assessed by 9 indices: fractional anisotropy (FA) maps, tensor mode (MO), mean diffusivity (MD), intracellular volume fraction (ICVF), isotropic volume fraction (ISOVF), mean eigenvectors (L1, L2, L3), and orientation dispersion index (OD), $315 = 35 \times 9$ dMRI endophenotypes are derived. Fig.3.7 summarize the methodological steps for the endophenotypes estimation.

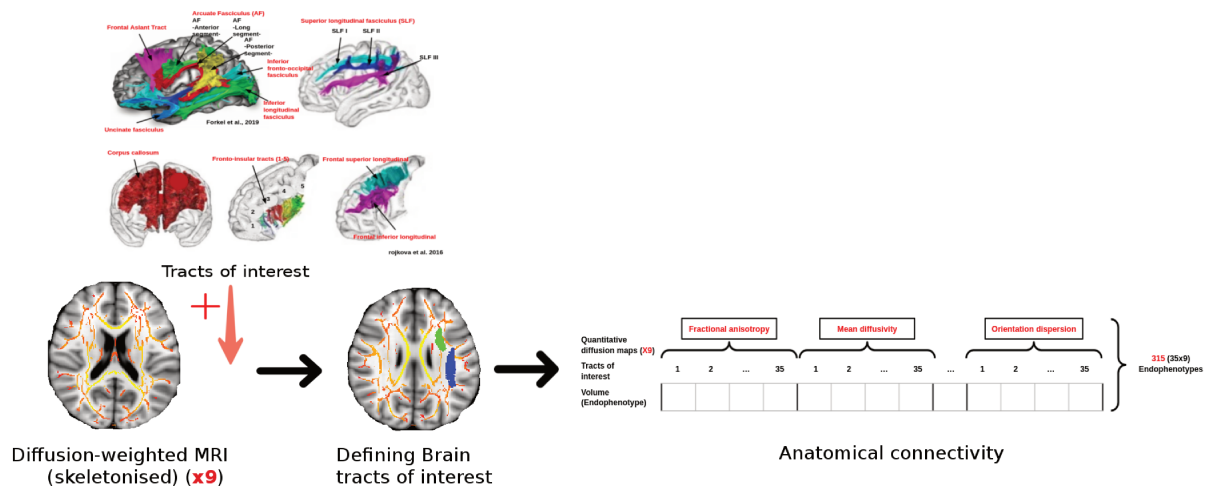


Figure 3.7 – **Anatomical connectome extraction pipeline** with one main steps consisting on mapping brain tracts of interest already defined on skeletonised images and extracting mean value for each tract of interest. At the end, we obtain an anatomical connectome for each participants where each cell represents an anatomical connectivity.

3.5 Conclusion

Multi-modal data collection in large-scale cohorts offers us a great opportunity to investigate open neuroscience questions from different perspectives. A valuable aspect of this work has been to exhibit different (endo)phenotypes regarding human language trait. These quantitative endophenotypes reflects human brain inter-individual variability. In order to ensure that these endophenotypes represents human language as cognitive process of interest, we adopted a region of interest approach. Whether for cortical regions or white matter tracts, the ROIs were carefully chosen. These ones were previously associated with language impairment, and/or identified in task based fMRI as supporting language and in case of WM tracts, reach language cortical regions. In the next chapter, we will investigate and show that many of these (endo)phenotypes are heritable and therefore worthy targets for more in-depth genetic analysis.

* * *
* *
*

Heritability of human language

Chapter Outline

4.1	Pedigree-based heritability of behavioural and neuroimaging (endo)phenotypes: a twin study	79
4.1.1	Cognitive behavioural assessment	79
4.1.2	Brain neural activation	80
4.1.3	Functional connectome	80
4.2	SNP-based heritability of neuroimaging endophenotypes: a general population study	81
4.2.1	Functional connectome	81
4.2.2	Anatomical connectome	83
4.3	Discussion	83
4.4	Are heritability estimates inflated in twin studies?	83
4.5	Conclusion	87

How do we decide if a cognitive trait such as language is genetically influenced? In quantitative genetics, heritability reflect cognitive traits variance that can be attributed to genetics versus environmental factors or their interactions. The traits may be either language related behaviour or language brain image derive measures (IDP). In chapter 1.4.1, the heritability concept has been detailed. Relevant to our study, heritability of human language has been extensively reported in the literature as revealed by reviews of the field [Stromswold, 2001]. Previous studies have mainly focused on twins, adoption, and linkage studies of language. The predominant genetics investigations concerns behavioural traits. With the recent availability of large-scale imaging-genetics cohorts, researches have addressed the heritability of language hypothesis regarding the left-right hemispheric asymmetry measured from neuroimaging [Carrion-Castillo et al., 2020, Le Guen et al., 2020, Sha et al., 2021]. Details about all these studies are provided in the chapter 1.6.1. However, we could not help but notice a lack of researches of the heritability of language regarding other available modalities. Thus, the present contribution seeks to provide elements of answers regarding this scientific question by taking advantage of the different (endo)phenotypes exhibited during this thesis and presented in the previous chapter. In this chapter, we use pedigree and real whole genome genotyping data as well as psychometric measures (chapter 3.1) and MRI data including task-based fMRI (see chapter 3.2), resting-state fMRI (see chapter 3.3), and diffusion-weighted MRI (see chapter 3.4) from the two cohorts introduced in the chapter 2.1; the Human Connectome Project (HCP) and the UK Biobank (UKB).

The objective of the present work is two-fold: *i)* to contribute to the study of language heritability and compare the obtained results with the literature; *ii)* to filter the endophenotypes for the multivariate genome-wide association study presented in both chapter 6 and 7.

Furthermore, the estimation of heritability reported by pedigree studies have been subject of multiple concerns as the estimation might be inflated [Hofer et al., 2018, Stromswold, 2001, van der Lee et al., 2017]. The possible origins of these discrepancies will be further discussed.

4.1 Pedigree-based heritability of behavioural and neuroimaging (endo)phenotypes: a twin study

Using SOLAR, we estimated the pedigree-based heritabilities of each endophenotypes. As covariates, the sex, the age, the ethnicity (Hispanic or not), and the first four principal components of the principal component analysis of the genetic data were considered. To account for multiple tests, a 0.05 threshold on False Discovery Rate (FDR) adjusted p-values was applied.

4.1.1 Cognitive behavioural assessment

The pedigree-based heritability (h^2) was estimated for each of the language scores we described in the chapter 3.1. After a FDR correction for multiple testing, all language scores have a significant pedigree-based heritability (Table 4.1), ranging from 19.4% for the language task math median reaction time to 71.7% for the oral reading recognition.

Table 4.1 – **Heritabilities of language related behavioral scores.** Endophenotypes in bold represent language related behavioral scores with significant heritability after false discovery rate correction at $p < 0.05$.

Endophenotype	heritability	P value	SE	FDR P value
Oral reading recognition	71.7%	4.74E-30	0.038	6.16E-29
Vocabulary Comprehension	64.2%	1.78E-24	0.045	1.16E-23
Fluid Intelligence	60.0%	9.54E-19	0.055	4.13E-18
Working Memory	50.2%	5.50E-15	0.058	1.79E-14
Language task math accuracy	43.5%	8.39E-12	0.061	2.18E-11
Language task overall accuracy	41.5%	3.94E-10	0.065	8.54E-10
Language task story difficulty level	36.9%	1.00E-07	0.069	1.86E-07
Fluid Intelligence reaction time	34.0%	3.00E-07	0.069	4.87E-07
Language task math difficulty level	34.1%	1.30E-06	0.071	1.88E-06
Language task story accuracy	19.9%	1.38E-03	0.068	1.64E-03
Language task overall median reaction	20.6%	1.38E-03	0.072	1.64E-03
Language task story median reaction time	19.5%	1.68E-03	0.069	1.82E-03
Language task math median reaction time	19.4%	2.87E-03	0.072	2.87E-03

4.1.2 Brain neural activation

The pedigree-based heritability (h^2) was estimated for each of the neural activations measured in the 25 SmPSS regions. In order to take into account the multiple tests, FDR correction on the p values was applied. This revealed 16 regions out of the 25, whose language activation displays a significant pedigree-based heritability (Table 4.2), ranging from 14.4% for the T2pR to 34.9% for the T1.

Table 4.2 – **Heritabilities of language task activation.** Bold endophenotypes represents language task activation with significant heritability after false discovery rate correction at $p < 0.05$.

Endophenotype	heritability	P value	SE	FDR P value
F2p	32.9%	2.00E-07	0.065	1.25E-06
T1R	32.3%	2.00E-07	0.064	1.25E-06
T1	34.9%	1.00E-07	0.065	1.25E-06
T2ml	33.0%	2.00E-07	0.064	1.25E-06
SMG	31.1%	5.00E-07	0.064	2.50E-06
T1a.HeschlR	30.5%	1.10E-06	0.066	4.58E-06
Pole	31.1%	4.90E-06	0.070	1.75E-05
AG	27.2%	7.90E-06	0.064	2.47E-05
T1a	27.9%	1.07E-05	0.066	2.97E-05
STSp	28.0%	7.34E-05	0.076	1.83E-04
PT	23.4%	1.47E-04	0.066	3.34E-04
PrF3op	21.2%	2.40E-04	0.062	5.00E-04
F3orb	19.8%	1.66E-03	0.070	3.19E-03
F3opdR	17.8%	3.22E-03	0.067	5.76E-03
F3opd	16.2%	5.39E-03	0.065	8.98E-03
T2pR	14.5%	8.50E-03	0.063	1.33E-02
F2antR	11.0%	4.59E-02	0.067	6.57E-02
Fusa	10.9%	4.73E-02	0.068	6.57E-02
PrecR	10.6%	6.05E-02	0.070	7.96E-02
F3td	8.7%	8.10E-02	0.064	1.01E-01
Prec	7.7%	1.05E-01	0.063	1.25E-01
T3p	7.5%	1.18E-01	0.065	1.34E-01
T2p	5.4%	1.99E-01	0.065	2.16E-01
F3tv	4.6%	2.40E-01	0.066	2.50E-01
RoIS	4.1%	2.54E-01	0.063	2.54E-01

4.1.3 Functional connectome

The pedigree-based heritability (h^2) was estimated for each of the 300 functional connectivity extracted from SmPSS set of ROIs. In order to take into account the multiple tests, FDR correction on the p values was applied. This revealed 217 FCs with a significant pedigree-based heritabilities (Fig. 4.1, Table B.1), ranging from 11.4% for the F2p↔F3td to 48.7% for the

F3tv↔F3orb FCs.

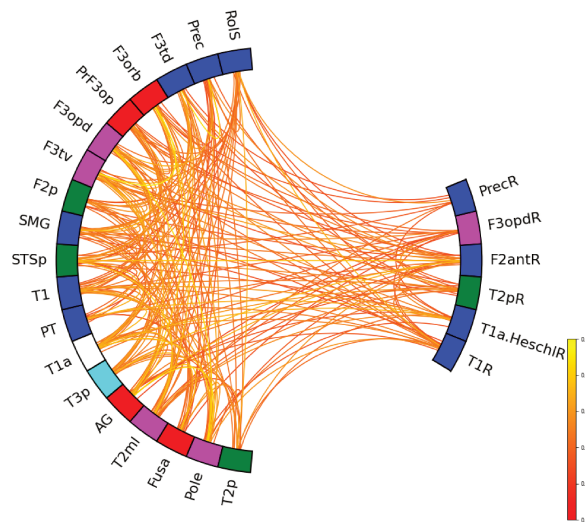


Figure 4.1 – Pedigree-based heritabilities of brain functional connectivities extracted using SmPSS ROIs. Circle plot where all 25 regions are represented (19 ROIs on the left represent the left hemisphere and 6 ROIs on the right represent the right hemisphere) illustrating the FCs with significant pedigree-based heritability after false discovery rate correction at $p < 0.05$. Each link between two regions represents an endophenotype. The intensity of the color between two regions represents the pedigree-based heritability of an endophenotype as defined by the color bar. Please refer to Table B.1 for all pedigree-based heritability values.

4.2 SNP-based heritability of neuroimaging endophenotypes: a general population study

Using GCTA, genomic-relatedness-based restricted maximum-likelihood (GREML) analyses were performed to estimate the SNP-based heritabilities of each endophenotypes. These ones were pre-residualised controlling for covariates including sex, genotype array type, age, recruitment site, and ten genetic principal components provided by UK Biobank. In addition, we performed a rank-based inverse-normal transformation of the residualised endophenotypes to ensure that the inputs are normally distributed. To account for multiple tests, a 0.05 threshold on False Discovery Rate (FDR) adjusted p-values was applied.

4.2.1 Functional connectome

The SNP-based heritability of the resting-state functional connectome was investigated using two set of ROIs introduced in the chapter 2.2.1. These are presented below.

4.2.1.1 Functional connectivity extracted using OcSS ROIs

The single-nucleotide polymorphism (SNP)-based heritability (h^2) was estimated for each of the 15 functional connectivity extracted from the OcSS set of ROIs. In order to take into account the multiple tests, FDR correction on the p values was applied. This revealed 14 FCs

with a significant SNP-based heritabilities (Fig. 4.2, Table B.2), ranging from 4.9% for the IFGtri \leftrightarrow pSTS to 12% for the IFGtri \leftrightarrow IFGorb FCs.

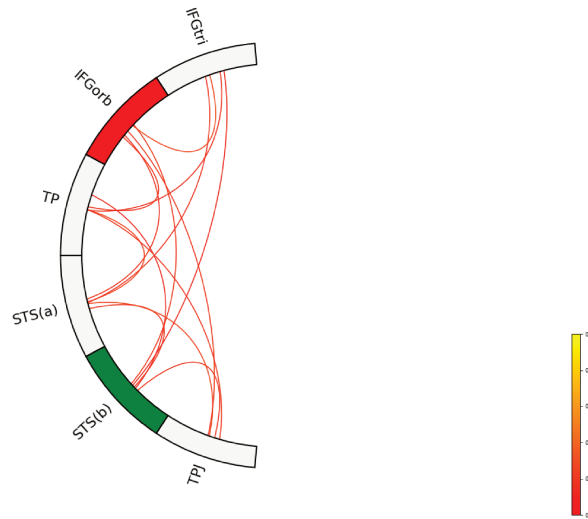


Figure 4.2 – **SNP-based heritabilities of brain functional connectivities extracted using OcSS ROIs.** Circle plot where all 6 regions are represented (6 ROIs on the left represent the left hemisphere. No region in the right hemisphere are included in OcSS ROIs.) illustrating the FCs with significant SNP-based heritability after false discovery rate correction at $p < 0.05$. Each link between two regions represents an endophenotype. The intensity of the color between two regions represents the SNP-based heritability of an endophenotype as defined by the color bar. Please refer to Table B.2 for all SNP-based heritability values.

4.2.1.2 Functional connectivity extracted using SmpSS ROIs

The single-nucleotide polymorphism (SNP)-based heritability (h^2) was estimated for each of the 300 functional connectivity extracted from SmpSS set of ROIs. In order to take into account the multiple tests, a FDR correction on the p-values was applied. This revealed 142 FCs with a significant SNP-based heritabilities (Fig. 4.3, Table B.3), ranging from 3.5% for the SMG \leftrightarrow T1 to 14.3% for the SMG \leftrightarrow F3opd FCs.

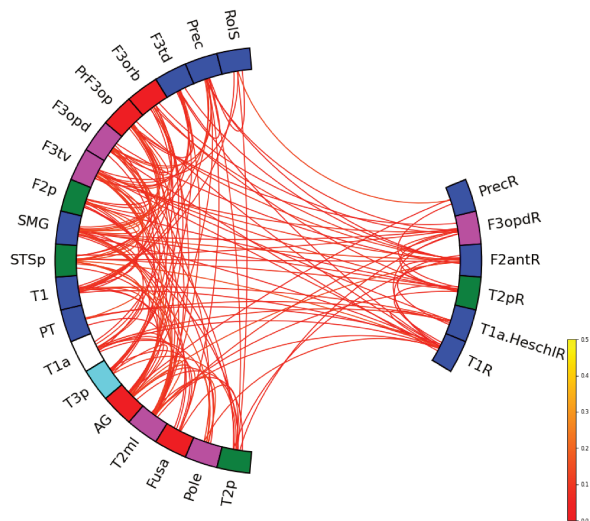


Figure 4.3 – SNP-based heritabilities of brain functional connectivities extracted using SmPSS ROIs. Circle plot where all 25 regions are represented (19 ROIs on the left represent the left hemisphere and 6 ROIs on the right represent the right hemisphere) illustrating the FCs with significant SNP-based heritability after false discovery rate correction at $p < 0.05$. Each link between two regions represents an endophenotype. The intensity of the color between two regions represents the SNP-based heritability of an endophenotype as defined by the color bar. Please refer to Table B.3 for all SNP-based heritability values.

4.2.2 Anatomical connectome

The single-nucleotide polymorphism (SNP)-based heritability (h^2) was estimated for each of the 315 bilateral bundles extracted measures of brain anatomical connectivity. All but two ACs (313) showed significant SNP-based heritability (False discovery rate (FDR)-corrected $p < 0.05$) (Table B.4), ranging from 3.8% for the left fronto Insular tract number 4 measured with OD to 62.7% for the corpus callosum measured with ICVF.

4.3 Discussion

A pivotal question in biology is the heritability of a particular trait. It provides a quantitative estimation of its association with a set of genetic markers and thus, informs us about the richness of its contribution in the trait of interest. We observed a hierarchy of language heritability estimates from the Human Connectome Project (twin study) with 19.4-71.7% for language-behavioural scores, 11.4-48.7% for resting-state FC, and 14.4-34.9% for language-task activations. In terms of heritability ranking, the resting-state functional connectivity edges show low levels of heritability compared to behavioural scores. This observation was also made regarding structural features such as cortical or white matter tracts features (in UKB [Elliott et al., 2018]). Overall, we show evidence that human language is, to some extent, under genetic control. Furthermore, we observed a discrepancy between both SNP-based and Pedigree-based heritabilities. The possible origins of these discrepancies will be further discussed.

4.4 Are heritability estimates inflated in twin studies?

We estimated the heritability of human language considering different endophenotypes. These ones displayed a substantial genetic influence. We noticed a discrepancy between heritability estimates depending on the cohort used, with higher estimates using the HCP compared to the UK Biobank. For the sake of clarity, we focus here on the heritability estimates of the brain functional connectivities extracted using SmPSS ROIs.

The origin of this discrepancy may be related to differences in either the design parameters of the analyses, or the cohorts' environmental conditions which differ significantly. As regards the design parameters like the length of the rsfMRI used to compute the FC, global effects can be seen in Fig. 4.4 which highlights the heritability differences. This figure also reports the global minor differences in heritability estimates when using SOLAR or GCTA in a same cohort (HCP) for which both genotyping and pedigree information were available.

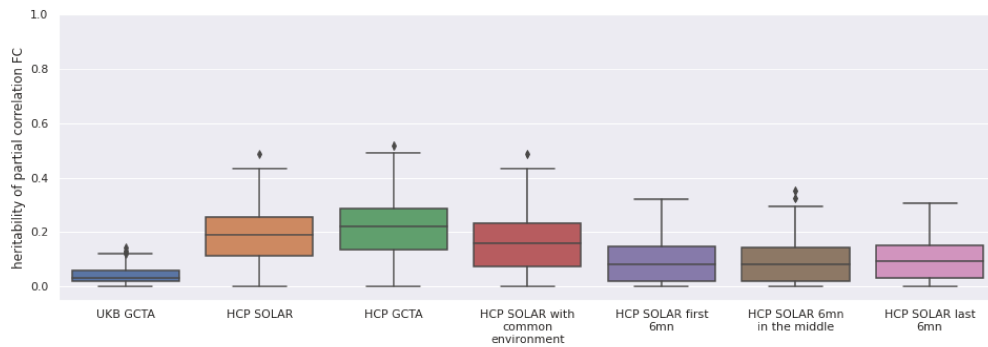


Figure 4.4 – **Comparison of heritabilities estimated according to different design parameters.** The endophenotypes extracted using SmpSS ROIs. Boxplots 1 and 2 display the heritability distributions estimated on UKB (using GCTA) and HCP (using SOLAR) respectively. Boxplot 3 : HCP-FCs using GCTA. Boxplot 4 : HCP-FCs using GCTA and considering the household effect. Boxplots 5 to 7 : HCP (using SOLAR) on rsfMRI sessions with the same duration as in UKB (6 min).

In order to shed lights on the origin of the observed heritability gap, we attempted to rule out a certain number of hypotheses supposed to be at the origin of these discrepancies. **First**, we investigated whether the observed gap is due to the difference of genetic information (pedigree/genotyping data) between both cohorts. As both genetic information are available in the HCP cohort, we computed the heritability of FCs using pedigree in one hand and using the genotyping data in an other hand. We observed that the heritability computed on HCP genotyped SNPs are slightly higher than those obtained with pedigrees (See Fig.4.4). Although, these estimations are highly correlated ($r = 0.97, p = 1.1e - 209$, Fig.4.5c), it is noteworthy that the SNP-based heritability estimation does not capture all genetic information. Indeed, only common SNPs are considered in our analysis. Rare SNPs variants, epigenetic variations, copy number variation are not modelised in the analysis. Additionally, genotype-based heritability with GCTA assume no shared environment. In practice, this is achieved by using unrelated individual. As HCP cohort is composed of twins and siblings, this HCP-GCTA analysis do not follow that standard and its derived results can be biased. These results should be taken with caution. Furthermore, while SNP-based GCTA heritability captures the SNP heritability, pedigree-based SOLAR heritability captures a much broader version of the heritability including for example non-additive effect [Manolio et al., 2009]. Starting from this observation, the comparison between both estimates became hard to do and irrelevant.

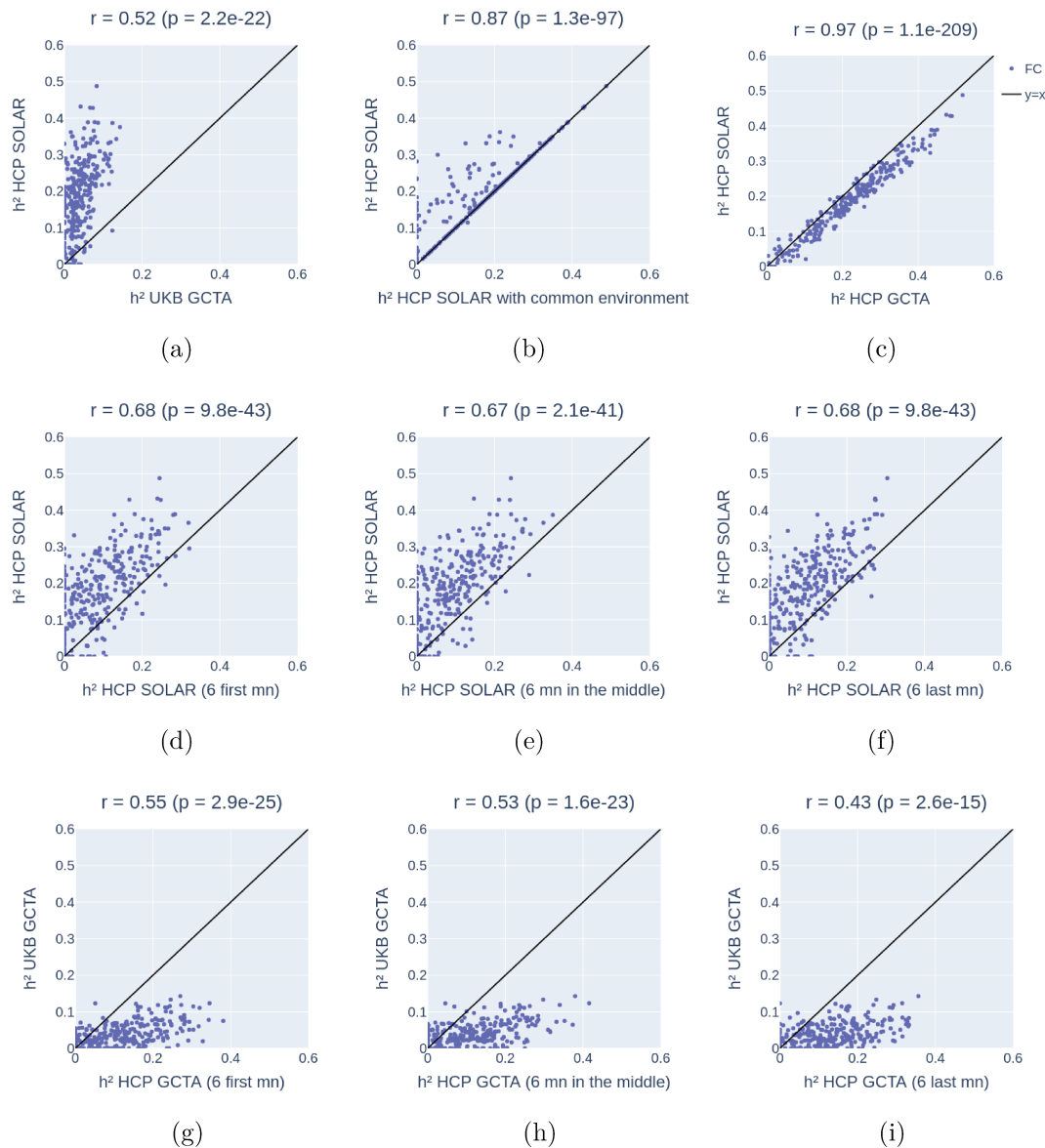


Figure 4.5 – Comparison between heritabilities obtained with SOLAR on HCP (concatenation of the 4 runs) and the ones obtained with : (a) GCTA on HCP, (b) SOLAR on HCP taking into account the shared environment, (c) GCTA on UKB, (d) SOLAR on HCP (6 first minutes of the signal), (e) SOLAR on HCP (6 minutes in the middle of the signal), (f) SOLAR on HCP (6 last minutes of the signal)

Second, we took into account the shared environment effect in the HCP cohort and found that the heritability estimations remained highly similar ($r = 0.87, p = 1.3e-97$) to the ones estimated without modeling the shared environment (See Fig.4.5b). These results could, to some extent, be compared to the adoption studies presented in chapter 1.6.1. The adoption studies compare adopted children’s linguistic abilities with those of their adopted and biological relatives. Taken as a whole, these studies suggest that genetic factors affect language abilities. This observation is in line with our findings. However, this results should be taken with caution. One question arises: To what extent the household feature capture the environment variable

and to what extent, should we compare it to adoption studies which are certainly more relevant for the estimation of the human language environment part?

Third, as highlighted by [Elliott et al., 2019], long resting-state time series allow a better estimation of intrinsic connectivity and thus higher estimates of heritability. The resting-state time series last longer in HCP compared to UKB. We thus cut the time series the HCP's participants and considered the 6 first minutes, the 6 minutes in the middle and finally the 6 last minutes as shown in the Fig.4.6. We calculated resting-state FC endophenotypes and estimated their heritabilities. As in [Elliott et al., 2019], we observe that the values of h^2 are higher for the whole uncut sessions compared to the 6-minute sessions in HCP (See Fig.4.5d, 4.5e, 4.5f). This suggest that long resting-state fMRI time series improve the reliability of intrinsic connectivity estimates and subsequently, it better captures inter-individual differences.

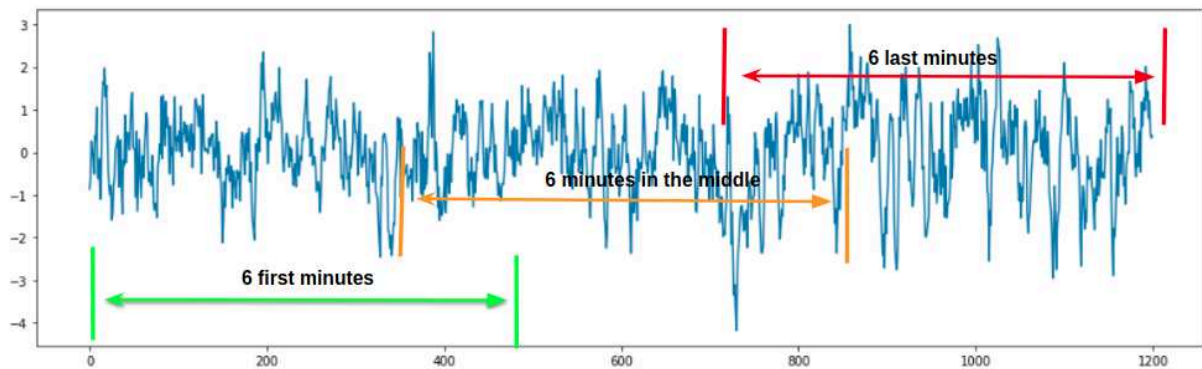


Figure 4.6 – Different cut of the resting-state time serie (6 minutes) obtained from HCP cohort.

Fourth, we compared the heritability of FC obtained from HCP cohort (6 first minutes, 6 minutes in the middle and 6 last minutes) using genotyping data (GCTA [Yang et al., 2011]) with the heritability of FC obtained from UKB cohort (6 minutes) using genotyping data as well (GCTA [Yang et al., 2011]). Although the session lengths are now comparable, a significant discrepancy between h^2 values remains (See Fig.4.5g, 4.5h, 4.5i). Nevertheless, the overall pattern remains correlated: UKB/HCP-(6 first minutes) $r = 0.55, p = 2.9e - 25$; UKB/HCP-(6 minutes in the middle) $r = 0.53, p = 1.6e - 23$; UKB/HCP-(6 last minutes) $r = 0.43, p = 2.6e - 15$. One potential explanation of this observed gap concerns the age of the participants in each cohort. Contrary to UK Biobank which constitutes a relatively old sample (mean age= 54 ± 7.45 years), participants in HCP cohort are younger (mean age= 29 ± 3.56 years) and thus, the influence of the environmental factors may differ. A growing body of literature examines changes in heritability estimates over time [Bergen et al., 2007]. For example [Bouchard, 2013] show evidence that the heritability of IQ increases with age. It reaches an asymptote at about 0.80 at 18~20 years of age and continuing at that level well into adulthood. Furthermore, the heritability of IQ appears to diminish with advancing age. In a sample of adult twins (aged from 27 to 59) and older cohort (aged from 60 to 94), a decrease of verbal and performance IQ heritability measures was observed with 70%, 73% and 56% and 60% respectively [Finkel and McGue, 1998]. [Bergen et al., 2007] suggests that beyond the tumultuous formative years, environmental effects may

begin to accumulate and form a larger proportion of the total phenotypic variance. Too little research has been done to deduce the overall pattern of heritability changes in later adulthood. Future studies that cover different developmental stages; i.e. adolescence and late adulthood, will inform us whether the observed heritability reflect some age-related specificity.

Overall, these different hypotheses contributed to explain the gap observed between both SNP-based and pedigree-based human language heritability estimates. The results obtained are in line with [Le Guen et al., 2019] work that used surface area and cortical thickness as endophenotypes. Although these different hypotheses helped to explain the observed discrepancy between the two heritability estimates, this experimental design does not allow us to present a definitive answer that would explain this phenomenon.

4.5 Conclusion

We observed significant heritabilities of the language (endo)phenotypes. These heritabilities display a large range of values depending on the trait considered but also the design of the cohort: twin versus general population cohorts. This work emphasises that human language brain organisation is under relatively strong genetic control. We found lower levels of SNP heritability for functional connectivity compared to heritability estimated in twin cohorts. This is a frequently voiced concern and typical of many traits in the literature [Elliott et al., 2018, Yang et al., 2015]. It may results from upward bias in twin study estimates due to gene–gene and gene–environment interactions [Purcell, 2002, Zuk et al., 2012], and/or downward bias of SNP heritability due to uncaptured rare genetic variation.

* * *
* *
*

Shared genetic variance between language-related performance, language-brain activations and language functional connectivity

Chapter Outline

5.1	Shared heritability between both resting-state functional connectivity and neural activations	91
5.2	Shared heritability between both resting-state functional connectivity and language-related scores	92
5.3	Shared heritability between both task activations and language-related scores	92
5.4	Discussion	93

In the previous chapter, we estimated the human language heritability in both psychometric and imaging approaches. We used two large-scale resources currently available, namely the Human Connectome Project (HCP) and the UK Biobank (UKB) introduced in the chapter 2.1 to contribute to the study of language heritability. We observed a hierarchy of language heritability estimates from the Human Connectome Project (twin study) with 19.4-71.7% for language-behavioural scores, 11.4-48.7% for resting-state FC, and 14.4-34.9% for language-task activations. In terms of heritability ranking, the resting-state functional connectivity edges show low levels of heritability compared to behavioural scores. This observation was also made regarding structural features such as cortical or white matter tracts features (in UKB [Elliott et al., 2018]).

Although functional MR imaging is recognised to produce valuable endophenotypes [Elliott et al., 2018], resting-state FCs between language regions represent remote measures of language as the participants are not engaged in any language task-experiments and simply stay at rest. As stated in the chapter 1.3.2.1, previous works reported that resting-state derived endophenotypes are correlated with behavioural language processing [Cheema et al., 2021, Cross et al., 2021, Koyama et al., 2011, Stevens et al., 2017]. However, with the current release, no behavioural language scores were provided in the UK Biobank cohort. We were thus unable to confirm the observed correlations between resting-state FC endophenotypes and language behavioural scores.

In this chapter, we want to take advantage of the Human Connectome Project cohort which comprises language-related scores, language task-based fMRI data as well as resting-state fMRI data to calibrate the relevance of the resting-state image derived endophenotypes in the human language context through a bivariate heritability study. To do so, we performed bivariate genetic analyses to quantify the shared genetic influence between human language measure. In this chapter, we use pedigree and psychometric measures (chapter 3.1) and MRI data including neural activation measured in language task fMRI experiment (see chapter 3.2), functional connectivity estimated in resting-state fMRI (see chapter 3.3) from the Human Connectome Project (HCP) cohorts introduced in the chapter 2.1. We restricted our analysis to a small number of cognitive performance scores measured during the MRI session and considered the most representative of the human language. We used the average of accuracy of both conditions (STORY and MATH) in the language task to characterise the individual performance. Regarding the NIH behavioural scores, the working memory, the oral reading recognition, and the vocabulary comprehension scores were further considered in this analysis. All of these features are well detailed in the chapter 3. A large number of endophenotypes are derived from these three modalities. This methodological choice aims to reduce the number of tests to be performed and alleviates the correction for multiple tests. Furthermore, and for the sake of multiple corrections, we first conducted a (endo)phenotypic correlation to quantify the relationship between the human language performances (scores), and the two functional traits derived from the imaging. The significant (endo)phenotype pairs were subjected to a bivariate genetic analysis. The latter one quantify the shared genetic part taken in the correlation of these (endo)phenotypes pairs. As covariates, the sex, age, the ethnicity (Hispanic or not), and the first four principal components

of the principal component analysis of the genetic data were considered. Furthermore, to define significantly phenotypically and genetically correlated (endo)phenotypes, a 0.05 threshold on False Discovery Rate (FDR) adjusted p-values was applied to account for multiple testing on the paired (endo)phenotypes for both phenotypic and genotypic correlation analysis.

5.1 Shared heritability between both resting-state functional connectivity and neural activations

We computed the endophenotypic correlation between resting-state functional connectivity (300 FCs) and neural activations (25). Table C.1 summarises the endophenotypic correlation and their associated p-values. P-values correction for multiple testing revealed 215 among 7500 = 300*25 paired FCs-neural activations endophenotypes ranging from $r = -0.34, p_{FDR} = 4.76e-19$ for the PrF3op↔SMG with PrF3op to $r = 0.11, p_{FDR} = 4.95e-02$ for the F3td↔RolS with F3td. Next, we performed a bivariate genetic analysis using these 215 paired endophenotypes. The shared genetic variance estimates for these endophenotypes are presented in the Table 5.1 P-values correction for multiple testing revealed 12 among 215 significantly correlated endophenotypes pairs ranging from $rg = 0.89, p_{FDR} = 3.19e-2$ for the T1↔F2antR with T1 to $rg = 0.48, p_{FDR} = 3.19e-02$ for the Pole↔AG with AG.

Table 5.1 – **Shared heritabilities between both resting-state functional connectivity and neural activations.** Bold rows represents pairs of endophenotypes with significant shared heritability after false discovery rate correction at $p < 0.05$. For the sake of readability, only significant results are displayed. Pearson Correlation between pairs of endophenotypes are also shown. The asterisk represent significant endophenotypic correlation after false discovery rate correction at $p < 0.05$. Table C.1 summarises the endophenotypic correlation and their associated p-values.

Endophenotype 1	Endophenotype 2	Endophenotypic correlation	Genetic correlation	P value	SE	FDR P value
F2p↔SMG	F3td	0.128*	100.0%	1.03E-03	Not Computable	3.19E-02
F3td↔T3p	F3td	-0.115*	-100.0%	2.16E-03	Not Computable	3.84E-02
T1↔F2antR	T1	-0.177*	-89.0%	1.51E-03	0.376	3.19E-02
Prec↔F2antR	PrF3op	-0.124*	-76.6%	6.29E-04	0.242	3.19E-02
T2ml↔F3opdR	T2ml	-0.137*	-74.9%	9.93E-04	0.263	3.19E-02
Pole↔F3opd	F3opd	0.254*	74.4%	1.10E-04	0.173	2.35E-02
F3opd↔SMG	F3opd	-0.238*	-65.2%	4.15E-04	0.172	3.19E-02
Fusa↔T3p	PT	0.137*	57.4%	1.22E-03	0.176	3.19E-02
T1↔T1a.HeschlR	T1	0.197*	53.6%	1.64E-03	0.172	3.19E-02
PT↔T1	T1R	0.181*	50.2%	1.36E-03	0.148	3.19E-02
AG↔F3orb	AG	0.149*	49.6%	1.37E-03	0.154	3.19E-02
Pole↔AG	AG	0.154*	48.2%	7.46E-04	0.138	3.19E-02

5.2 Shared heritability between both resting-state functional connectivity and language-related scores

We computed the endophenotypic correlation between human language performances (4 language-related scores), and the functional connectivity estimated in resting-state fMRI (300 FCs). Table C.2 summarises the endophenotypic correlation and their associated p-values. None of the $1200 = 300 * 4$ pairs FCs-language-related scores endophenotypes passed the correction for multiple testing.

5.3 Shared heritability between both task activations and language-related scores

We computed the endophenotypic correlation between human language performances (4 language-related scores) and neural activations (25). Table C.3 summarises the endophenotypic correlation and their associated p-values. P-values correction for multiple testing revealed 28 among $100 = 25 * 4$ pairs FCs-neural activations endophenotypes ranging from $r = 0.24, p_{FDR} = 1.09e-09$ for the vocabulary comprehension and Pole to $r = 0.09, p_{FDR} = 3.78e-02$ for the oral reading recognition and PT. Next, we performed a bivariate genetic analysis using these 28 pairs of endophenotypes. The shared genetic variance estimates for these endophenotypes are presented in the Table 5.2. P-values correction for multiple testing revealed 23 significantly correlated endophenotypes pairs ranging from $rg = 0.61, p_{FDR} = 4.05e-4$ for the Language task accuracy and the temporal pole to $rg = 0.26, p_{FDR} = 2e-02$ for the vocabulary comprehension and T2ml.

Table 5.2 – **Shared heritabilities between both task activations and language-related scores.** Bold rows represents pairs of endophenotypes with significant shared heritability after false discovery rate correction at $p < 0.05$. For the sake of readability, only significant results are displayed. Pearson Correlation between pairs of endophenotypes are also shown. The asterisk represent significant endophenotypic correlation after false discovery rate correction at $p < 0.05$. Table C.3 summarises the endophenotypic correlation and their associated p-values.

Endophenotype 1	Endophenotype 2	Endophenotypic correlation	Genetic correlation	P value	SE	FDR P value
Pole	Language task overall accuracy	0.217*	61.8%	7.23E-05	0.155	4.05E-04
STSp	Language task overall accuracy	0.130*	61.1%	2.18E-04	0.175	8.72E-04
T1R	Working Memory	0.147*	59.3%	5.83E-06	0.134	8.16E-05
Pole	Working Memory	0.148*	57.0%	3.23E-05	0.142	2.26E-04
Pole	Vocabulary Comprehension	0.239*	56.3%	7.36E-07	0.112	2.06E-05
T1R	Language task overall accuracy	0.095*	53.3%	5.67E-04	0.167	1.30E-03
T2ml	Working Memory	0.145*	53.0%	2.41E-05	0.125	2.25E-04
PT	Working Memory	0.115*	50.9%	4.75E-04	0.146	1.23E-03
T2ml	Language task overall accuracy	0.142*	50.8%	3.83E-04	0.143	1.19E-03
STSp	Working Memory	0.133*	49.9%	4.82E-04	0.145	1.23E-03
T1a	Language task overall accuracy	0.175*	49.0%	2.27E-03	0.162	4.24E-03
STSp	Vocabulary Comprehension	0.183*	44.8%	3.48E-04	0.131	1.19E-03
Pole	Oral reading recognition	0.176*	41.8%	1.15E-04	0.108	5.36E-04
T1	Language task overall accuracy	0.093*	39.6%	3.82E-03	0.139	6.69E-03
T1a	Working Memory	0.142*	35.6%	1.04E-02	0.132	1.46E-02
T1R	Vocabulary Comprehension	0.105*	34.9%	1.78E-03	0.115	3.56E-03
T2ml	Oral reading recognition	0.142*	34.8%	6.02E-04	0.103	1.30E-03
T1	Working Memory	0.113*	34.7%	4.59E-03	0.122	7.14E-03
STSp	Oral reading recognition	0.108*	31.2%	6.97E-03	0.120	1.03E-02
T1R	Oral reading recognition	0.165*	30.0%	4.12E-03	0.000	6.79E-03
T1a	Vocabulary Comprehension	0.138*	29.9%	1.31E-02	0.122	1.74E-02
T1a.HeschlR	Working Memory	0.093*	28.0%	3.45E-02	0.132	4.21E-02
T2ml	Vocabulary Comprehension	0.126*	26.4%	1.58E-02	0.110	2.01E-02

5.4 Discussion

In the previous chapter, we showed that language related psychometric measures and image derived phenotypes are heritable. Although resting state fMRI data are reported as heritable, little is known about their relation to the language-related cognitive process. In this chapter, we address this question regarding the shared genetics between these ones and direct measures of language as cognitive process. To do so, we estimated phenotypic correlations between cognitive abilities, neural activations and task-free data in language. Moreover, we estimated the genetic proportion in these phenotypic correlations. Hence, we showed a shared aetiology between resting-state functional connectivity and neural activations. This work advocate for the existence of a link between resting-state and task-based data both phenotypically and genetically regarding human language brain organisation.

* * *
* *
*

The genetic architecture of human language functional connectome

Chapter Outline

6.1	Multivariate genome-wide association analysis	97
6.2	Unraveling the genome: Evaluation of results from mvGWAS study of human language.	107
6.3	Limitations.	121
6.4	Conclusions	122

In the chapter 3, we exhibited different (endo)phenotypes including resting-state functional connectome regarding human language traits. In the chapter 4, we brought out that resting-state FC are, to some extent, under genetic control. This reinforces our belief that resting state FC endophenotypes capture inter-individual variation in brain function, and thus, are worthy targets for more in-depth genetic analysis. To date, no gene discovery experiments have been reported using resting-state functional MRI derived traits. Our study is therefore an attempt to fill this observed gap in the literature and to verify the relevance of task-free paradigm for the study of complex cognitive trait such as language.

Notwithstanding that rsfMRI FC are heritable, they are reported as showing the lowest levels of SNP heritability among available IDPs in UKB [Elliott et al., 2018]. In a massively univariate framework, it would be really difficult to disentangle the genetic associations with each FC signal and reach genome-wide statistical significance. Language-related brain regions share information across components and scales that support language, and genetic variants are supposed to have distributed effect across regions. Contrary to the massive univariate approach, multivariate GWASes have the advantage of leveraging the distributed nature of genetic effects and the presence of pleiotropy across modalities and is thus, more consistent with biology. Therefore, we consider globally this synergistic system and perform a multivariate approach with MOSTest [van der Meer et al., 2020]. This method introduced in the chapter 2.6 considers the distributed nature of genetic signals shared across brain regions and aggregates effects across spatially distributed traits of interest.

As mentioned in the chapter 2.2.1, in this thesis, two sets of regions of interest were employed for defining the language network. The first one named -OcSS- is defined based on the probabilistic activation map of a single fMRI contrast and includes the perisylvian regions which cover both Broca and Wernicke's areas. The second one named -SmPSS- represents a more extended language network and is a result of a meta-analysis regrouping and analysing a significant number of scientific reports. The latter one offers a more general view of the language regions, a rich interpretation of the regions' classification regarding language components (phonology, semantic and syntax) and more importantly, it offers potentially a more consensual set of ROIs within the neuroscientific community. In the chapter 4, we estimated the human language heritability using both resting-state functional connectivity extracted these two sets of ROIs. In this chapter, the same experimental design was applied to these two sets of endophenotypes. The obtained results will be compared to each other and will allow us to judge how representative the SmPSS set of ROIs actually is. Therefore, we performed a multivariate genetic association of the significantly heritable FC endophenotypes of each set separately. These consist of 142 FCs and 14 FCs extracted from both SmPSS and OcSS ROIs (see chapter 4.2.1). The results from these analyses were subjected to a replication study in an independent sample ($N=4,754$). Additionally, as the connections between different language regions are ensured by the white matter fiber bundles [Catani and Forkel, 2019, Catani et al., 2005], we hypothesised that the hit SNPs associated with the hit FCs could be associated with neuroanatomical white matter tracts that support the information transmission between the regions that compose these hit-FCs. Therefore, we tested the potential associations of the hit SNPs with the following white matter bundles: the corpus callosum, the left frontal aslant tract, the left arcuate anterior/long/posterior segment,

the left inferior fronto-occipital fasciculus, the left uncinate tract (See Chapter 2.2.2 for more details). We used the endophenotypes resulting from the TBSS analysis described in the chapter 3.4.

6.1 Multivariate genome-wide association analysis

We performed a multivariate genome-wide association studies (mvGWAS) between the filtered imputed genotypes and the two sets of significantly heritable endophenotypes highlighted in the chapter 4.2.1.2, using the Multivariate Omnibus Statistical Test (MOSTest) [van der Meer et al., 2020]. As stated in the chapter 2.6, all endophenotypes were pre-residualised with covariates including sex, genotype array type, age, recruitment site, and ten first genetic principal components provided by UK Biobank. In addition, MOSTest performs a rank-based inverse-normal transformation of the residualised endophenotypes to ensure that the inputs are normally distributed. The distributions across the participants of all endophenotypes were visually inspected before and after covariate adjustment. MOSTest generated multivariate summary statistics that capture the significance of the association across all heritable language endophenotypes. To account for multiple testing over the whole genome, statistically significant SNPs were considered as those reaching the genome-wide threshold $p = 5e-8$.

GWAS using FCs derived from SmPSS set of ROIs. Concerning the FCs derived using SmPSS ROIs, we performed a mvGWAS with the 142 FCs with significant SNP-based heritability. This analysis tested each SNP separately for its simultaneous association with the 142 FCs and yielded 4566 significant SNPs at a genomic threshold, distributed on chromosomes 2, 3, 5, 6, 10, 11, 14, 15, 17, 18 and 22. FUMA [Watanabe et al., 2017] software was used to analyse mvGWAS results and identified lead SNPs at each associated locus. Considering the genome-wide significance threshold $p = 5e-8$, there were 20 distinct genomic loci distributed on the 11 chromosomes, associated with different aspects of language FC (Fig.6.1a, Table 6.1 and Supplementary Fig. 6.2, 6.3, 6.1b, and 6.5) and represented by 20 lead SNPs.

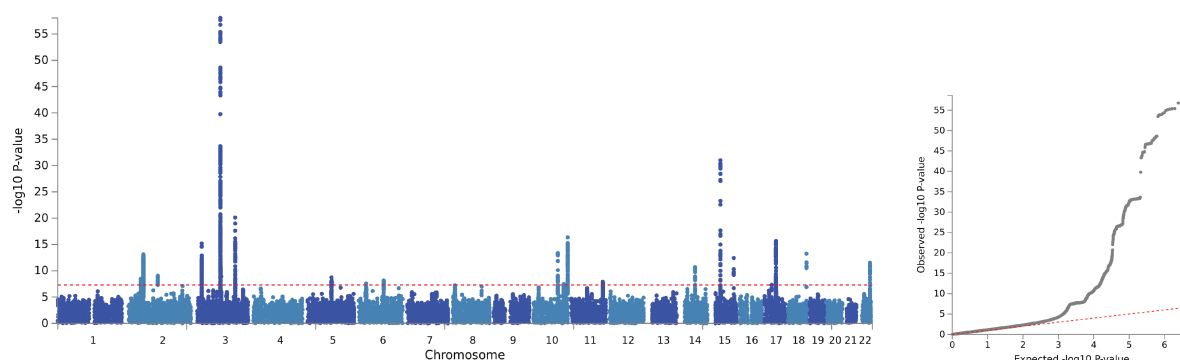
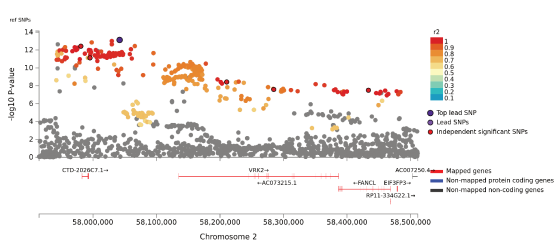
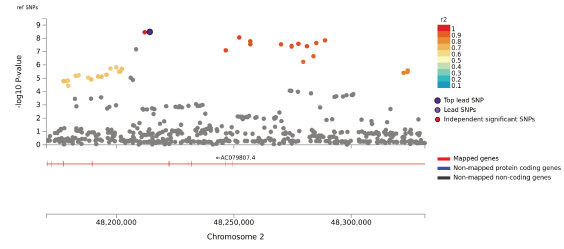


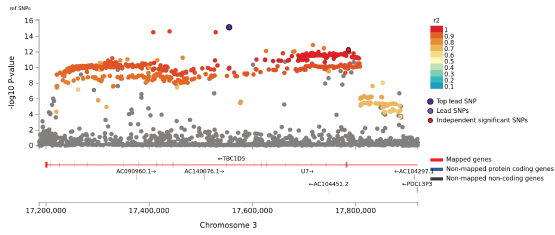
Figure 6.1 – **Multivariate GWAS analysis of the SmPSS resting state functional connectivity in 32,186 participants.** Manhattan plot for multivariate GWAS across 142 FCs. The red dashed line indicates the genome-wide significance threshold $p = 5e-8$. The Quantile-quantile plot is also shown.



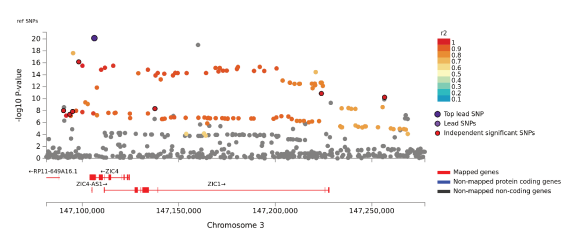
lead SNP: rs2717046



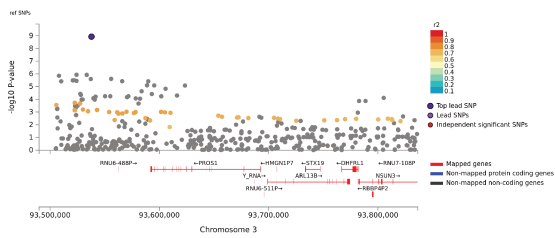
lead SNP: rs62141276



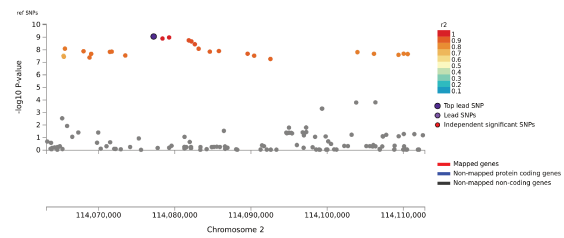
lead SNP: rs67851870



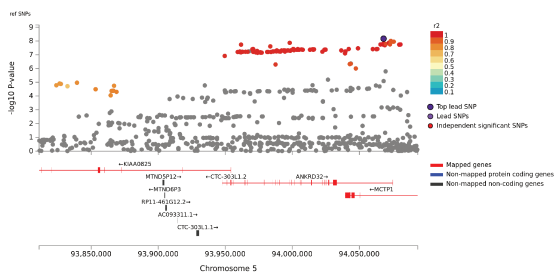
lead SNP: rs2279829



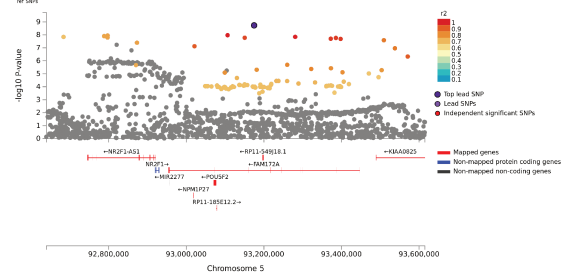
lead SNP: rs62266110



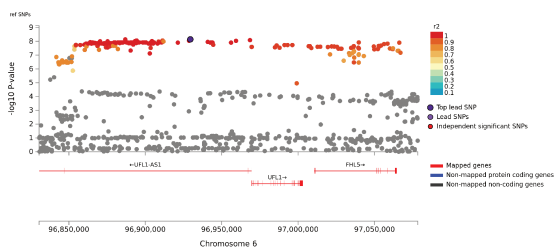
lead SNP: rs62158166



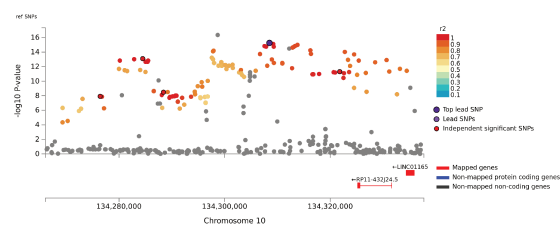
lead SNP: 5:94068140_AC_A



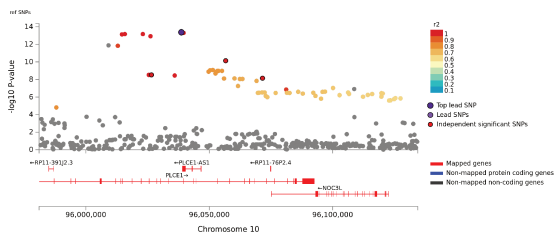
lead SNP: rs145120402



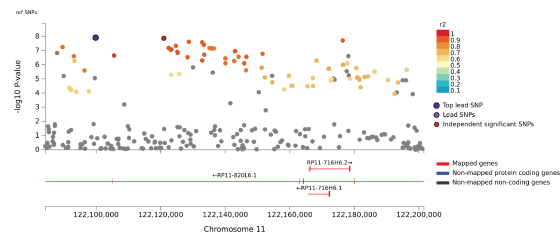
lead SNP: rs4262195



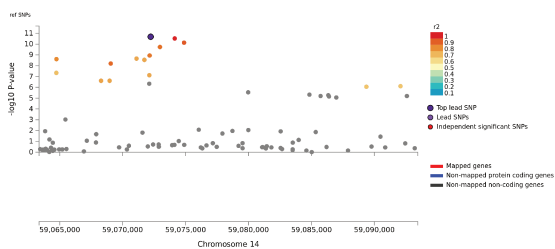
lead SNP: rs11146399



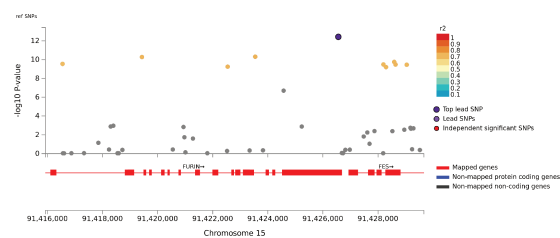
lead SNP: rs11187838



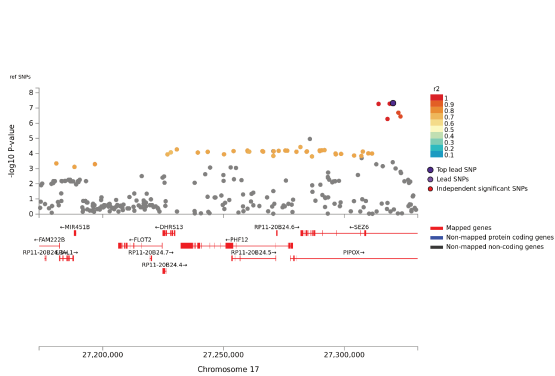
lead SNP: rs11218577



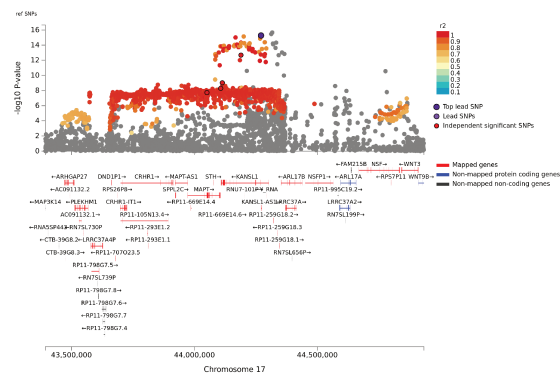
lead SNP: rs186347



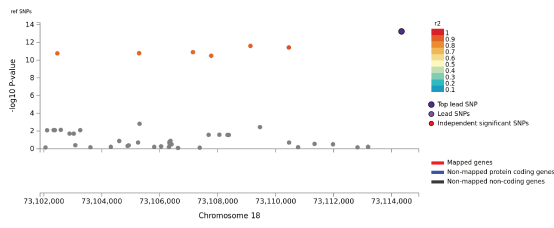
lead SNP: rs4702



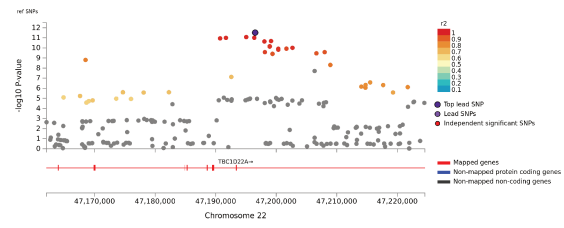
lead SNP: rs34039488



lead SNP: rs1744270659_G_A



lead SNP: rs7234875



lead SNP: rs2542028

Figure 6.2 – Locus Zoom of the significant loci identified by the multivariate GWAS for functional connectivity.

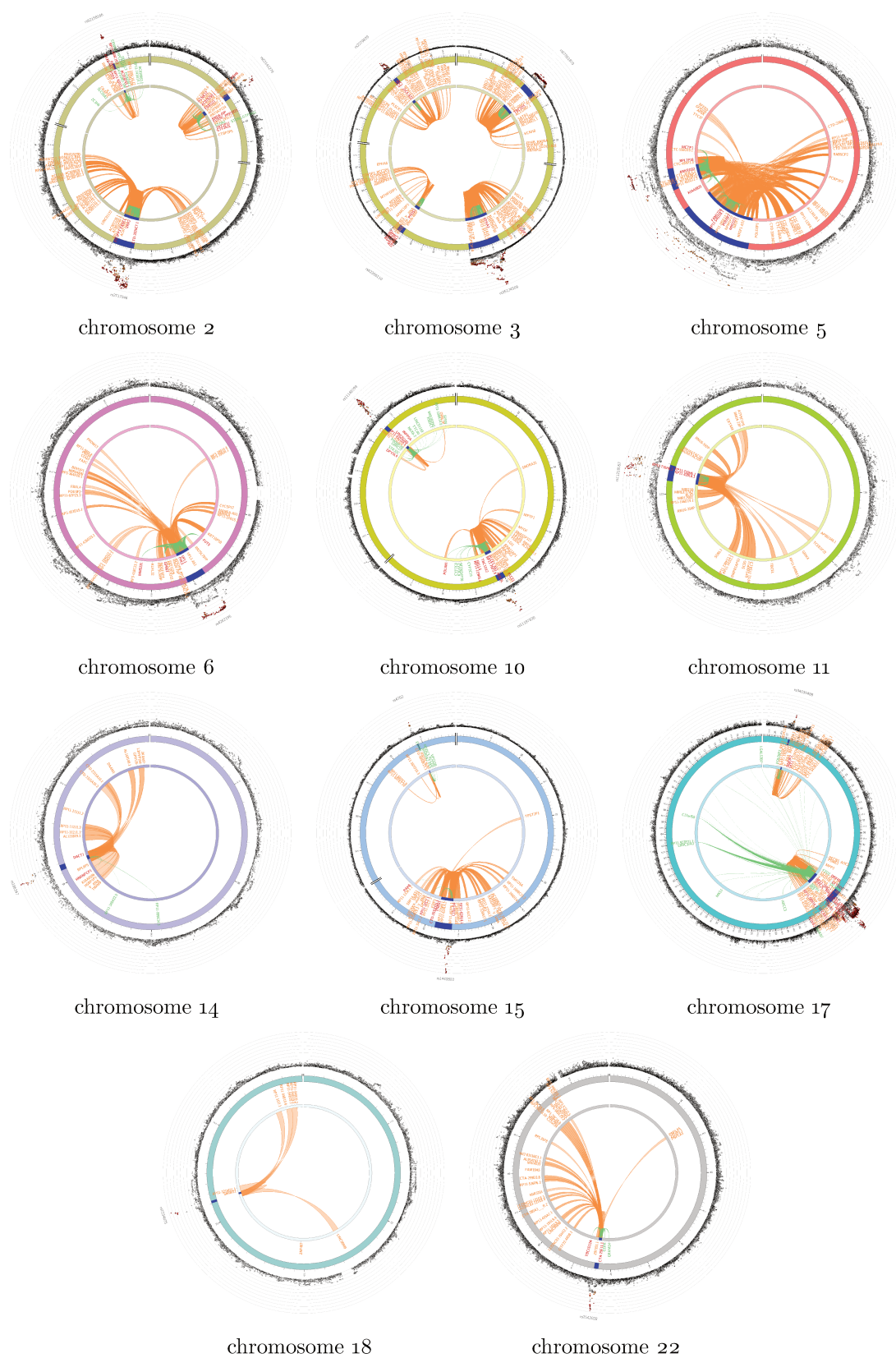


Figure 6.3 – **Genomic loci, eQTL associations and chromatin interactions identified via multivariate GWAS for functional connectivity.** Circos plot representing the genomic risk loci, and the genes associated with the loci by chromatin interactions and eQTLs. From outer layer to inner layer: Manhattan plot. Genomic risk loci are in blue. Genes mapped by chromatin interaction are in orange. Genes mapped by eQTL are in green. Genes mapped by both are in red. Chromatin interaction and eQTLs links follows the same color coding presented above.

GWAS using FCs derived from OcSS set of ROIs. Concerning the FCs derived using OcSS ROIs, we performed a multivariate genome-wide association study with the 14 heritable FCs. This analysis yielded 125 significant SNPs at a genomic threshold, distributed on chromosomes 3, 10, 14. FUMA identified 4 distinct genomic loci distributed on the 3 chromosomes, associated with different aspects of brain language cognitive activities (Fig.6.4, Table 6.2) and represented by 4 lead SNPs.

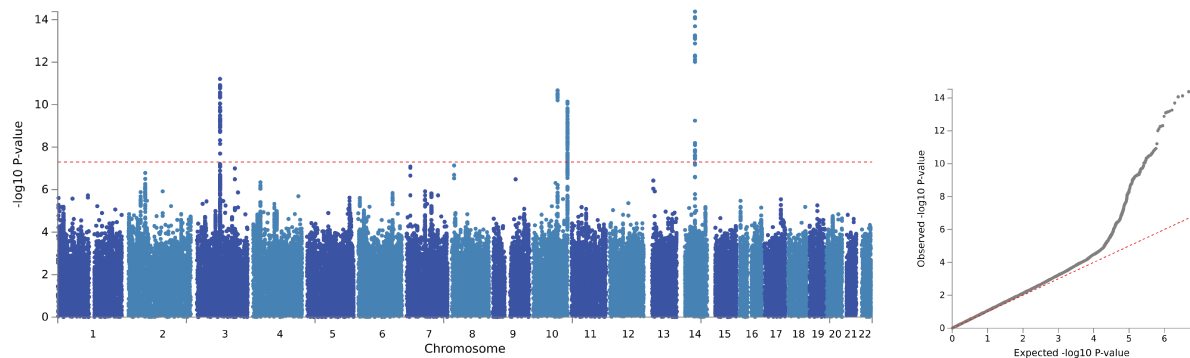


Figure 6.4 – Multivariate GWAS analysis of the OcSS resting state functional connectivity in 32,186 participants. Manhattan plot for multivariate GWAS across 14 FCs. The red dashed line indicates the genome-wide significance threshold $p = 5e-8$. The Quantile-quantile plot is also shown.

Validation of lead SNPs associated with rsfMRI FCs. The multivariate genome-wide association results were replicated in an independent non-British sample considering the nominal significance threshold $p < 0.05$. Following the same pre-processing steps as for the primary sample, the non-British replication sample consists in 4,754 individuals with a mean age of 53 years (± 7.55) and 2,601 female.

Concerning the **-SmPSS-FCs**, three lead SNPs out of twenty were replicated at the nominal significance level ($p < 5e-2$) on multivariate test in the independent non-British replication dataset: $rs1440802$ ($p = 9.58e-3$), $rs35124509$ ($p = 3.25e-3$), $rs11187838$ ($p = 2.92e-2$). Table 6.1 summarises these results. Moreover, these lead SNP showed association at $p < 0.05$ on univariate testing of all but three specific central traits identified in the discovery mvGWAS.

Table 6.1 – Genomic loci associated with heritable language functional connectivities (SmPSS ROIs) using MOSTest. Lead SNP: ID of the lead SNPs within each locus. Position: position of the SNP in the hg19 human reference genome. mvgwasP discovery -British-: MOSTest association P value obtained using the discovery sample. mvgwasP replication -non British-: MOSTest association P value obtained using the independent replication sample. Functional category: Functional consequence of the SNP on the gene obtained from ANNOVAR. 'Central' phenotypes: the phenotypes that contributed most to the multivariate association considering the genome-wide association threshold ($p = 5e - 8$).

Genomic Locus	Lead SNP	Chr	Position	Functional Category	non effect allele	effect allele	MAF	mvgwasP (discovery -British-)	mvgwasP (replication -non British-)	Nearest Gene	'central' phenotypes
1	rs62141276	2	48214217	ncRNA intronic	A	G	0.367	$p = 3.26e-9$	$p = 0.27$	<i>AC079807.4</i>	-
2	rs2717046	2	58041936	intergenic	T	C	0.380	$p = 7.50e-14$	$p = 0.95$	<i>CTD-2026C7.1</i>	-
3	rs62158166	2	114077218	intergenic	C	G	0.223	$p = 8.69e-10$	$p = 0.38$	<i>PAX8</i>	RoIS ↔ PrecR
4	rs67851870	3	17554860	intronic	G	A	0.322	$p = 6.57e-16$	$p = 0.35$	<i>TBC1D5</i>	-
5	rs35124509	3	89521693	exonic	C	T	0.401	$p = 8.95e-59$	$p = 3.25e-3$	<i>EPHA3</i>	AG ↔ F3orb, pSTS ↔ Pole, Pole ↔ T2ml, T1a ↔ STSp, STSp ↔ F3orb, SMG ↔ T3p, AG ↔ STSp, F2p ↔ AG, T2ml ↔ SMG
6	rs62266110	3	93537923	intergenic	A	G	0.319	$p = 1.17e-09$	$p = 0.93$	<i>RNU6-488P</i>	-
7	rs2279829	3	147106319	UTR3	T	C	0.212	$p = 7.57e-21$	$p = 0.68$	<i>ZIC4</i>	T2ml ↔ PrF3op, Prec ↔ F3opd, F2p ↔ STSp, SMG ↔ F2antR,
8	rs145120402	5	93174765	intronic	C	A	0.0433	$p = 1.83e-9$	$p = 0.10$	<i>FAM172A</i>	-
9	5:94068140_AC_A	5	94068140	intronic	A	AC	0.209	$p = 6.79e-9$	$p = 0.30$	<i>ANKRD32:MCTP1</i>	-
10	rs4262195	6	96929475	ncRNA intronic	C	T	0.181	$p = 7.19e-9$	$p = 0.70$	<i>UFL1-AS1</i>	-
11	rs11187838	10	96038686	intronic	A	G	0.435	$p = 4.29e-14$	$p = 2.92e-2$	<i>PLCE1</i>	-
12	rs11146399	10	134308479	intergenic	T	C	0.457	$p = 5.50e-16$	$p = 0.28$	<i>RP11-432J24.5</i>	RoIS ↔ PrecR
13	rs11218557	11	122099839	ncRNA intronic	C	T	0.4579	$p = 1.24e-8$	$p = 0.77$	<i>RP11-820L6.1</i>	-
14	rs186347	14	59072226	intergenic	T	G	0.458	$p = 2.08e-11$	$p = 0.92$	<i>DACT1</i>	Pole ↔ T2ml
15	rs1440802	15	39635124	ncRNA intronic	C	T	0.090	$p = 1e-31$	$p = 9.58e-3$	<i>RP11-624L4.1</i>	Prec ↔ F3opd, PrecR ↔ RoIS
16	rs4702	15	91426560	UTR3	A	G	0.442	$p = 3.77e-13$	$p = 0.42$	<i>FURIN</i>	-
17	rs34039488	17	27320232	intronic	A	G	0.162	$p = 4.74e-8$	$p = 0.46$	<i>PIPOX:SEZ6</i>	-
18	17:44270659_G_A	17	44270659	intronic	A	G	0.399	$p = 5.36e-16$	$p = 0.45$	<i>KANSL1</i>	-
19	rs7234875	18	73114340	intergenic	C	T	0.399	$p = 5.71e-14$	$p = 0.82$	<i>RP11-321M21.3</i>	F3tv ↔ SMG
20	rs2542028	22	47196524	intronic	G	A	0.268	$p = 3.06e-12$	$p = 0.60$	<i>TBC1D22A</i>	-

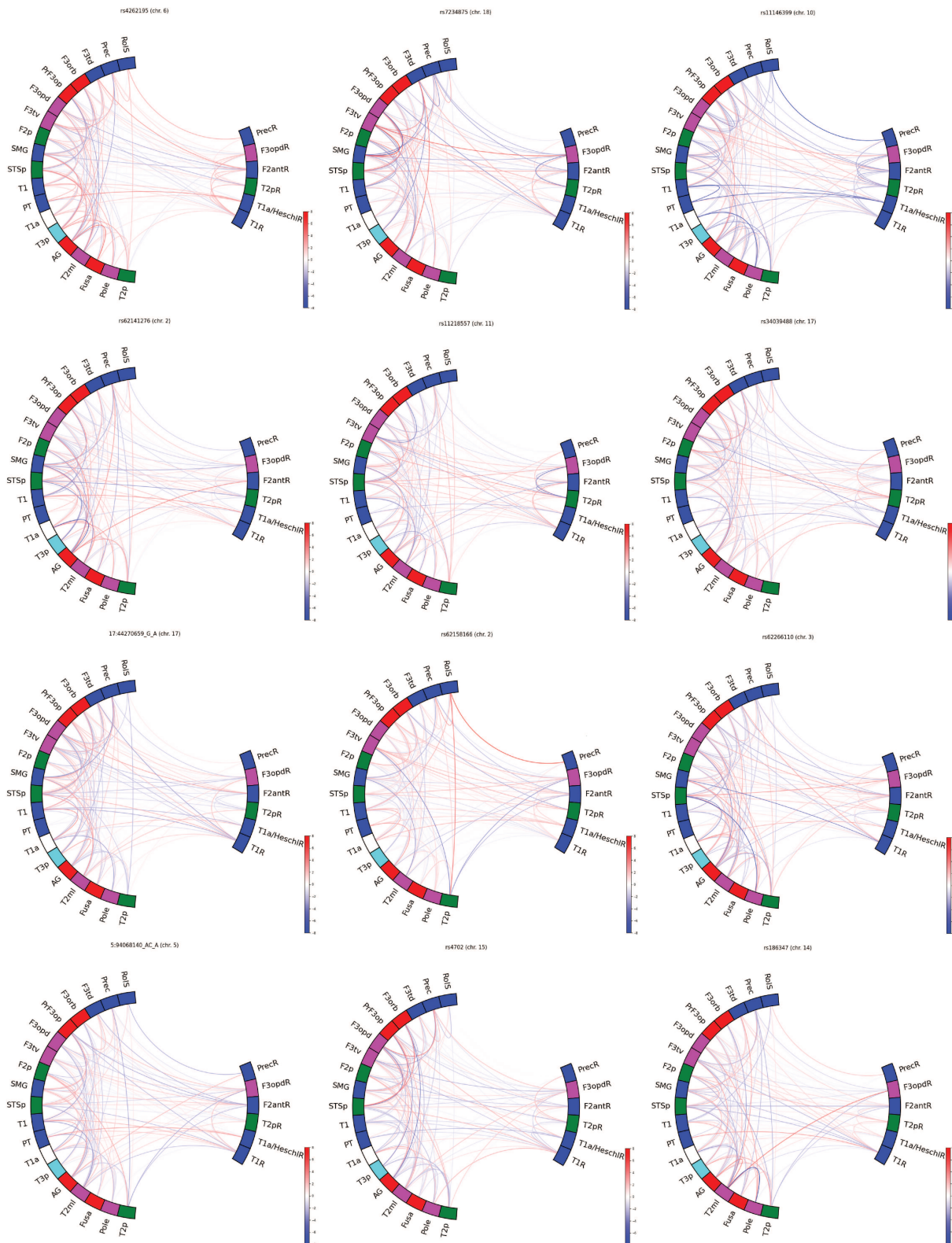
Concerning the **-OcSS-FCs**, one lead SNP was replicated at the nominal significance level ($p < 5e-2$) on multivariate test in the independent non-British replication dataset ($rs7917322$ ($p = 3.16e-2$)) while the lead SNP $rs7650184$ ($p = 6.43e-2$) showed a tentative association in the multivariate level. Table 6.2 summarises these results.

Table 6.2 – **Genomic loci associated with heritable language functional connectivities (OcSS ROIs) using MOSTest.** Lead SNP: ID of the lead SNPs within each locus. Position: position of the SNP in the hg19 human reference genome. mvgwasP discovery -British-: MOSTest association P value obtained using the discovery sample. mvgwasP replication -non British-: MOSTest association P value obtained using the independent replication sample. Functionnal category: Functional consequence of the SNP on the gene obtained from ANNOVAR. 'Central' phenotypes: the phenotypes that contributed most to the multivariate association considering the genome-wide association threshold ($5e - 8$).

Genomic Locus	Lead SNP	Chr	Position	Functional Category	non effect allele	effect allele	MAF	mvgwasP (discovery -British-)	mvgwasP (replication -non British-)	Nearest Gene	'central' phenotypes
1	rs35124509	3	3p11	exonic	C	T	0.4018	6.125755e-12	3.74e-1	<i>EPHA3</i>	-
2	rs11187838	10	10q23.33	intronic	A	G	0.4354	2.122796e-11	5.64e-2	<i>PLCE1</i>	TP↔aSTS
3	rs67221163	10	10q26.3	intergenic	G	C	0.4547	1.841408e-10	4.65e-1	<i>RP11-432J24.5</i>	-
4	rs160459	14	14q23	intergenic	C	A	0.4588	4.155101e-15	3.59e-4	<i>DACT1</i>	TP↔pSTS

The obtained results regarding the multivariate genome-wide association study with the 14 heritable FCs -OcSS- represented by 4 distinct genomic loci distributed on the 3 chromosomes are included in the results obtained from the multivariate genome-wide association study with the 142 heritable FCs -SmPSS-. In the following, for the sake of readability, we will only present the results of the FCs from this set ROIs.

Regional effect. Such a multivariate approach has the advantage of leveraging the distributed nature of genetic effects and the presence of pleiotropy across endophenotypes. Loci respectively identified by MOSTest as associated with several FCs made clear that these SNPs have distributed effects, often with mixed directions, across regions and FCs. Fig. 6.1 shows the FCs associations with each of the twenty lead SNPs loci.



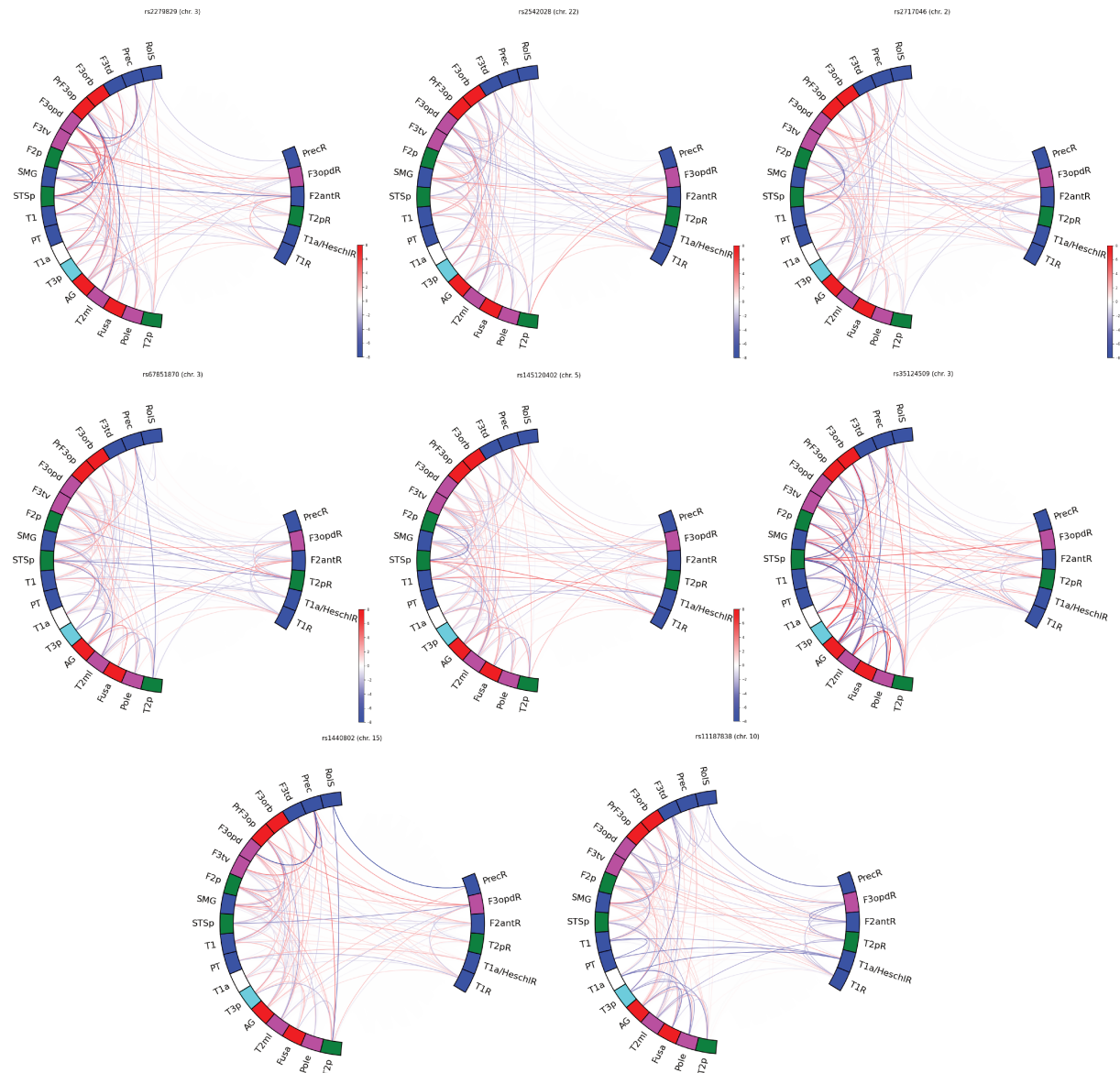


Figure 6.5 – **Regional effects.** Circle plot where all 25 regions are represented (19 ROIs on the left represent the left hemisphere and 6 ROIs on the right represent the right hemisphere) illustrating the lead SNPs identified from the multivariate GWAS for functional connectivity. Each link between two regions represents an endophenotype. The intensity of the color between two regions represents the Z-values from the univariate GWAS for each FCs as defined by the color bar. The absolute Z-values scaling is clipped at 8 ($p = 1.2e-15$). Positive effects of carrying the minor allele are shown in red, and negative in blue.

6.2 Unraveling the genome: Evaluation of results from mvGWAS study of human language

In order to interpret this wealth of results, we proceed as follow:

- First, for each lead SNP, we defined the 'central' endophenotypes that contributed the most in the multivariate association by using the individual univariate summary statistics performed by MOSTest and by considering the genome-wide significance threshold ($p < 5e - 8$) (Supplementary Table 6.1).
- Second, the SNP-based genetic correlation was estimated (using GCTA [Lee et al., 2012] software) for each pair of central FCs associated to each genetic loci.
- Third, we hypothesised that the genetic variants significantly associated with the language FCs could be associated with neuroanatomical tracts that support the information transmission between language areas. Therefore, we tested the potential associations between the hit SNPs with the average values of dMRI relevant white matter tracts.

In what follows, we will review the associations that were replicated which includes these three genomic loci: 15q14, 3p11.1 and 10q23.33.

6.2.1 *Perceptual motor interactions process driven by THBS1 gene*

Identification of 'central' endophenotypes associated with 15q14 genomic risk region.

On 15q14, the lead SNP rs1440802 had two central FCs (that contributed the most in the multivariate association, see 6.2). The minor allele was associated with the partial correlation between:

- (i) The precentral gyrus and the dorsal pars opercularis (Prec↔F3opd). Both connected regions are in the left frontal lobe, and are labelled with a phonological linguistic component (Prec) and multi-labelled with semantic and sentence language processing (F3opd).
- (ii) The (PrecR↔RolS) corresponds to the partial correlation between the precentral gyrus and the Rolandic sulcus. Both regions are identified in the right and left frontal lobes respectively, and are labelled as phonological linguistic component (Fig. 6.6a and Supplementary Table 6.1).

These edges have previously been described in FC studies dedicated to language and more specifically in the perceptual motor interactions [Fridriksson et al., 2009, Nishitani and Hari, 2000, Schwartz et al., 2008, 2012, Turner et al., 2009].

SNP-based genetic correlation of the 'central' endophenotypes.

Both (Prec↔F3opd) and (PrecR↔RolS) 'central' endophenotypes displays a non-significant genetic correlation ($r_G = 16.9 \pm 0.13$, $p = 9.80e - 02$, $N = 32, 186$). Table 6.3 reports genetic correlations for the 'central' endophenotypes identified in the three replicated genomic loci.

Table 6.3 – The SNP-based genetic correlation analysis was estimated using GCTA software, for each pair of central FCs associated to 15q14, 3p11.1 or 10q23.33 genetic loci. Bold phenotypes represents pair of FCs with significant genetic correlation after false discovery rate correction at $p < 0.05$.

Endophenotype 1	Endophenotype 2	P value	rG	SE	FDR P value
T1a↔STSp	Pole↔STSp	2.29E-05	54.8%	0.132	8.47E-04
SMG↔T3p	AG↔STSp	9.78E-05	65.6%	0.192	1.81E-03
STSp↔F3orb	Pole↔STSp	2.21E-04	40.4%	0.111	2.72E-03
AG↔F3orb	T2ml↔SMG	1.01E-03	-46.9%	0.158	9.39E-03
AG↔F3orb	SMG↔T3p	3.76E-03	53.6%	0.211	1.39E-02
Pole↔T2ml	STSp↔F3orb	3.25E-03	-31.4%	0.113	1.39E-02
AG↔F3orb	T1a↔STSp	2.84E-03	-52.5%	0.196	1.39E-02
Pole↔T2ml	AG↔STSp	2.85E-03	36.2%	0.132	1.39E-02
Pole↔T2ml	T2ml↔SMG	2.52E-03	-30.5%	0.109	1.39E-02
Pole↔T2ml	Pole↔STSp	3.53E-03	-28.7%	0.099	1.39E-02
AG↔F3orb	Pole↔STSp	6.34E-03	-36.5%	0.150	1.80E-02
Pole↔T2ml	SMG↔T3p	6.81E-03	35.1%	0.146	1.80E-02
F2p↔AG	AG↔STSp	6.39E-03	54.5%	0.250	1.80E-02
SMG↔T3p	STSp↔F3orb	5.53E-03	-39.6%	0.160	1.80E-02
F2p↔AG	T1a↔STSp	1.07E-02	-51.2%	0.225	2.58E-02
AG↔F3orb	STSp↔F3orb	1.12E-02	-38.2%	0.142	2.58E-02
T1a↔STSp	STSp↔F3orb	1.21E-02	32.9%	0.153	2.63E-02
AG↔F3orb	F2p↔AG	1.39E-02	53.5%	0.243	2.85E-02
Pole↔STSp	T2ml↔SMG	1.68E-02	22.9%	0.109	3.16E-02
T1a↔STSp	SMG↔T3p	1.71E-02	-38.3%	0.187	3.16E-02
F2p↔AG	T2ml↔SMG	2.40E-02	-34.9%	0.180	4.04E-02
SMG↔T3p	T2ml↔SMG	2.33E-02	-29.0%	0.144	4.04E-02
STSp↔F3orb	T2ml↔SMG	3.98E-02	20.7%	0.120	6.40E-02
AG↔STSp	T2ml↔SMG	4.84E-02	-22.3%	0.137	7.46E-02
T1a↔STSp	T2ml↔SMG	5.60E-02	21.9%	0.137	8.28E-02
SMG↔T3p	Pole↔STSp	6.05E-02	-21.8%	0.142	8.61E-02
AG↔STSp	Pole↔STSp	7.06E-02	-19.3%	0.127	9.68E-02
AG↔STSp	STSp↔F3orb	8.66E-02	-19.6%	0.139	1.14E-01
F2p↔AG	SMG↔T3p	9.02E-02	31.1%	0.232	1.15E-01
Prec↔F3opd	RoIS↔PrecR	9.80E-02	16.9%	0.131	1.21E-01
Pole↔T2ml	T1a↔STSp	1.08E-01	-16.6%	0.135	1.29E-01
AG↔F3orb	AG↔STSp	1.91E-01	16.0%	0.184	2.21E-01
F2p↔AG	Pole↔STSp	2.08E-01	14.0%	0.176	2.33E-01
Pole↔T2ml	F2p↔AG	2.15E-01	-13.7%	0.176	2.34E-01
T1a↔STSp	AG↔STSp	2.22E-01	-12.8%	0.164	2.34E-01
F2p↔AG	STSp↔F3orb	3.88E-01	5.4%	0.189	3.99E-01
Pole↔T2ml	AG↔F3orb	4.56E-01	1.6%	0.148	4.56E-01

Validation of lead SNPs using diffusion imaging derived endophenotypes. we tested the potential associations between the hit SNPs with the average values of dMRI relevant white matter tracts; 3 white matter tracts to be tested with locus on 15q14: the white matter tracts linking the regions of the (Prec↔F3opd) consists of the *i*) arcuate anterior segment fasciculus

(AF) dorsal pathway [Catani et al., 2005], *ii*) the frontal aslant tract (FAT) which is reported as connecting Broca's region (BA44/45) with dorsal medial frontal areas including supplementary and pre-supplementary motor area (BA6) [Catani and Forkel, 2019, Rojkova et al., 2016] while the anatomical connectivity underlying the (PrecR \leftrightarrow RolS) FC endophenotype consists of *iii*) the corpus callosum which interconnects both hemispheres.

The MO measured in the FAT and the OD measured in the anterior segment of AF are associated with the rs1440802 SNP with $p = 3.33e-6$ and $p = 2.47e-65$, respectively. The corpus callosum exhibits no significant association (See Table 6.4).

Table 6.4 – Univariate associations of 2 lead SNPs with diffusion MRI indices. Bold phenotypes represents tracts significantly associated with the lead SNP after Bonferroni correction ($p = 6.94e - 4(0.05/(3 * 9 + 5 * 9))$)

CHR	SNP	Endophenotype	Metric	Beta	P	Z		
15	rs1440802	Corpus callosum	L3	-0.034427736	0.0128030706756605	-2.4892006		
		Corpus callosum	ISOVF	-0.03349917	0.015433186438535	-2.422051		
		Corpus callosum	FA	0.032574	0.0185153496000616	2.3551476		
		Corpus callosum	MD	-0.0320895	0.0203348456808069	-2.3201115		
		Corpus callosum	L2	-0.03150293	0.0227447878217582	-2.2776945		
		Corpus callosum	L1	-0.024891464	0.0719201471926044	-1.7996233		
		Corpus callosum	ICVF	0.020728022	0.1339805874968674	1.4985878		
		Corpus callosum	MO	-0.007192162	0.6030911735857749	-0.5199604		
		Corpus callosum	OD	-0.0071261316	0.6064225847816228	-0.51518667		
		Frontal Aslant Tract Left	MO	-0.064280435	3.339583634083396e-06	-4.6487465		
		Frontal Aslant Tract Left	OD	0.05600727	5.119567913113231e-05	4.0501003		
		Frontal Aslant Tract Left	L1	-0.05497112	7.033938936284984e-05	-3.9751348		
		Frontal Aslant Tract Left	ISOVF	-0.03451673	0.0125731467908473	-2.4956362		
		Frontal Aslant Tract Left	MD	-0.030007247	0.030041806799683	-2.1695387		
		Frontal Aslant Tract Left	L3	-0.019929217	0.1496320442788888	-1.4408324		
		Frontal Aslant Tract Left	ICVF	0.009302874	0.5012290542417216	0.6725572		
		Frontal Aslant Tract Left	FA	-0.00893621	0.5182478779915272	-0.6460486		
		Frontal Aslant Tract Left	L2	-0.0038335118	0.7816691369948107	-0.27714452		
		Arcuate Anterior Segment Left	OD	-0.23503968	2.4731891328416e-65	-17.070208		
		Arcuate Anterior Segment Left	MO	0.20940934	3.854779219502659e-52	15.194348		
		Arcuate Anterior Segment Left	FA	0.14841966	5.953674748880915e-27	10.749573		
		Arcuate Anterior Segment Left	L2	-0.14630064	3.124967969997189e-26	-10.595551		
		Arcuate Anterior Segment Left	L1	0.10824838	4.748243823784488e-15	7.833407		
		Arcuate Anterior Segment Left	L3	-0.048699692	0.000429210896869	-3.5214393		
		Arcuate Anterior Segment Left	ISOVF	-0.04646888	0.000779221912092	-3.360072		
		Arcuate Anterior Segment Left	MD	-0.03227475	0.0196215356697533	-2.3335073		
		Arcuate Anterior Segment Left	ICVF	0.01887147	0.1724557985221415	1.3643552		
		3	rs35124509	Inferior Fronto Occipital fasciculus Left	MO	0.041906517	2.497821751857874e-07	5.1578646
				Inferior Fronto Occipital fasciculus Left	OD	-0.032216765	7.352872189541101e-05	-3.9645689
				Inferior Fronto Occipital fasciculus Left	L1	0.020603567	0.0112419137853251	2.5350878
				Inferior Fronto Occipital fasciculus Left	FA	0.011022241	0.1750699434483398	1.3560919
				Inferior Fronto Occipital fasciculus Left	MD	0.008882381	0.2744777383394011	1.0928088
				Inferior Fronto Occipital fasciculus Left	ISOVF	0.004858134	0.5500444614237823	0.5976935
				Inferior Fronto Occipital fasciculus Left	L2	-0.0047178925	0.5616182288559746	-0.58043957
				Inferior Fronto Occipital fasciculus Left	L3	0.004688471	0.5640612077910263	0.5768198
				Inferior Fronto Occipital fasciculus Left	ICVF	-0.0037015	0.6488269473186558	-0.45539242
				Uncinate Left	MO	0.04215374	2.1220118223943872e-07	5.1883187
				Uncinate Left	OD	-0.023885116	0.003293203051305	-2.9389555
				Uncinate Left	L1	0.022804746	0.0050162697429167	2.8059871
				Uncinate Left	FA	0.0124250045	0.1263415538835854	1.528689
				Uncinate Left	MD	0.009754712	0.2300860378190222	1.2001373
				Uncinate Left	ICVF	-0.008374233	0.3028745596840176	-1.0302886
				Uncinate Left	L3	0.0072252634	0.3740427072882268	0.88892627
				Uncinate Left	L2	-0.0059861178	0.4614441477476281	-0.73647094
				Uncinate Left	ISOVF	-0.00015708078	0.9845814785319852	-0.019325454
				Arcuate Anterior Segment Left	MO	-0.023604663	0.0036791304163527	-2.904438
				Arcuate Anterior Segment Left	L2	0.02263006	0.0053612311006976	2.784488
				Arcuate Anterior Segment Left	FA	-0.016705215	0.0398443764601674	-2.0553586
Arcuate Anterior Segment Left	ISOVF			0.012345138	0.1287972035080454	1.5188621		
Arcuate Anterior Segment Left	MD			0.010159739	0.2113101729791165	1.2499709		
Arcuate Anterior Segment Left	ICVF			-0.010145878	0.2119338859324724	-1.2482654		
Arcuate Anterior Segment Left	L3			0.007130031	0.3803728653956867	0.8772095		
Arcuate Anterior Segment Left	OD			0.0058026165	0.4752924638259838	0.7138944		
Arcuate Anterior Segment Left	L1			-0.00077840226	0.9237065215197204	-0.095765874		
Arcuate Long Segment Left	OD			-0.036813475	5.882222834891848e-06	-4.530579		
Arcuate Long Segment Left	L1			0.020133851	0.0132387238045088	2.477282		
Arcuate Long Segment Left	MO			0.014724737	0.0700387951827605	1.8116597		
Arcuate Long Segment Left	MD			0.012832667	0.1143707059249424	1.578849		
Arcuate Long Segment Left	ISOVF			0.012217956	0.1327840217022739	1.5032134		
Arcuate Long Segment Left	L3			0.007494609	0.3564947860978851	0.9220648		
Arcuate Long Segment Left	ICVF			-0.006647461	0.4134499954108012	-0.8178375		
Arcuate Long Segment Left	FA			0.003683449	0.6504251731527789	0.4531716		
Arcuate Long Segment Left	L2			0.0017721133	0.827412710308558	0.21802105		

CHR	SNP	Endophenotype	Metric	Beta	P	Z
		Arcuate Posterior Segment Left	L3	0.041222345	3.903850074370563e-07	5.0735874
		Arcuate Posterior Segment Left	ICVF	-0.038005184	2.90620129195227e-06	-4.67734
		Arcuate Posterior Segment Left	MD	0.036748115	6.110479990305916e-06	4.52253
		Arcuate Posterior Segment Left	FA	-0.03642834	7.3549024549896465e-06	-4.4831505
		Arcuate Posterior Segment Left	L2	0.030796966	0.0001507877719798	3.7897685
		Arcuate Posterior Segment Left	L1	0.021983918	0.0068311464928437	2.7049656
		Arcuate Posterior Segment Left	MO	0.016925715	0.0372975642974694	2.0824919
		Arcuate Posterior Segment Left	ISOVF	0.014387828	0.0766931917480869	1.7702038
		Arcuate Posterior Segment Left	OD	0.00464435	0.5677342741085722	0.5713915

Functional annotation: Locus in 15q14 associated to (Prec↔F3opd) and (PrecR↔RoIS) endophenotypes. Four independent SNPs were identified in locus 15q14 (rs1440802, rs11629938, rs773225188, rs34680120) (Fig. 6.6c). Regarding eQTL annotations, we explored tissue-specific gene expression resources, including both brain tissues and blood - considered as a good proxy when brain tissues are not available [Qi et al., 2018]. Significant results were obtained:

- The four independent SNPs are cis-eQTL of *THBS1* gene in eQTLGen, BIOSQTL and GTEx/v8. Additionally, rs34680120 is eQTL of *RP11-37C7.1* gene ($p_{adj} < 1.02e-3$) in PsychENCODE and eQTL of *CTD-2033D15.1* gene ($p_{adj} < 6.0e-6$) in BIOSQTL; see Fig.6.6c. Overall, the variants of this genomic risk region are found 72 times as eQTL of genes from different data sources.
- Based on the human gene expression data from the Brainspan database, we found that *THBS1* gene has relatively high mRNA expression during early mid-prenatal to late prenatal stages, from 16 to 37 post-conceptual weeks; see Fig. 6.6e.
- Indirect predictions might be added from the following annotation. For instance, *RASGRP1*, identified by chromatin interaction mapping and which also appears to be under control of temporal expression during neurodevelopment, is reported as over-expressed in the perisylvian language areas [Johnson et al., 2009] and as up-regulated in the dorsal striatum [Cirnaru et al., 2020].

Fig. 6.6 summarises these results, found by mvGWAS, associated to (Prec↔F3opd) and (PrecR↔RoIS) FC endophenotypes. These pinpoint *THBS1* as the possible gene underlying this association signal.

Locus regulating *THBS1* associated with the perceptual motor interactions process

In our results, a locus in 15q14 was associated with the precentral-opercularis FC (Prec↔F3opd) and the precentral-Rolandic FC endophenotypes (PrecR↔RoIS). In the literature, the L|R Prec regions in the ventral precentral gyrus are both associated with phonology language component and considered relevant for pharynx and tongue fine-movement coordination in the human and non-human primates [Belyk and Brown, 2017, Kumar et al., 2016, Vigneau et al., 2006]. RoIS in the dorsal Rolandic sulcus is attributed to the phonology component and matches the mouth primary motor area but also the perception of syllables [Fadiga et al., 2002, Vigneau et al., 2006, Wilson et al., 2004]. F3opd in the dorsal pars opercularis (BA44/45) is associated with

semantic/sentence processing. The motor theory of speech perception has been quite an old debate [Flinker et al., 2015, Galantucci et al., 2006, Liberman and Mattingly, 1985, Schwartz et al., 2008, Whalen, 2019]. In this study, we report a locus in 15q14 (lead SNP rs1440802) associated with both this FC between the motor and Broca’s areas and the frontal aslant tract connecting directly (pre)supplementary motor area with the opercular part of inferior frontal gyrus [Catani et al., 2012, Vergani et al., 2014], in line with this perception–motor link.

SNPs in high LD with rs1440802 in the genomic region have already been linked to several other structural features (surface area and cortical thickness) including primary motor cortex, primary somatosensory cortex [Elliott et al., 2018, van der Meer et al., 2020], supramarginal, and pars opercularis [van der Meer et al., 2020], supporting a common genetic influence of the sensory-motor interaction.

The lead SNP rs1440802 and SNPs in LD that we have shown to be associated with both (Prec↔F3opd) and (PrecR↔RoIS) are found to be eQTL of *THBS1* gene in the blood with high confidence. The thrombospondin-1 protein encoded by *THBS1* gene is a member of the thrombospondin family, a glycoprotein expressed in the extracellular matrix. It has been implicated in synaptogenesis [Christopherson et al., 2005] and regulates the differentiation and proliferation of neural progenitor cells [Lu and Kipnis, 2010], and has been involved in human neocortical evolution [Cáceres et al., 2003, 2007]. Other members of the thrombospondin’s family, *THBS2* and *THBS4*, have been shown to be over-expressed in the adult human cerebral cortex compared to chimpanzees and macaques [Cáceres et al., 2007]. Their increased expression suggests that human brain might display distinctive features involving enhanced synaptic plasticity in adulthood which may contribute to cognitive and linguistic abilities [Sherwood et al., 2008]. From a developmental point of view, *THBS1* appears to be under control of temporal expression during development, as revealed by BrainSpan data (See Fig. 6.6e. *THBS1* expression was studied from the longitudinal transcriptomic profile resource of the developing human brain (18, 19, 21, 23 weeks of gestation) [Johnson et al., 2009]. Its expression is reported as over-expressed in the neocortex, including the perisylvian language areas, compared to phylogenetically older parts of the brain such as the striatum, thalamus and cerebellum [Johnson et al., 2009]. Thrombospondin-1 have also been linked to Autism spectrum disorder [Lu et al., 2014], Alzheimer’s disease [Ko et al., 2015], and Schizophrenia [Park et al., 2012].

Taken together, these results indicate that *THBS1*, modulated by a lead SNP in the 15q14 locus, could be prioritised in the study of key genes playing a role in the functional connectivity part of the *perceptual motor* interaction required for language, and with the anatomical connectivity that supports their interactions.

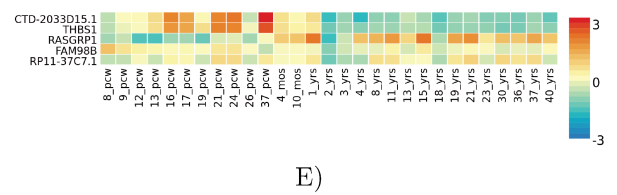
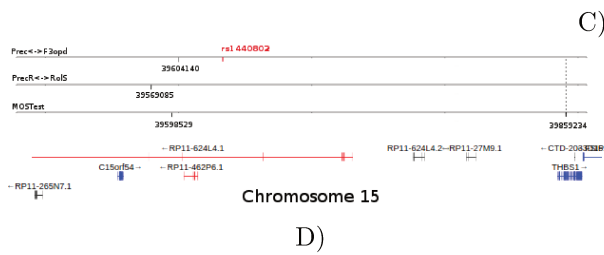
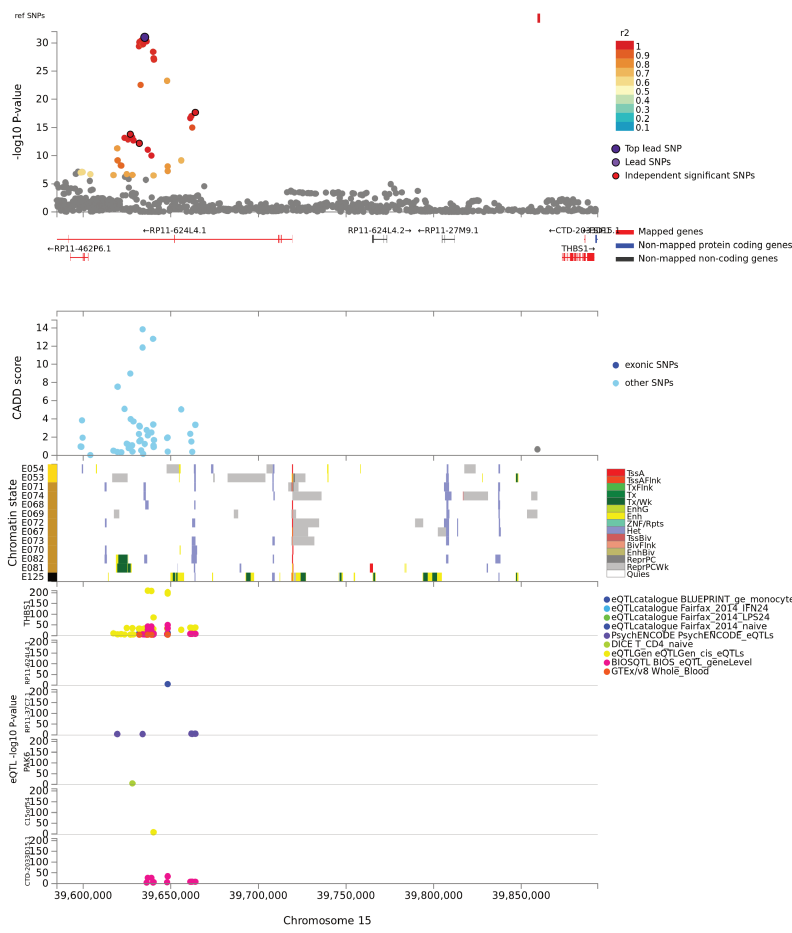
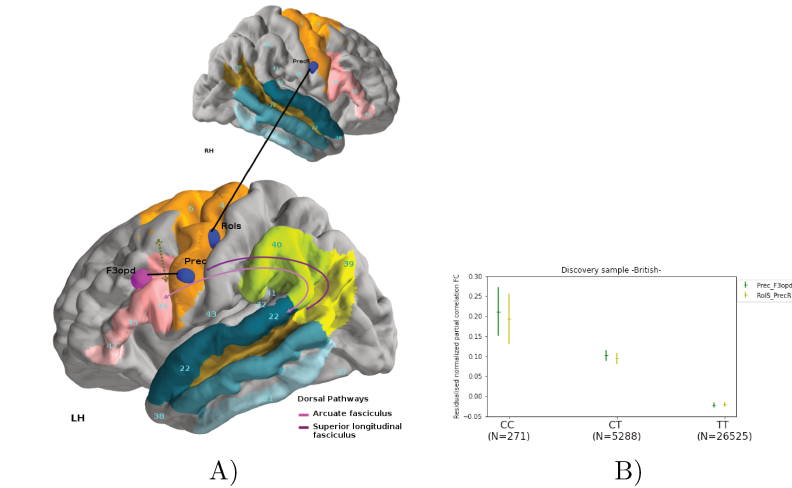


Figure 6.6 – Main results for the 15q14 locus. A) The two pairs of ROIs that forms the endpoints of the associated FCs reported as black bold lines. B) Effect sizes of the SNP rs1440802 for the two connections: (Prec↔F3opd) FC in green and (PrecR↔RolS) FC in yellow. C) Locus Zoom of the genomic region identified by the mvGWAS. Chromatin state of the genomic region. Brain tissue name abbreviations are the following; E054: Ganglion Eminence derived primary cultured neurospheres, E053: Cortex derived primary cultured neurospheres, E071: Brain Hippocampus Middle, E074: Brain Substantia Nigra, E068: Brain Anterior Caudate, E069: Brain Cingulate Gyrus, E072: Brain Inferior Temporal Lobe, E067: Brain Angular Gyrus, E073: Brain Dorsolateral Prefrontal Cortex, E070: Brain Germinal Matrix, E082: Fetal Brain Female, E081: Fetal Brain Male, E125: NH-A Astrocytes Primary Cells. The state abbreviations are the following; TssA: active transcription start site (TSS), TssFlnk: Flanking Active TSS, TxFlnk: Transcription at gene 5' and 3', Tx: Strong transcription, TxWk: Weak transcription, EnhG: Genic enhancers, Enh: Enhancers, ZNF/Rpts: ZNF genes & repeats, Het: Heterochromatin, TssBiv: Bivalent/Poised TSS, BivFlnk: Flanking Bivalent TSS/Enh, EnhBiv: Bivalent Enhancer, ReprPC: Repressed PolyComb, ReprPCWk: Weak Repressed PolyComb, Quies: Quiescent/Low. Expression quantitative trait loci (eQTL) associations (data source: eQTLGen [Võsa et al., 2018], PsychENCODE [Wang et al., 2018], DICE [Schmiedel et al., 2018], BIOS QTL browser [Zhernakova et al., 2017], GTEx/v8 [Consortium et al., 2017], eQTLcatalogue). D) Overlap of the genomic region risk region identified from FUMA for MOSTest results, (Prec↔F3opd) and (PrecR↔RolS). E) Gene expression from BrainSpan for the interesting genes prioritised by FUMA.

6.2.2 Fronto-temporal semantic network driven by *EPHA3* gene

Identification of 'central' endophenotypes associated with 3p11.1 genomic risk region.

On 3p11.1, the lead SNP rs35124509 had nine central FCs (that contributed the most in the multivariate association, see 6.2). The minor allele was associated with the partial correlation between:

- the left posterior part of the superior temporal sulcus and the left temporal pole (Pole↔STSp),
- the left temporal pole and the lateral/middle part of the middle temporal gyrus (Pole↔T2ml),
- the angular gyrus and the pars orbitalis of the left inferior frontal gyrus (AG↔F3orb),
- the anterior part of the Superior temporal gyrus and the left posterior part of the superior temporal sulcus (T1a↔STSp),
- the left posterior part of the superior temporal sulcus and the pars orbitalis of the left inferior frontal gyrus (STSp↔F3orb),
- the supramarginal gyrus and the posterior part of the left inferior temporal gyrus (SMG↔T3p),
- the angular gyrus and the left posterior part of the superior temporal sulcus (AG↔STSp),
- the Posterior part of the middle frontal gyrus and the angular gyrus (F2p↔AG),
- the lateral/middle part of the middle temporal gyrus and the supramarginal gyrus (T2ml↔SMG)

See also Fig. 6.7a and Supplementary Table 6.1. These connected regions are located across the left parieto-fronto-temporal lobe, and are mainly labelled as semantic language processing. These edges have previously been described in FC studies dedicated to language and especially to the semantic component. This component typically includes the inferior frontal gyrus, the left temporal cortex (i.e. temporal pole, middle temporal gyrus, fusiform gyrus) and the left angular gyrus [Binder et al., 2009, Jackson et al., 2016, Vigneau et al., 2006]. At the univariate level, these loci associated to central endophenotypes display an important overlap; See Fig 6.7d.

SNP-based genetic correlation of the 'central' endophenotypes.

The SNP-based genetic correlation was estimated for each pair of central FCs associated to 3p11.1 genetic loci, indicating overlapping genetic contributions among several FCs (Table 6.3). A negative genetic correlation between some FCs has been observed which indicates that variants can have antagonistic effects on the co-activations of these regions.

Validation of lead SNPs using diffusion imaging derived endophenotypes.

We tested the potential associations between the hit SNPs with the average values of dMRI relevant white matter tracts; 5 white matter tracts, to be tested with locus on 3p11.1, were selected according to the nine central endophenotypes associated with the 3p11.1 locus. The anatomical connectivity underlying these connections consists of the *i*) inferior fronto-occipital fasciculus (IFOF) which connects the inferior frontal regions with the temporal and occipital cortex [Forkel et al., 2014], *ii*) uncinate fasciculus (UF) which is reported to connect the anterior temporal lobe to the orbital region and part of the inferior frontal [Catani and De Schotten, 2008, Catani and Forkel, 2019, Friederici, 2017, Vigneau et al., 2006], and *iii*) arcuate long segment fasciculus (AF), *iv*) anterior segment of arcuate fasciculus, *v*) posterior segment of arcuate fasciculus [Catani et al., 2005].

The MO measured in the IFOF and UF is associated with the rs35124509 SNPs with $p = 2.49e-7$ and $p = 2.12e-7$, respectively. Both long and posterior segment of AF are associated with rs35124509 SNPs with $p = 5.88e-6$ (OD) and $p = 3.90e-7$ (L3), while the anterior segment of AF exhibits no significant association (See Table 6.4).

Functional annotation: Locus in 3p11.1 associated to semantic-language related endophenotypes.

Fourteen independent SNPs were identified in locus 3p11.1 (Fig. 6.7c).

- The rs35124509 SNP is a non-synonymous variant within exon 16 of *EPHA3* protein-coding gene. The subregion around rs35124509 and rs113141104 has its chromatin state annotated as (weak) actively-transcribed states (Tx, TxWk) in the brain tissues, specifically in the Brain Germinal Matrix, the Ganglion Eminence derived primary cultured neurospheres, and in the Fetal Brain Female.
- Concerning the subregion around rs6551410, it has its chromatin state annotated as Weak transcription (TxWk) in the Fetal Brain Female, enhancer (enh) in the Brain Germinal Matrix and Repressed PolyComb (ReprPC) in both the Ganglion Eminence and Cortex derived primary cultured neurospheres. Additionally, the subregion around rs6551407 has its chromatin state annotated as Weak transcription (TxWk) in the Brain Germinal Matrix, Fetal Brain Male and Fetal Brain Male. Overall, this reveals a genomic region involved in fine regulation mechanisms of brain development.
- Considering the rs35124509 SNP and variants in linkage disequilibrium (LD) with it in the genomic risk region, we scrutinised CADD and RDB scores, precise genomic positions and risk prediction, and we noticed some remarkable SNPs. We observed two exonic variants: *i*) The SNP rs1054750 ($LD_{rs35124509} r^2 > 0.99$, $p_{mvGWAS} = 6.65e - 34$) is a synonymous variant within exon 16 of *EPHA3*. *ii*) the already mentioned non-synonymous lead SNP

rs35124509 ($p_{mvGWAS} = 8.95e - 59$), the minor allele results in a substitution in the protein from tryptophan (W) residue (large size and aromatic) into an arginine (R) (large size and basic) at position 924 (W924R, p.Trp924Arg) in the Sterile Alpha Motif (SAM) domain. This SNP is not predicted to alter protein function (Polyphen-2 = "benign") but is predicted to be potentially a regulatory element by several tools (RDB score = $3a$, CADD = 22.3 - when $CADD_{thresh} = 12.37$ for deleterious effect as suggested by Kircher et al. [2014])

- Moreover, we observed eight SNP (rs28623022, rs7650184, rs7650466, rs73139147, rs3762717, rs73139144, rs73139148, rs566480002) ($LD_{rs35124509} r^2 > 0.73$, $p_{mvGWAS} < 4.46e - 20$) located in 3'-UTR of *EPHA3* which could affect its expression by modulating miRNA binding [Popp et al., 2016].
- The hit-SNP rs35124509 and the rest of highlighted SNPs act as eQTL for *EPHA3* in different tissues including brain cerebellum ($p_{FDR} < 5e-2$ in GTEx/v8 data source).

Fig. 6.7 summarises the functional annotations in *3p11.1* associated to multiple FC endophenotypes in semantic component of language. These functional characterization supports *EPHA3* as a possible gene with a key role in language development in humans.

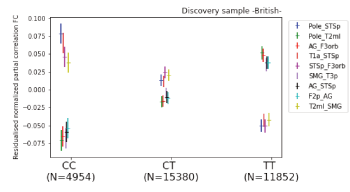
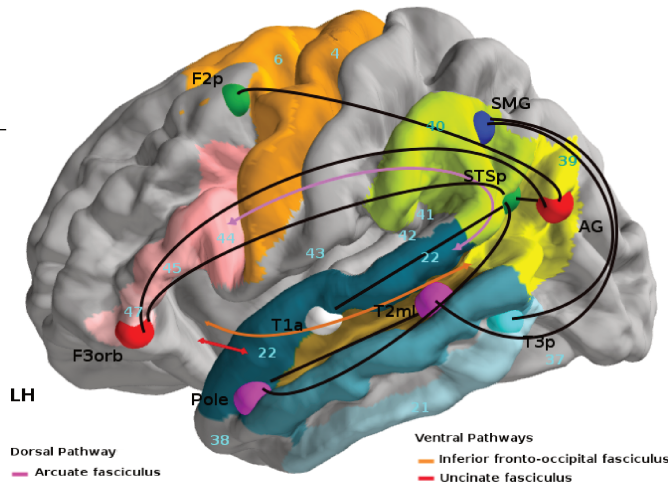
Locus in *EPHA3* associated with the fronto-temporal semantic network In our results, a locus in *3p11.1* is found associated with nine fronto-parietal-temporal endophenotypes. In the literature, The angular gyrus (AG) has been shown to activate during functional imaging tasks probing semantics and involved in conceptual knowledge [Vigneau et al., 2006]. F3orb in the pars orbitalis in the inferior frontal gyrus is labelled semantic for its involvement in semantic retrieval in spoken and sign language [Rönnberg et al., 2004]. It has also been associated with categorisation, association, and word generation tasks [Booth et al., 2002, Gurd et al., 2002, Noppeney and Price, 2004]. The temporal pole region, located in the anterior temporal lobe, is associated with semantic and sentence processing [Vigneau et al., 2006] and the posterior superior temporal sulcus (pSTS) is reported to be implicated in syntactic complexity [Constable et al., 2004] but also process the semantic integration of complex linguistic material [Vigneau et al., 2006]. Both pSTS and the angular gyrus overlap with the Geschwind's territory (See Fig. 2.2). The lateral/middle part of the middle temporal gyrus is devoted to verbal knowledge [Vigneau et al., 2006]. These regions and their corresponding endophenotypes fit rather well with the *fronto-temporal semantic system* described in [Vigneau et al., 2006] facilitating the association of integrated input messages with internal knowledge. The anterior part of the superior temporal gyrus and the posterior part of the inferior temporal gyrus are phonological-semantic interface areas processing. [Vigneau et al., 2006] propose that these ones are transitional zones between the perception and semantic integration of language stimuli and are crucial during the development of language.

SNPs of this genomic region in high LD with the lead SNP rs35124509 have already been found associated with: rsfMRI ICA functional connectivity (edge 387, 383, 399, and ICA-features 3); see [Elliott et al., 2018]. The ICA maps used for these FC estimations partially-overlap

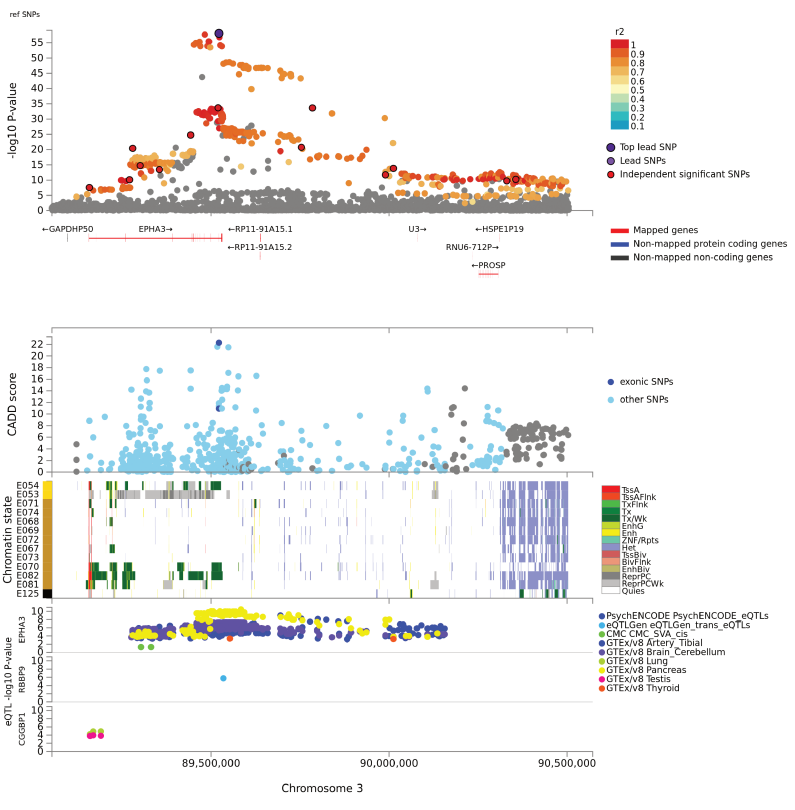
semantic language areas including the angular gyrus, the most anterior part of the STS, the anterior fusiform gyrus, the lateral-middle part of T2, the ventral part of the pars triangularis and the pars orbitalis of the left inferior frontal gyrus. Regarding cognitive traits, this locus was associated to intelligence [Savage et al., 2018]. Finally, other SNPs, in strong LD with the lead SNP rs35124509, consistently act as an eQTL of *EPHA3* in brain tissues.

The ephrin type-A receptor 3 protein encoded by *EPHA3* gene belongs to the ephrin receptor family that can bind the ephrins subfamily of the tyrosine kinase protein family. EPH receptors and their ligands were found to play important roles in multiple developmental processes, including tissue morphogenesis, embryogenesis, neurogenesis, vascular network formation, neural crest cell migration, axon fasciculation, axon guidance, and topographic neural map formation [Gerstmann and Zimmer, 2018, Gibson and Ma, 2011, Pasquale, 2008]. *EPHA3* binds predominantly EFNA5 and plays a role in the segregation of motor and sensory axons during neuromuscular circuit development [Lawrenson et al., 2002]. In [Johnson et al., 2009], *EPHA3* is reported as over-expressed in the fetal rhesus macaque monkey neocortex (NCTX) and especially in the occipital lobe compared to the other NCTX areas. Noticeably, its ligand *EFNA5* is over-expressed in perisylvian areas and is located in a human accelerated conserved non-coding sequence (haCNS704) [Johnson et al., 2009]. EPH receptors have been linked to neurodevelopmental disorders, including schizophrenia [Zhang et al., 2010] and autism spectrum disorder [Casey et al., 2012]. Moreover, in [Rudov et al., 2013], *EPHA3* is found *in silico*, as putative gene implicated in dyspraxia, dyslexia and specific language impairment (SLI). Finally, we observed that *EPHA3* is expressed in the human brain, in a consistent manner across developmental stages from early prenatal to late-mid prenatal (8-24 pcw, BrainSpan ; see Fig. 6.7e.

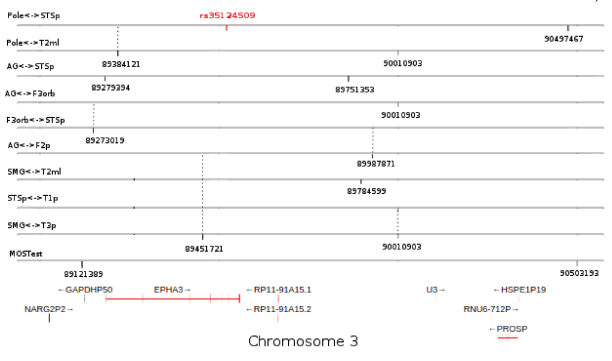
Taken together, these results indicate that *EPHA3* in the 3p11.1 locus, could be prioritised in the study of key genes playing a role in the *fronto-temporal semantic* network, and with the anatomical connectivity support of this network.



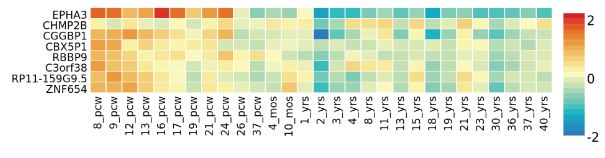
A) B)



C)



D)



E)

Figure 6.7 – **Main results for the 3p11.1 locus.** A) The pairs of ROIs that forms the endpoints of the associated FCs reported as black bold lines. B) Effect sizes of the SNP rs35124509 for the nine connections: (AG↔F3orb), (Pole↔STSp), (Pole↔T2ml), (T1a↔STSp), (STSp↔F3orb), (SMG↔T3p), (AG↔STSp), (F2p↔AG) and (T2ml↔SMG) FCs. C) Locus Zoom of the genomic region identified by the mvGWAS. Chromatin state of the genomic region. Brain tissue name abbreviations are the following; Eo54:Ganglion Eminence derived primary cultured neurospheres, Eo53: Cortex derived primary cultured neurospheres, Eo71: Brain Hippocampus Middle, Eo74: Brain Substantia Nigra, Eo68: Brain Anterior Caudate, Eo69: Brain Cingulate Gyrus, Eo72: Brain Inferior Temporal Lobe, Eo67:Brain Angular Gyrus, Eo73: Brain Dorsolateral Prefrontal Cortex, Eo70: Brain Germinal Matrix, Eo82: Fetal Brain Female, Eo81: Fetal Brain Male, E125: NH-A Astrocytes Primary Cells. The state abbreviations are the following; TssA: active transcription start site (TSS), TssFlnk: Flanking Active TSS, TxFlnk: Transcription at gene 5' and 3', Tx: Strong transcription, TxWk: Weak transcription, EnhG: Genic enhancers, Enh: Enhancers, ZNF/Rpts: ZNF genes & repeats, Het: Heterochromatin, TssBiv: Bivalent/Poised TSS, BivFlnk: Flanking Bivalent TSS/Enh, EnhBiv: Bivalent Enhancer, ReprPC: Repressed PolyComb, ReprPCWk: Weak Repressed PolyComb, Quies: Quiescent/Low. Expression quantitative trait loci (eQTL) associations (data source: eQTLGen [Vösa et al., 2018], PsychENCODE[Wang et al., 2018], DICE [Schmiedel et al., 2018], BIOS QTL browser [Zhernakova et al., 2017], GTEx/v8 [Consortium et al., 2017], eQTLcatalogue). D) Overlap of the genomic region risk region identified from FUMA for MOSTest results and the nine FCs mentioned above. E) Gene expression from BrainSpan for the interesting genes prioritised by FUMA.

6.2.3 PrecR↔RoIS and T1a↔T1a/HeschlR endophenotypes driven by *PLCE1*, *NOC3L* and *HELLS* genes

Identification of 'central' endophenotypes associated with 10q23.33 genomic risk region.

A locus in 10q23.33 was highlighted by the mvGWAS. At the univariate level, no endophenotype reached the genome-wide significance threshold for the lead SNP in this locus (rs11187838). Consequently, we did not perform the genetic correlation study, nor perform a *post hoc* study on potential supporting anatomical tracts from dMRI data.

Functional annotation: Locus in 10q23.33 associated to PrecR↔RoIS T1a↔T1a/HeschlR endophenotypes. Four independent SNPs were identified in locus 10q23.33 (rs11187838, rs17109875, rs11187844, rs20772180).

- The subregion around all four SNPs has its chromatin state annotated as (weak) actively-transcribed states (Tx, TxWk) in the brain tissues, specifically in the ganglion eminence and cortex derived primary cultured neurospheres, hippocampus (middle), substantia nigra, anterior caudate, angular gyrus, Dorsolateral/Prefrontal cortex, brain germinal matrix, fetal brain female/male and NH-A astrocytes primary cells.
- Two exonic variants are noteworthy: The rs2274224 ($LD_{rs11187838} r^2 > 0.99$, $p_{mvGWAS} = 5.04e - 14$) and rs11187895 ($LD_{rs17109875} r^2 > 0.6$, $p_{mvGWAS} = 3.08e - 7$) SNPs are nonsynonymous SNV within exon 19 of *PLCE1* and exon 11 of *NOC3L* and are both not predicted to alter protein function (Polyphen-2="benign") but are predicted to have a deleterious effect (CADD = 17.35, CADD = 19.24).
- Moreover, we observed three SNP (rs11187870, rs11187877, rs145707916) ($LD_{rs17109875} r^2 > 0.66$, $p_{mvGWAS} < 7.546e - 7$) located in 3'-UTR of *PLCE1:NOC3L*.
- Regarding eQTL annotations, the variants in the 10q23.33 locus act as eQTL for *HELLS*, *NOC3L* and *PLCE1* genes in different brain tissues including brain cerebellum, brain cerebellar hemisphere, Brain nucleus accumbens basal ganglia, hippocampus ($p_{FDR} < 5e-2$ in GTEx/v8 data source).

These functional characterisation highlight these three genes (*HELLS*, *NOC3L* and *PLCE1*) that may influence the FCs related to language processing in humans.

Locus regulating *PLCE1*, *NOC3L* and *HELLS* associated with PrecR↔RolS and T1a↔T1a.HeschlR.

A locus in 10q23.33 was highlighted by the mvGWAS. At the univariate level, no endophenotype reached the genome-wide significance threshold. But looking at the suggestive threshold $p = 1e - 5$, we pinpoint putative 'central' endophenotypes to aid interpretation of the processes underlying this association signal. Two bilateral fronto-temporal endophenotypes were the most associated to rs11187838: the precentral-Rolandic FC endophenotypes (PrecR↔RolS, $p = 1.85e - 07$) and the right anterior part of the superior temporal gyrus (T1aR) overlapping Heschl's gyrus (T1a/HeschlR) and its homotopic areas of LH primary auditory regions (T1a↔T1a/HeschlR, $p = 9.61e - 06$). All these regions participate in an elementary audio-motor loop involved in both comprehension and production of syllables forming a bilateral fronto-temporal network activated by the auditory representation of speech sounds [Vigneau et al., 2006, 2011]. SNPs of this genomic region in high LD with the lead rs11187838 act as an eQTL of *HELLS*, *NOC3L*, *PLCE1* genes in multiple brain tissues. The *HELLS* gene encodes the lymphoid-specific helicase (Lsh), a member of the SNF2 helicase family of chromatin remodeling proteins. Patients with a genetic mutation of *HELLS* present psychomotor retardation including slow cognitive, motor development and psychomotor impairment [Thijssen et al., 2015]. The Lsh protein might play a role as epigenetic regulator in neural cells [Han et al., 2017]. Finally, we observed that the three genes (*NOC3L*, *PLCE1*, *HELLS*) are expressed in the human brain, across developmental stages from early prenatal to early mid prenatal (8-17 pcw, BrainSpan).

Taken together, these results indicate that the three highlighted genes (*PLCE1*, *NOC3L* and *HELLS*) in the 10q23.33 locus, as potential candidates in the study of key genes playing a role in the *bilateral fronto-temporal auditory-motor* network.

6.3 Limitations

Although functional MR imaging is recognised to produce valuable endophenotypes [Elliott et al., 2018], resting-state FC represent remote measures of language as the participants are not engaged in any language task-experiments and simply stay at rest. Yet, as stated in the introduction 1.3.2.1, previous works reported that resting-state derived endophenotypes are correlated with behavioural language processing [Cheema et al., 2021, Cross et al., 2021, Koyama et al., 2011, Stevens et al., 2017]. However, with the UK Biobank release at hand in this study, no behavioural language scores were provided. We were thus unable to check the correlations between our resting-state FC endophenotypes and language behavioural scores. Moreover, we observed a hierarchy of language heritability estimates from the Human Connectome Project (twin study) with 18-72% for language-behavioural scores, 11.6-34.6% for language-task activations [Le Guen et al., 2018] and 11.6-49.8% for resting-state FC, making resting-state FC not the most suitable endophenotype for an association study about language.

The lack of a large, age-matched replication sample represents one major limitation of the present study in the sense that we could not reproduce all our results.

Multivariate methods display a modest improvement in power as compared to univariate approach, and sometimes display lower power. Additionally, the results are less straightforward to interpret. We have addressed this issue by assessing each of the prioritized loci at the univariate level, to pinpoint at central endophenotypes that are contributing the most to the multivariate signal. Such a complex trait as language may be driven by a lot of interacting genes, a multivariate approach on the SNPs side is highly desired to uncover relevant gene pathways in language development and processing.

Compared to structural endophenotypes, the FCs have low amplitude which hinders the study in terms of statistical power. This observation constitute another limitation that is somehow surpassed when working on large scale cohorts and using multivariate approaches.

Another potential limitation is the UK Biobank dataset in which this study is based. It should be noted that the UKB constitutes a relatively old sample. Future studies in other developmental stages (i.e. children, adolescent, young-adult) will inform us whether the observed associations are stable across development, or whether they reflect some age-related specificity.

6.4 Conclusions

In this study, we extracted language endophenotypes from rsfMRI acquisitions in the largest imaging-genetic cohort on general population to date. To make these endophenotypes as closer as possible to language function in the brain, we adopted a ROI-based approach for FC estimation with language ROIs derived from an comprehensive meta-analysis of tbfMRI studies on language. This approach makes the endophenotypes comparable across individuals and thus suitable for an association study. After filtering on heritability significance, we performed a multivariate GWAS technique in order to take advantage of the correlation structure among rsfMRI FC and uncover potential pleiotropic loci.

Thereby, we highlighted potential key genes related to language processing including *EPHA3* gene in the 3p11.1 locus with a role in the *fronto-temporal semantic* network, *THBS1* gene, modulated by the 15q14 locus, associated with the functional connectivity part of the *perceptual motor* interaction required for language, and *PLCE1* gene in the 10q23.33 locus with a potential role in the *bilateral fronto-temporal auditory-motor* network. These genes could be prioritized to study language in suitable genetic model.

Furthermore, two results that have not shown significant association in the replication sample, have been reported in previous works; a 3q24 locus in *ZIC4* which is involved in visual and auditory pathway development [Horng et al., 2009] has been associated with brain asymmetry [Sha et al., 2021]; a 14q23.1 locus near *DACT1* has been reported to be associate with the STAP [Le Guen et al., 2018] and superior temporal sulcus surface area [Sha et al., 2021]. Altogether, theses results provide an novel insight into the genetic architecture of the language in humans. A growing number of works claim that language studies should consider task-free fMRI data in general population in order to really focus on the neurobiological organization of language and how it supports natural language, i.e. as it is used in everyday life [Hasson et al., 2018]. On a clinical side, such research settings are highly desirable as there is a need to map language areas in patients unable to perform language tasks [Branco et al., 2016, Klingbeil et al., 2019, Park et al., 2020, Ramage et al., 2020]. By validating our findings in rsfMRI FC endophenotypes with dMRI-derived endophenotypes, we tried to provide both a functional and structural connectome point of view of the language organisation in the brain.

Finally, recent imaging-genetic approaches such as multivariate GWAS, allowed us to increase in statistical power and circumvent the small effect sizes of these original endophenotypes. We believe that such approaches are really promising and will broadly disseminate in the imaging-genetic field and beyond. With the presented approach, we tried to contribute to these innovative trends and to pave the way to other alternative task-free approaches to study natural language and its genetic underpinnings.

* * *
* *
*

The genetic architecture of anatomical connectivity in human language

Chapter Outline

7.1	Multivariate genome-wide association studies	124
7.2	Validation of lead SNPs associated with ACs.	128
7.3	Functional annotations of genomic loci associated with language	133
7.4	Discussion	139

In the chapter 6, we used resting-state fMRI in a discovery sample of 32,186 healthy volunteers from the UK Biobank [Sudlow et al., 2015], the compiled information of a large-scale meta-analysis on language components [Vigneau et al., 2006, 2011] and performed a multivariate genetic association of human language functional connectome endophenotypes filtered based on heritability significance presented in the chapter 4.

In the present chapter, the same experimental design is implemented regarding anatomical connectivity (AC) as it can be measured from dMRI. Using the significantly heritable AC endophenotypes, we performed a multivariate genome-wide association study on 31,775 participants from UK Biobank. Briefly, these ACs endophenotypes consists on nine measures (MD, FA, ICVF, ISOVF, L1, L2, L3, MO, OD) of the following bilateral white matter tracts: arcuate anterior/long/posterior segment, the frontal aslant tract, the frontal inferior longitudinal, the frontal superior longitudinal, fronto insular tract 1 to 5, the inferior fronto occipital fasciculus, inferior longitudinal, superior longitudinal fasciculus I-III, uncinate and the corpus callosum.

The results from this analysis were subjected to a replication study in an independent sample ($N=4,672$). Finally, these results were functionally annotated and compared with genes identified as language-related in the literature.

7.1 Multivariate genome-wide association studies

We performed a multivariate genome-wide association studies (mvGWAS) between the filtered imputed genotypes and the set of significantly heritable endophenotypes ACs highlighted in the chapter 4.2.2, using the Multivariate Omnibus Statistical Test (MOSTest) [van der Meer et al., 2020].

As described in the Chapter 2.6, all endophenotypes were pre-residualised with covariates including sex, genotype array type, age, recruitment site, and ten first genetic principal components provided by UK Biobank. In addition, MOSTest performs a rank-based inverse-normal transformation of the residualised endophenotypes to ensure that the inputs are normally distributed. The distributions across the participants of all endophenotypes were visually inspected before and after covariate adjustment. MOSTest generated multivariate summary statistics that capture the significance of the association across all heritable language endophenotypes. To account for multiple testing over the whole genome, statistically significant SNPs were considered as those reaching the genome-wide threshold $p = 5e-8$.

More specifically, we performed a mvGWAS with the 313 ACs with significant SNP-based heritability. This analysis tested each SNP separately for its simultaneous association with the 313 ACs and yielded 34,881 significant SNPs at a genomic threshold, distributed on all autosomal chromosomes.

FUMA [Watanabe et al., 2017] software was used to analyse mvGWAS results and identify lead SNPs at each associated locus. Considering the genome-wide significance threshold $p = 5e-8$, there were 285 distinct genomic loci distributed on 21 chromosomes, associated with different aspects of brain-language AC and represented by 367 lead SNPs. (Fig.7.1, 7.2, Table 7.1).

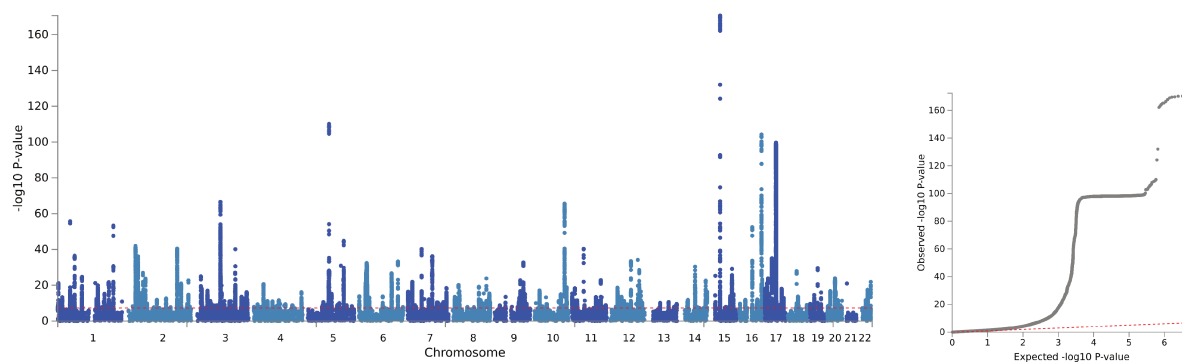
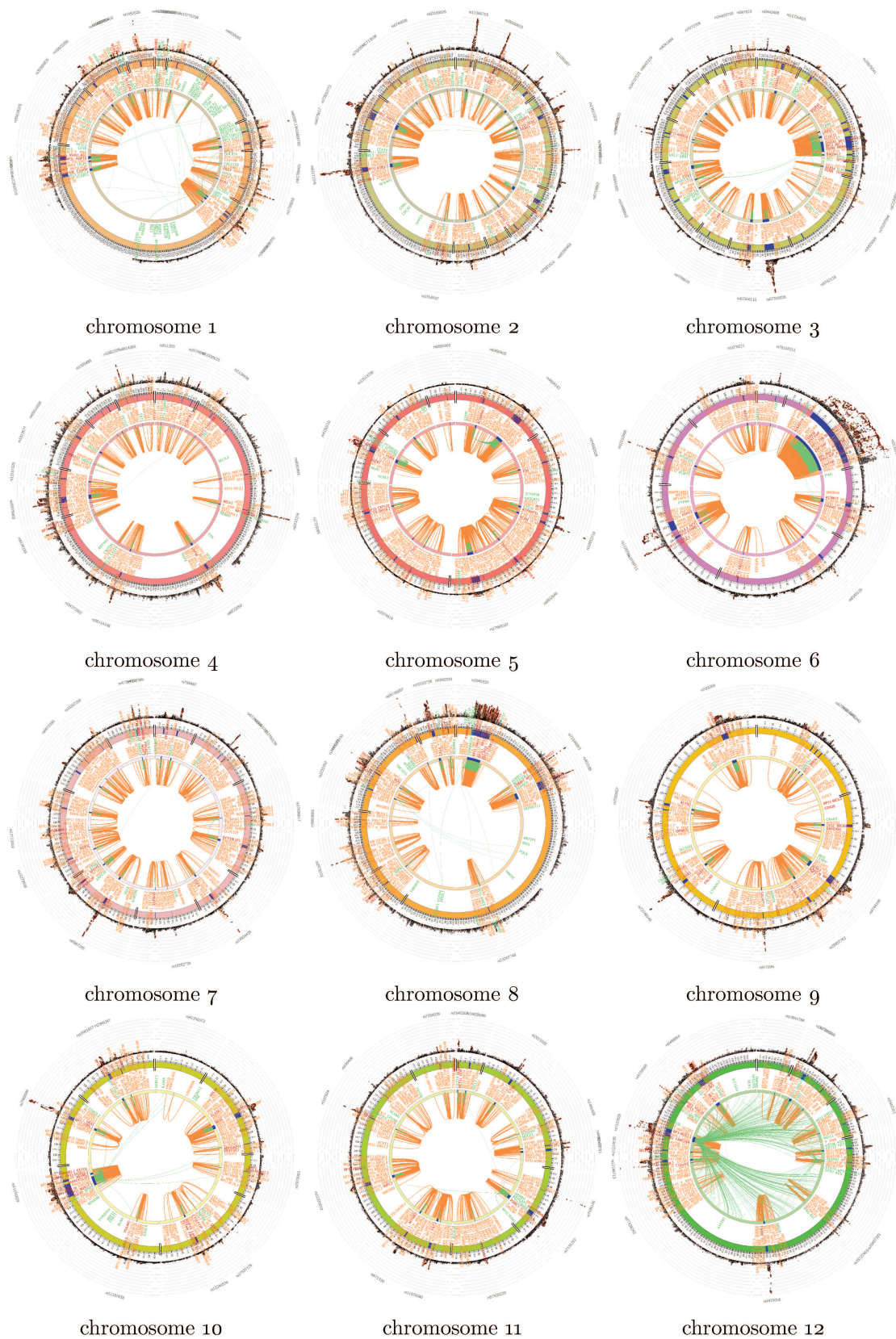


Figure 7.1 – **Multivariate GWAS analysis of the anatomical connectivity in 31,753 participants.** Manhattan plot for multivariate GWAS across 313 FCs. The red dashed line indicates the genome-wide significance threshold $p = 5e-8$. The Quantile-quantile plot is also shown.



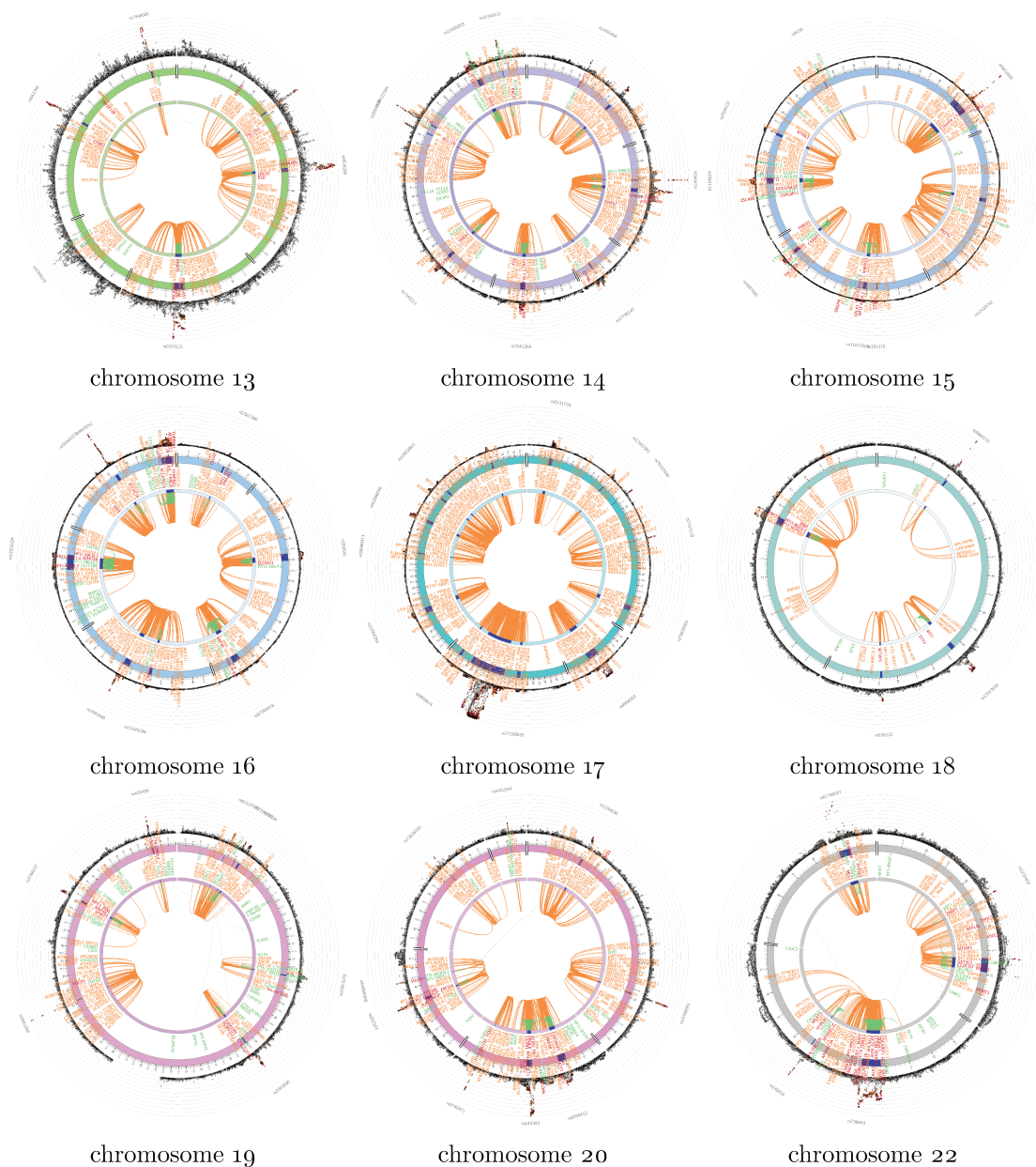


Figure 7.2 – **Genomic loci, eQTL associations and chromatin interactions identified via multivariate GWAS for anatomical connectivity.** Circos plot representing the genomic risk loci, and the genes associated with the loci by chromatin interactions and eQTLs. From outer layer to inner layer: Manhattan plot. Genomic risk loci are in blue. Genes mapped by chromatin interaction are in orange. Genes mapped by eQTL are in green. Genes mapped by both are in red. Chromatin interaction and eQTLs links follows the same color coding presented above. Due to a high number of genes identified using eQTLs mapping (361) and chromatin interaction (874), the chromosome 17 circos plot represent only chromatin interaction.

7.2 Validation of lead SNPs associated with ACs.

The multivariate genome-wide association results were replicated in an independent non-British sample considering the nominal significance threshold $p < 0.05$. Following the same pre-processing steps as for the primary sample, the non-British replication sample consists in 4,672 individuals with a mean age of 53 years (± 7.55) and 2,572 female.

Seventy-one lead SNPs (out of the 367 unveiled by the discovery study) were replicated at the nominal significance level ($p < 5e-2$) on multivariate test in the independent non-British replication dataset. Table 7.1 summarises these results.

Table 7.1 – Genomic loci associated with heritable language anatomical connectivities using MOSTest. Lead SNP: ID of the lead SNPs within each locus. Position: position of the SNP in the hg19 human reference genome. mvqwasP discovery -British-: MOSTest association P value obtained using the discovery sample. mvqwasP replication -non British-: MOSTest association P value obtained using the independent replication sample. Functional category: Functional consequence of the SNP on the gene obtained from ANNOVAR. 'Central' endophenotypes: number of the endophenotypes that contributed most to the multivariate association considering the genome-wide association threshold ($5e - 8$). 'Central' endophenotypes can be appreciated in the Table D.1.

Genomic Locus	Lead SNP	Chr	Position	Functional Category	non effect allele	effect allele	MAF	mvqwasP (discovery -British-)	mvqwasP (replication -non British-)	Nearest Gene	'central' endophenotypes
1	rs4970394	1	962891	intronic	T	C	0.421	4.01E-09	1.68E-01	AGRN	-
2	rs13303058	1	2005909	intronic	T	C	0.037	6.05E-10	1.43E-01	PRKCCZ	-
3	rs11580768	1	2799480	intergenic	T	C	0.459	2.03E-21	4.12E-01	RP11-740P5.3	-
4	rs10779704	1	8693537	intronic	C	A	0.371	2.43E-08	3.57E-01	REER	-
5	rs4920605	1	17315425	ncRNA intronic	A	G	0.431	5.07E-14	3.78E-01	RP1-37C10.3;ATP13A2	-
6	rs6658111	1	47980916	intergenic	G	T	0.360	2.01E-56	4.27E-02	AL358458.1	1
7	rs10888700	1	51886337	intronic	C	T	0.433	1.31E-10	4.33E-01	FAF1	-
8	rs61788401	1	59317534	ncRNA intronic	A	T	0.109	3.63E-14	2.67E-01	LINC01135	1
9	1:64228261 CCTCT C	1	64228261	intergenic	C	CCTCT	0.168	6.33E-12	6.04E-01	ROR1	-
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.94E-04	JAK1	40
11	rs59670370	1	92922490	intergenic	T	C	0.107	4.81E-08	3.87E-01	GFL1	-
12	rs3787067	1	94002333	intronic	G	A	0.391	2.00E-25	1.71E-02	FNBPL1	3
13	rs114088373	1	95709197	ncRNA intronic	T	C	0.017	2.26E-08	3.22E-01	TMEM56;RWDD3;RWDD3;RP11-57H12.5	-
14	rs6693567	1	150510660	intergenic	T	C	0.265	1.65E-09	1.21E-01	RP11-54A4.2	-
14	rs72992015	1	150842696	intronic	T	C	0.212	3.47E-13	3.01E-01	ARNT	-
15	rs3841064	1	153919858	upstream:downstream	AC	A	0.491	1.14E-12	4.69E-02	CRT2	4
16	rs4745	1	155106227	exonic	T	A	0.479	1.20E-20	1.60E-01	EFNA1	1
16	rs72704130	1	155207030	intronic	A	G	0.029	4.53E-10	6.66E-01	GBA	-
16	rs372518691	1	155856347	intergenic	C	CT	0.318	1.57E-12	7.94E-01	SYT11	-
17	rs35039375	1	175168863	intronic	G	A	0.091	1.67E-08	3.57E-01	PRRC2C	-
18	1:178770951 AAGGG A	1	178770951	intronic	A	AAGGG	0.429	8.46E-12	2.57E-01	RALGPS2	-
19	rs35306826	1	180956936	intronic	A	T	0.408	6.65E-16	3.00E-03	STX6	1
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	18
20	rs17641524	1	197704717	exonic	T	C	0.209	1.16E-12	3.48E-02	DENND1B	-
21	rs4950982	1	205078247	intronic	C	T	0.122	2.02E-09	1.08E-01	RBBP5	-
22	rs1772159	1	205759195	UTR3	A	G	0.399	3.57E-10	2.32E-01	SLC41A1	-
23	rs2480403	1	206614441	intronic	C	T	0.203	1.14E-10	4.97E-01	SRGAP2	-
24	rs1452628	1	215139887	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	59
24	rs2601645	1	215185552	intronic	T	C	0.084	2.12E-21	6.34E-02	KCNK2	16
25	rs11380753	2	27148654	intronic	CT	C	0.385	1.04E-42	1.99E-03	DPYSL5	3
25	rs41288789	2	27167617	exonic	A	G	0.051	8.28E-09	8.12E-01	DPYSL5	-
26	rs35050623	2	37063240	intergenic	CT	C	0.462	7.04E-37	2.08E-01	AC007382.1	2
27	2:42278500 GGGGCCCCAGCCCTGGAT G	2	42278500	intronic	G	GGGGCCCC AGCCCTGGAT	0.234	1.74E-14	9.75E-02	PKDCC	-
28	rs11891807	2	48288816	ncRNA intronic	G	T	0.385	2.29E-16	5.22E-02	AC079807.4	1
29	2:56116193 TTGATTCTTGGCTCCCTCC T	2	56116193	intronic	T	TTGATTCTTGG CCTCCCTCC	0.090	1.13E-27	2.69E-03	EFEMP1	8
30	rs200493119	2	58154132	intronic	CG	C	0.399	2.04E-18	3.64E-01	VPRK2	-
31	rs13011022	2	60626822	intergenic	A	G	0.409	1.55E-11	8.67E-02	AC007381.2	-
32	rs7600385	2	65984026	ncRNA intronic	T	A	0.401	2.03E-24	2.78E-02	AC074391.1	2
33	rs11688599	2	66756550	intronic	T	C	0.320	3.63E-08	2.92E-01	MEIS1	-
34	rs6719862	2	71958862	intergenic	G	A	0.375	1.62E-10	6.97E-01	DYSF	-
35	rs143238565	2	98430405	intronic	TA	T	0.076	2.26E-08	6.92E-01	TMEM131	-
36	rs60240458	2	109949227	intronic	C	A	0.488	1.37E-12	8.10E-02	SH3RF3	-
37	rs7581524	2	113234316	intergenic	T	C	0.275	3.44E-08	7.36E-01	TTL	-
38	rs368600848	2	119755955	intergenic	A	G	0.014	4.33E-08	6.75E-01	MARCO	-
39	rs144495495	2	145692649	ncRNA intronic	CAT	C	0.250	2.36E-13	4.22E-01	TEXA1	9
40	rs1558957	2	162890175	intronic	C	T	0.411	1.97E-08	3.07E-01	DPP4	-
41	rs56136540	2	187501682	intronic	A	G	0.404	7.32E-16	2.60E-01	ITGAV	-
41	rs62172370	2	188289174	ncRNA intronic	A	C	0.303	2.99E-41	6.17E-03	AC007319.1;CALCRL	-
42	rs4676517	2	200342749	intergenic	G	A	0.161	5.07E-12	5.38E-01	SATB2-AS1	-
43	rs34349435	2	203619455	intronic	A	C	0.034	2.03E-09	6.95E-01	FAM117B	4
43	rs72932772	2	203682304	intronic	G	C	0.129	1.05E-21	5.23E-02	ICAI1	10
43	rs113796732	2	204323814	intronic	C	G	0.039	2.50E-12	3.13E-02	RAPH1	-
44	rs142787179	2	204840948	intergenic	C	G	0.028	5.42E-09	7.51E-01	ICOS	-
45	rs7593551	2	225231713	intergenic	C	T	0.299	3.17E-08	8.04E-01	FAM124B	-
46	rs1156040	2	227057413	intergenic	A	G	0.133	6.60E-10	2.97E-01	AC068138.1	-
46	rs2713536	2	227163642	intergenic	T	C	0.363	2.93E-10	5.08E-01	AC068138.1	-
47	rs6749936	2	232055684	intergenic	T	C	0.211	2.23E-23	1.12E-02	ARMC9	21
48	rs62189826	2	237501894	intergenic	C	T	0.181	9.16E-09	8.97E-01	ACKR3	-
49	rs13079464	3	13822439	intergenic	C	G	0.467	7.77E-12	3.48E-01	LINC00620	4
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	107
50	rs11720653	3	17439510	intronic	C	T	0.316	1.04E-18	1.95E-01	TBC1D5	1
51	rs111710612	3	33577384	intronic	C	A	0.018	4.70E-08	6.36E-01	CLASP2	-
51	3:33846846 CA C	3	33846846	intronic	C	CA	0.342	2.90E-08	7.07E-01	PDCD61P	-
52	rs142003705	3	39479611	intergenic	G	A	0.020	1.26E-09	8.81E-02	AC099332.1	8
52	rs10675541	3	39510286	intronic	TAC	T	0.480	2.16E-16	3.10E-01	MOBP	5
53	3:45795731 CTCTT C	3	45795731	intergenic	C	CTCTT	0.104	2.21E-11	2.18E-01	SLC6A20	-
54	3:47312035 AAAAAAG A	3	47312035	intronic	A	AAAAAAG	0.402	5.85E-09	9.80E-02	KIF9	-
55	rs59623077	3	49106542	intronic	TA	T	0.336	6.27E-09	8.95E-01	QRICH1	1
55	rs7627910	3	49971831	intergenic	C	T	0.458	2.73E-08	8.03E-01	MON1A	-
55	rs71326918	3	50174844	ncRNA intronic	A	C	0.110	6.40E-12	5.53E-02	RP11-493K19.3	-

Genomic Locus	Lead SNP	Chr	Position	Functional Category	non effect allele	effect allele	MAF	mvgwasP (discovery -British-)	mvgwasP (replication -non British-)	Nearest Gene	'central' endophenotypes
55	rs9832454	3	50217558	intronic	T	C	0.089	1.35E-11	5.60E-03	SEMA3F	1
55	rs62260798	3	50274278	intronic	T	C	0.080	4.74E-08	9.39E-01	GNAI2	-
56	rs762174010	3	52572841	intronic	T	TTTG	0.453	1.26E-10	6.46E-01	SMIM4	-
57	rs2688771	3	57937992	intergenic	T	C	0.435	1.13E-09	6.99E-01	PPIAP16	-
58	rs62243554	3	62568709	intronic	T	A	0.339	5.43E-09	2.23E-01	CADPS	-
59	rs76207096	3	66044816	intergenic	A	C	0.303	2.77E-09	6.57E-02	MAG1	-
60	rs3925059	3	69463714	intronic	G	C	0.319	2.46E-09	7.16E-01	FRMD4B	-
61	rs6782116	3	77176032	intronic	T	C	0.427	1.55E-09	1.80E-01	ROBO2	-
61	rs553398120	3	77340010	intronic	C	A	0.421	5.46E-09	8.86E-01	ROBO2	-
62	rs10638977	3	88781746	intergenic	AAC	A	0.499	8.44E-12	1.64E-01	NDUF5P5	-
62	rs58378065	3	88943830	intergenic	AT	A	0.190	1.50E-08	8.65E-01	NARG2P2	-
62	3:89257759	AGAAG	A	intronic	A	AGAAG	0.290	2.56E-08	3.31E-02	EPHA3	-
62	rs67316928	3	89482234	intronic	C	T	0.398	2.94E-07	3.17E-05	EPHA3	1
63	rs62266110	3	93537923	intergenic	A	G	0.319	1.33E-20	1.39E-01	RNU6-488P	-
64	rs6788676	3	104674040	intergenic	A	G	0.355	6.95E-12	1.76E-01	ALCAM	-
65	rs373566812	3	124462488	intronic	AAAAT	A	0.165	1.02E-11	7.18E-01	UMPS	3
66	rs747333548	3	129568656	intronic	TAC	T	0.082	9.17E-12	2.32E-01	TMCC1	-
67	rs10490862	3	132160713	intronic	G	A	0.120	4.45E-09	5.03E-01	DNAJC13	-
68	rs9881400	3	135861666	intronic	T	G	0.265	1.43E-09	8.39E-02	PPP2R3A	-
69	rs36008198	3	146362717	intergenic	TGTA	T	0.394	4.23E-09	1.09E-01	PLSCR5	-
69	rs9847505	3	146369166	intergenic	G	C	0.462	9.03E-28	6.43E-02	PLSCR5	-
70	rs62275311	3	147034001	ncRNA intronic	C	A	0.245	4.95E-20	2.74E-02	RP11-649A16.1	1
70	rs2279829	3	147106319	UTR3	T	C	0.212	7.46E-41	2.42E-02	ZIC4	-
71	rs34737522	3	168892763	intronic	CT	C	0.144	2.67E-08	6.56E-01	MECOM	-
72	3:170987040	ACTCT	A	intronic	A	ACTCT	0.455	2.71E-14	7.93E-02	TNIK	2
73	rs16856586	3	171436724	intronic	A	G	0.302	1.42E-09	8.16E-01	PLD1	-
73	rs9647379	3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	28
74	rs4241484	3	177316085	ncRNA intronic	A	C	0.406	2.40E-12	6.99E-01	LINC00578	-
75	rs3772705	3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	25
76	rs34859790	3	190652830	intergenic	T	C	0.384	7.78E-17	3.23E-01	GMNC	-
77	rs687610	3	193515781	intergenic	G	C	0.426	1.54E-12	2.74E-02	RP11-528A4.3	1
78	rs811316	4	1689377	intergenic	C	T	0.190	1.23E-08	6.53E-01	FAM53A	-
79	rs3774893	4	5887388	intronic	C	A	0.418	2.87E-08	5.41E-01	CRMP1	-
80	rs11930623	4	7931682	intronic	C	A	0.405	4.95E-08	2.20E-01	AFAP1	-
81	rs1134634	4	15603069	UTR3	C	G	0.415	4.61E-10	3.81E-01	CC2D2A	-
82	4:20318933	GA	G	intronic	G	GA	0.062	8.34E-09	6.02E-01	SLIT2	-
83	rs4692493	4	30949606	intronic	T	A	0.495	1.08E-08	6.54E-02	PCDH7	-
84	rs6812216	4	38651939	ncRNA intronic	A	G	0.354	2.24E-21	3.35E-01	RP11-617D20.1:AC021860.1	-
85	rs10032190	4	53734940	intergenic	A	C	0.243	3.20E-08	6.83E-01	RASL11B	-
85	rs6822050	4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	61
86	rs36114335	4	73319708	intronic	AT	A	0.350	1.59E-11	4.62E-01	ADAMTS3	-
87	rs34272002	4	77657959	intronic	CAG	C	0.385	2.39E-08	4.06E-01	SHROOM3	-
88	rs4148155	4	89054667	intronic	G	A	0.114	4.35E-09	8.15E-01	ABCG2	-
89	4:96797672	TA	T	intergenic	T	TA	0.099	4.26E-08	3.10E-01	RNU6-1059P	-
90	rs35976463	4	97926008	intergenic	CATG	C	0.316	2.24E-12	6.82E-01	COX7A2P2	10
90	rs62315846	4	98455696	ncRNA intronic	G	A	0.135	2.15E-09	5.51E-02	STPG2:RP11-681L8.1	-
91	rs13107325	4	103188709	exonic	T	C	0.076	3.20E-10	1.30E-01	SLC39A8	1
92	rs1913577	4	119297657	intergenic	G	A	0.365	2.13E-09	4.93E-01	PRSS12	-
93	rs45510500	4	123179900	exonic	T	C	0.056	3.91E-10	3.13E-03	KIAA1109	-
94	4:149409090	CAG	C	intergenic	C	CAG	0.465	1.70E-09	8.38E-01	RP11-655B23.1	-
95	rs2305980	4	151171269	intronic	A	T	0.257	1.74E-09	2.73E-01	DCLK2	-
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	30
97	rs6814390	4	187835465	ncRNA intronic	A	G	0.161	1.94E-08	2.12E-01	F11-AS1:RP11-215A19.2	-
98	rs6450402	5	56040390	intergenic	T	A	0.226	1.26E-08	5.12E-01	AC008940.1	-
99	rs72755166	5	59982433	intronic	G	A	0.134	3.85E-10	2.60E-01	DEPDCB	-
99	rs4604142	5	60642595	intronic	T	C	0.499	2.71E-11	4.09E-01	ZSWIM6	-
100	rs34592828	5	75996909	exonic	A	G	0.045	6.33E-10	4.43E-01	IQGAP2:CTD-2384B11.2	1
101	rs145824047	5	81800313	intergenic	A	G	0.396	4.50E-10	8.02E-01	CTD-2015A6.1	2
101	rs752021310	5	81866998	intergenic	C	CTAAT	0.447	3.20E-23	8.87E-03	CTD-2015A6.2	2
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	145
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	140
102	rs309588	5	82852578	ncRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN:VCAN-AS1	55
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN:VCAN-AS1	203
103	rs6882046	5	87968864	ncRNA exonic	G	A	0.262	4.02E-12	5.02E-01	LINC00461	-
104	rs17669337	5	92187932	intergenic	T	C	0.412	4.08E-08	3.21E-01	CTC-458G6.4	-
105	rs73135321	5	92928643	intronic	A	G	0.093	3.45E-09	6.63E-01	NR2F1	-
105	rs115136616	5	93021116	intronic	T	C	0.047	6.85E-16	1.73E-01	FAM172A	1
106	rs2974619	5	114470013	intronic	G	T	0.400	1.90E-09	1.11E-01	TRIM36	-
107	rs11353435	5	116608074	intergenic	AAGAT	A	0.330	7.03E-10	1.46E-01	RPL35AP15	-
108	rs7729285	5	122754416	intronic	G	A	0.431	9.63E-11	2.12E-02	CEP120	1
109	rs10037401	5	139568762	intronic	T	A	0.208	1.48E-24	8.21E-03	CYSTM1	-
109	rs4150210	5	13972193	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	38
110	rs12518238	5	151259002	intronic	A	G	0.250	1.62E-09	3.86E-01	GLRA1	-
111	rs6555969	5	171128464	intergenic	C	T	0.320	4.50E-09	2.17E-01	AC011410.1	-
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	28
113	rs940404	6	25343444	intronic	A	T	0.068	4.17E-08	1.84E-02	LRRc16A	-
113	rs1800562	6	26093141	exonic	A	G	0.078	1.83E-10	5.26E-02	HFE	2
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	17
113	rs34196306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	32
113	rs71546548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	8.72E-01	HNRNP1P1	15
113	rs34676049	6	28453618	intergenic	G	T	0.108	1.93E-28	4.67E-04	GPX6	40
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	55
114	rs1520	6	39526867	intronic	G	A	0.374	6.46E-14	9.02E-01	KIF6	-
114	rs9380899	6	39794806	intronic	G	C	0.271	8.76E-12	5.17E-01	DAAM2	13
115	rs9345125	6	92002653	intergenic	G	A	0.171	1.90E-14	2.68E-01	RP1-17710.1	-
116	rs770637388	6	98561749	ncRNA intronic	C	CAATT	0.486	2.91E-08	1.82E-02	RP11-436D23.1	-
117	rs3799711	6	126075313	intronic	C	T	0.480	2.29E-27	3.15E-01	HEY2	1
118	rs75343521	6	126732511	intergenic	T	C	0.041	8.00E-09	2.25E-01	CENPW	-
118	rs1120786	6	126808637	intergenic	G	T	0.451	2.50E-22	4.85E-01	MIR588	2
119	rs12192990	6	151021978	intronic	G	C	0.412	5.37E-34	3.98E-01	PLEKHG1	1
120	rs1078211	6	170201300	ncRNA exonic	C	T	0.403	1.51E-08	1.99E-01	LINC00574	-
121	rs201865165	7	1087676	intronic	C	CA	0.422	1.27E-08	9.37E-01	C7orf50:GPR146	-
122	rs798487	7	2802943	intronic	A	G	0.297	4.19E-17	1.46E-01	AMZ1:GNA12	8
123	rs67248060	7	18899163	intronic	T	C	0.366	1.33E-22	4.78E-01	HDAC9	15
124	rs10950739	7	19628139	intergenic	A	G	0.374	4.38E-10	8.63E-01	AC007091.1	-
125	rs12700239	7	20865979	upstream	A	C	0.211	6.75E-09	4.69E-01	RPL23P8	-

Genomic Locus	Lead SNP	Chr	Position	Functional Category	non effect allele	effect allele	MAF	mvgwasP (discovery -British-)	mvgwasP (replication -non British-)	Nearest Gene	'central' endophenotypes
126	rs17686617	7	29346518	intronic	T	C	0.313	1.36E-09	1.94E-01	CHN2	2
127	7:33344749	7	33344749	intronic	T	TAA	0.304	1.02E-08	4.20E-01	BBS9	-
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	67
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	16
129	rs10252734	7	81418596	intergenic	G	T	0.400	5.23E-12	2.30E-01	HGF	1
130	rs761326546	7	83628746	intronic	T	TAATTA	0.212	6.84E-13	2.51E-01	SEMA3A	-
131	rs6967285	7	96200086	nan	C	T	0.308	6.42E-37	3.87E-02	SHFM1	2
132	7:99965328	7	99965328	ncRNA exonic	G	GA	0.189	1.71E-08	2.09E-01	STAG3L5P-PVRIG2P-PILRB:PILRB:PILRA	-
133	rs10279936	7	114147217	intronic	A	C	0.491	4.01E-11	2.96E-01	FOXP2	-
134	rs12113919	7	117612315	intergenic	G	C	0.410	4.78E-11	4.22E-01	AC003084.2	-
135	rs566523025	7	120759189	intronic	CT	C	0.365	4.06E-09	1.19E-02	CPED1	-
136	7:132558516	7	132558516	intronic	T	TATC	0.253	3.12E-08	3.40E-01	CHCHD3	-
137	rs6973330	7	134388544	intergenic	T	G	0.358	2.90E-10	3.26E-01	AC009276.4	7
138	rs10267280	7	137111423	intronic	T	C	0.366	2.71E-09	4.83E-01	DGKI	-
139	rs4725995	7	150819169	intronic	T	C	0.305	3.86E-19	3.41E-01	AGAP3	10
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	23
141	rs2945232	8	8098038	ncRNA exonic	C	T	0.472	3.23E-14	7.48E-01	FAM86B3P:ALG1L13P	-
141	rs4075359	8	9487813	intronic	C	T	0.377	1.73E-13	7.90E-02	TNKS	2
141	rs10108511	8	11435516	ncRNA intronic	C	T	0.469	6.75E-14	6.52E-01	LINC00208	-
142	rs71548883	8	23101514	intronic	CT	C	0.384	8.31E-21	3.02E-01	CHMP7	1
143	rs17055142	8	26201601	intronic	T	C	0.076	4.76E-10	6.70E-02	PPP2R2A	-
143	rs11421016	8	26327197	intronic	GA	G	0.174	3.05E-10	2.17E-01	BNIP3L	-
143	rs565497548	8	26452540	intronic	A	AAATC	0.000	2.33E-14	1.77E-02	DPYSL2	-
143	rs403185	8	26476721	intronic	G	A	0.408	1.84E-17	5.67E-02	DPYSL2	1
144	rs13207748	8	57185107	intergenic	A	G	0.219	1.50E-08	6.04E-01	RP11-140H16.2	-
145	8:82719186	8	82719186	ncRNA intronic	T	TA	0.414	1.62E-08	2.21E-01	RP13-923O23.6:SNX16	-
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	36
147	rs5893855	8	108394765	intronic	GA	G	0.421	2.00E-12	6.38E-02	ANGPT1	-
148	rs2721952	8	116645056	intronic	T	C	0.401	1.02E-10	1.16E-02	TRPS1	-
149	rs7463176	8	119957625	intronic	G	A	0.455	2.24E-09	5.54E-01	TNFRSF11B	-
150	rs10283100	8	120596023	exonic	G	A	0.060	3.68E-16	2.47E-03	ENPP2	19
150	rs1992721	8	120597417	intronic	G	A	0.283	2.72E-09	7.31E-01	ENPP2	-
151	rs55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CCDC26	81
152	rs45613837	8	141744917	intronic	A	G	0.452	2.94E-14	6.80E-01	PTK2	-
152	rs10102728	8	142205964	downstream	A	G	0.409	3.98E-16	4.91E-02	DENND3	2
153	rs6992333	8	144993377	exonic	G	A	0.430	3.58E-11	1.96E-01	PLEC	1
154	rs679038	9	22029080	ncRNA intronic	A	G	0.421	1.68E-08	1.62E-01	RP11-145E5.5:CDKN2B-AS1	14
155	rs693370	9	22446297	upstream	A	G	0.118	1.86E-09	6.32E-01	DMRTA1	-
156	rs842842	9	2267821	ncRNA intronic	C	A	0.421	2.53E-08	3.74E-01	RP11-399D6.2:RP11-370B11.3	-
157	rs1329733	9	92208227	intergenic	G	A	0.480	1.04E-08	6.07E-01	GADD45G	-
158	rs4744186	9	95932546	intergenic	T	C	0.431	2.93E-11	5.89E-02	WNK2	1
158	rs10992828	9	96415368	intronic	A	G	0.272	3.54E-08	8.46E-01	PHF2	-
159	rs28457763	9	98260514	intronic	A	G	0.328	8.49E-14	2.38E-01	PTCH1	-
160	rs911934	9	101637477	intergenic	G	T	0.313	1.75E-23	6.15E-01	RP11-92C4.3	3
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	25
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	59
162	9:118957629	9	118957629	intronic	A	AAAG	0.424	1.55E-08	3.09E-02	PAPPA	3
163	rs7030607	9	119245183	intronic	A	G	0.354	7.46E-16	7.30E-02	ASTN2	1
164	rs783200	9	125857032	intronic	A	G	0.136	2.58E-10	2.17E-01	RABGAP1	-
165	rs41298373	10	7622009	exonic	A	G	0.100	9.46E-11	9.34E-02	ITIH5	-
166	10:21878831	10	21878831	intronic	T	TCCCTTC	0.298	8.21E-18	1.16E-01	MLLT10	-
167	rs77445421	10	25041543	intergenic	G	GAGACAGTCAAGTC	0.352	2.01E-09	2.43E-01	ARHGAP21	10
168	rs3709083	10	33676655	intergenic	T	A	0.291	5.05E-09	6.41E-01	RP11-90B22.1	-
169	rs17537178	10	48442894	intergenic	G	A	0.267	3.61E-10	2.84E-01	GDF10	1
170	rs12240534	10	50318141	intronic	T	C	0.349	3.28E-10	1.67E-01	VSTM4	1
171	rs1187838	10	96938686	intronic	A	G	0.435	6.80E-12	2.43E-01	PLCE1	-
172	rs3740405	10	103908080	intronic	A	G	0.454	4.10E-08	6.11E-01	PPRC1	-
172	rs758741000	10	104299191	intronic	A	ATTGT	0.319	1.10E-15	7.39E-01	SUFU	6
172	rs1163084	10	105039048	intronic	C	T	0.456	2.15E-15	1.09E-01	INA	6
173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	44
173	rs1191829	10	105614452	intronic	T	G	0.353	3.25E-13	1.84E-01	SH3PXD2A	8
174	rs7089699	10	118692698	intronic	G	A	0.249	2.67E-66	9.69E-03	KIAA1598	3
175	rs10901817	10	126443404	intronic	A	G	0.427	1.26E-19	5.77E-01	RP11-12J10.3:METTL10	2
176	rs2366347	10	127674508	intronic	A	G	0.421	1.04E-11	6.53E-02	FANK1	-
177	rs1045502	11	252818	UTR3	A	G	0.086	1.81E-14	4.19E-01	PSMD13	7
178	rs10835060	11	1488143	intergenic	G	T	0.191	2.50E-09	3.98E-01	MOB2	-
179	rs2071020	11	10478228	intronic	T	C	0.452	4.99E-19	7.47E-01	AMPD3	23
179	rs4442541	11	10669172	intronic	G	A	0.311	5.91E-12	2.23E-03	MRV1	-
180	rs1468489	11	17022894	intronic	C	T	0.378	7.34E-09	9.20E-01	PLEKHA7	-
181	rs4488177	11	19454969	intronic	A	C	0.343	2.59E-12	1.49E-01	NAV2	-
182	rs2585753	11	19895756	intronic	G	C	0.226	1.52E-11	4.18E-03	NAV2	5
182	rs35891966	11	20129311	exonic	A	G	0.071	3.87E-08	2.17E-01	NAV2	-
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	89
184	rs7131262	11	47836302	intronic	A	T	0.288	1.44E-09	2.19E-01	NUP160	-
185	rs57630229	11	82409631	ncRNA intronic	A	G	0.113	4.35E-08	3.58E-01	RP11-179A16.1	15
186	rs77407913	11	91968738	intergenic	C	CTATT	0.181	3.96E-14	3.68E-03	RPL7AP57	-
187	11:92504928	11	92504928	intronic	T	TA	0.329	5.33E-10	1.15E-01	FAT3	-
187	rs1020080	11	92578985	intronic	A	G	0.387	9.29E-12	7.11E-02	FAT3	2
188	rs671320	11	95578780	intronic	T	G	0.465	2.40E-12	6.32E-01	MTMR2	-
189	rs11225553	11	102989206	intronic	G	A	0.076	3.24E-08	6.80E-01	DYNC2H1	-
190	rs167684	11	110962880	ncRNA exonic	G	A	0.314	1.38E-23	7.87E-02	RP11-89C3.4	-
191	rs636406	11	114031092	intronic	G	C	0.358	2.80E-09	6.12E-01	ZBTB16	2
192	rs7936928	11	130279168	intronic	C	T	0.395	1.40E-09	4.03E-01	ADAMTS8	-
193	rs10841784	12	21441577	intronic	C	T	0.283	9.57E-09	5.77E-01	SLCO1A2	1
194	rs2075378	12	27788021	exonic	A	G	0.355	2.20E-19	6.03E-01	PPFIBP1	1
195	rs7305286	12	28282047	intergenic	G	A	0.307	3.24E-17	4.99E-04	CCDC91	1
196	12:32495060	12	32495060	intronic	G	GTA	0.210	5.34E-11	2.20E-01	BICD1	-
197	rs2407281	12	43609752	intergenic	C	T	0.253	2.92E-14	1.35E-01	RP11-118A3.1	-
198	rs35427608	12	53826327	intergenic	AT	A	0.177	4.60E-11	8.43E-01	AMHR2	-
199	rs10876870	12	56478002	intronic	C	A	0.425	4.56E-09	2.20E-01	ERBB3	1
200	rs79487293	12	65905126	ncRNA intronic	T	C	0.325	5.11E-10	5.89E-01	RP11-230G5.2	-
201	rs35227403	12	68216239	intergenic	A	T	0.057	1.45E-09	2.06E-01	RP11-43N5.1	-
202	rs4479056	12	79938856	ncRNA intronic	A	G	0.307	3.57E-34	3.65E-01	RP11-359M6.1	4
203	rs77436242	12	99640159	exonic	C	G	0.048	3.03E-10	5.36E-01	ANKS1B	8
204	rs12146713	12	106476805	intronic	C	T	0.096	7.00E-35	1.99E-01	NUAK1	1
205	rs536700128	12	107063485	intronic	GA	G	0.243	1.13E-13	3.57E-01	RP11-144F15.1:RFX4	2
206	rs1114035	12	109102983	intronic	A	T	0.438	1.01E-08	8.17E-01	CORO1C	1
207	rs7137828	12	11932800	intronic	T	C	0.484	2.99E-29	6.03E-01	ATXN2	1
207	rs10219716	12	112503474	ncRNA exonic	A	T	0.085	1.60E-09	3.98E-01	NAA25:RP11-267L14.3	2
208	rs9739560	12	118461574	intronic	A	T	0.417	3.29E-17	1.33E-01	RFC5	33
209	rs6488864	12	123618362	intronic	C	G	0.199	5.66E-11	4.31E-01	PITPNM2	-
210	13:44887065	13	44887065	intergenic	A	AT	0.250	2.04E-08	1.66E-02	SERP2	-
211	rs9526208	13	47208767	intronic	G	A	0.245	7.67E-11	4.92E-02	LRCH1	-
212	rs6563121	13	80175000	intergenic	C	T	0.478	2.38E-11	8.49E-01	LINC01068	-
213	rs2254579	13	101045359	intronic	C	G	0.429	1.30E-			

Genomic Locus	Lead SNP	Chr	Position	Functional Category	non effect allele	effect allele	MAF	mvgwasP (discovery -British-)	mvgwasP (replication -non British-)	Nearest Gene	'central' endophenotypes
214	rs943799	13	107724282	intergenic	A	G	0.484	2.34E-09	2.70E-01	AL445204.1	-
215	rs11838776	13	111040681	intronic	A	G	0.277	3.42E-10	2.91E-03	COL4A2	1
215	rs7996686	13	111093237	intronic	G	A	0.429	2.89E-10	1.48E-01	COL4A2	-
216	rs1955454	14	29226194	ncRNA intronic	G	T	0.438	2.24E-11	1.32E-01	RP11-96617.1	-
217	14:31370483 TA T	14	31370483	intronic	T	TA	0.450	1.93E-12	2.91E-01	STRN3	-
218	rs100459	14	59074130	intergenic	C	A	0.459	4.19E-31	5.26E-01	DACT1	7
219	14:59631075 GTTGT G	14	59631075	intergenic	G	GTTGT	0.120	1.60E-19	4.58E-01	CTD-2315A10.1	-
219	rs17833810	14	59697746	intronic	G	A	0.432	1.79E-09	9.85E-01	DAAM1	-
220	rs17756147	14	69036127	intronic	A	G	0.246	2.17E-08	1.99E-01	RAD51B	-
221	rs7141254	14	75190527	intronic	T	G	0.459	4.12E-08	2.77E-01	FCF1	-
222	rs7142211	14	91910995	intergenic	A	G	0.443	4.74E-08	4.17E-01	SMEK1	1
223	rs12895055	14	99712794	intronic	T	C	0.437	1.10E-09	1.40E-01	BCL11B	-
223	rs2664299	14	99742187	intergenic	C	T	0.420	1.36E-10	4.81E-01	BCL11B	-
224	rs11846963	14	100271376	intronic	C	T	0.403	8.38E-11	8.04E-01	EML1	-
224	rs35522344	14	100283856	intronic	T	C	0.382	2.60E-23	3.21E-01	EML1	1
225	rs12889303	14	104161666	intronic	A	G	0.312	2.80E-08	8.48E-01	KLC1	-
226	rs10140111	14	105644282	ncRNA intronic	G	C	0.384	5.59E-17	1.16E-02	RP11-44N21.4:NUDT14	-
227	rs2929686	15	39085781	intergenic	T	C	0.387	7.99E-10	6.86E-01	RP11-326N17.1	-
227	rs716980	15	39374094	ncRNA intronic	C	T	0.493	6.98E-11	3.94E-01	RP11-624L4.1	-
227	rs12907993	15	39590057	ncRNA intronic	G	A	0.466	1.00E-10	5.06E-01	RP11-624L4.1	-
227	rs5812099	15	39633877	ncRNA intronic	AG	A	0.090	2.15E-171	2.63E-06	RP11-624L4.1	9
228	rs4046625	15	58256296	intronic	T	C	0.463	3.16E-08	6.10E-03	ALDH1A2	-
229	rs12438742	15	61947280	ncRNA intronic	G	C	0.429	1.36E-08	2.57E-01	RP11-507B12.1	6
230	rs351175	15	74528985	UTR5	G	T	0.281	1.35E-08	2.11E-01	CCDC33	-
231	rs11072509	15	75114374	ncRNA intronic	G	C	0.266	4.59E-08	6.90E-01	LMAN1L:RP11-414J4.2	-
232	rs4887098	15	79057950	intronic	C	G	0.429	7.08E-14	3.27E-02	ADAMTS7	-
233	rs2342118	15	85693016	intergenic	A	G	0.325	7.28E-30	1.84E-03	PDE8A	6
234	rs2584137	15	87988620	ncRNA intronic	G	A	0.371	2.60E-10	4.34E-01	RP11-648K4.2	-
235	rs6224	15	91423543	intronic	T	G	0.475	9.79E-16	6.63E-01	FURIN	1
236	rs2387280	16	4339684	intergenic	T	A	0.247	2.19E-10	2.98E-02	TFAP4	-
237	rs200541	16	24733141	intergenic	G	A	0.196	1.55E-12	2.86E-01	TNRC6A	-
237	rs72770483	16	24926314	intergenic	G	T	0.065	1.03E-08	8.15E-02	SLC5A11	-
238	rs67264578	16	30879395	intronic	A	G	0.264	4.94E-08	8.05E-01	BCL7C	-
239	rs11649294	16	50091335	ncRNA intronic	T	C	0.205	3.75E-08	9.06E-01	RP11-429P3.3	-
240	16:51185924 TA T	16	51185924	upstream	T	TA	0.203	1.11E-08	8.61E-01	SALL1	-
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	95
240	16:51495932 CAA C	16	51495932	intergenic	C	CAA	0.362	1.21E-08	5.29E-02	RP11-437L7.1	4
241	rs12325224	16	70675363	intronic	T	C	0.441	8.68E-13	3.18E-02	IL34	-
242	rs55845372	16	86606321	intergenic	C	T	0.188	1.25E-09	4.79E-03	FOX1	1
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	45
243	rs4060100	16	87302354	ncRNA intronic	A	C	0.299	2.12E-08	2.25E-01	C16orf95:RP11-178L8.5	-
244	rs9939914	16	89939929	upstream	A	T	0.237	1.36E-12	7.84E-02	TCF25	-
244	rs11138992	16	89992613	intronic	T	G	0.131	1.17E-31	9.85E-04	TUBB3:TUBB3	6
245	rs2131704	17	1970898	intronic	T	C	0.425	8.01E-19	6.48E-03	SMG6	10
245	rs11078865	17	2109109	intronic	A	T	0.371	1.11E-12	1.74E-01	SMG6	3
246	rs12451961	17	7351479	intronic	G	A	0.362	6.65E-13	4.45E-01	CHRN1	-
247	rs76653948	17	9151562	intergenic	G	C	0.177	1.03E-21	8.99E-02	STX8	-
248	rs7210118	17	13047713	intergenic	T	C	0.414	1.54E-24	6.24E-02	RP11-488L1.1	3
249	rs79626005	17	19210026	intronic	A	G	0.080	7.45E-10	3.47E-01	EPN2	-
250	rs4994353	17	28386280	intronic	A	T	0.443	9.63E-36	9.98E-02	EFCAB5	6
251	rs77259033	17	41351400	intronic	A	T	0.324	3.36E-08	6.25E-01	NBR1	-
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	76
252	17:42986266 AAAGG A	17	42986266	intronic	A	AAAGG	0.116	2.36E-10	1.85E-01	GFAP	-
252	rs12936234	17	4317951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	86
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	93
252	17:44259690 T G	17	44259690	intronic	G	T	0.186	2.32E-42	3.85E-01	KANSL1	-
252	rs62074125	17	44852612	intronic	C	A	0.257	3.84E-27	5.19E-02	WNT3	-
252	rs3933653	17	44888742	intronic	T	C	0.265	2.22E-10	3.07E-01	WNT3	-
252	rs56325564	17	45766771	intergenic	A	G	0.480	6.03E-12	5.02E-01	KPNB1	-
253	rs9889674	17	47010008	ncRNA intronic	A	G	0.265	1.78E-08	5.76E-01	SNF8:AC091133.1	-
254	rs11869269	17	57800388	intronic	A	G	0.157	1.21E-09	8.66E-01	VMP1	-
255	rs36090413	17	62984671	intergenic	C	G	0.081	3.35E-10	1.56E-01	AMZ2P1	1
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	68
257	17:68282406 GA G	17	68282406	intergenic	G	GA	0.243	1.56E-12	3.10E-01	CTD-2378E21.1	-
258	rs10652863	17	69179642	ncRNA intronic	CAG	CT	0.418	4.98E-08	7.49E-01	CASC17	-
259	rs770068798	17	73866798	intergenic	G	A	0.183	1.36E-13	4.55E-02	TRIM47	16
260	rs9966170	18	35029506	intronic	G	A	0.360	1.24E-28	5.40E-01	CELF4	1
261	rs13133639	18	40318077	intergenic	A	G	0.277	2.05E-14	8.55E-01	RIT2	-
262	rs630122	18	42429610	intronic	C	G	0.315	1.97E-09	3.95E-01	SETBP1	-
263	18:74537564 TA T	18	74537564	intronic	T	TA	0.171	9.73E-10	2.54E-01	ZNF236	-
264	rs62127993	19	3298646	ncRNA intronic	A	G	0.135	2.58E-10	3.58E-01	AC010649.1	2
265	rs7256295	19	4071814	intergenic	A	G	0.223	6.17E-09	9.02E-01	ZBTB7A	3
266	rs73234	19	4394310	intronic	C	G	0.456	9.25E-09	3.11E-01	SH3GL1	-
267	rs10403210	19	13103638	intergenic	C	T	0.466	4.69E-08	4.17E-02	NFIX	1
268	rs2303696	19	18548884	UTR5	T	A	0.379	2.82E-15	1.16E-01	ISYNA1	-
269	rs2043294	19	32207820	intergenic	C	A	0.178	2.72E-30	3.92E-01	RNA5SP471	-
270	rs3786837	19	39172042	intronic	C	G	0.476	8.88E-10	6.20E-02	ACTN4	-
271	rs429358	19	45411941	exonic	C	T	0.156	1.20E-16	5.37E-02	APOE	11
272	rs1159530	20	7103709	ncRNA intronic	C	G	0.324	3.66E-09	5.44E-01	RP4-764O22.1	-
273	rs16996551	20	15773571	intronic	T	C	0.371	1.94E-17	1.99E-01	MACROD2	-
274	rs6059412	20	32280829	intergenic	T	G	0.292	3.32E-09	5.99E-01	E2F1	-
275	rs143383	20	34025983	UTR5	G	A	0.363	1.31E-24	2.93E-03	GDF5	-
276	rs3746471	20	36841914	exonic	A	G	0.467	4.17E-17	2.22E-01	KIAA1755	-
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	23
277	rs2694896	20	45237143	intronic	G	A	0.385	4.95E-08	3.64E-01	SLC13A3	13
277	rs6011998	20	45269867	intronic	T	C	0.042	9.31E-10	3.66E-01	SLC13A3	9
278	rs66806308	20	46360465	intronic	T	C	0.209	4.12E-08	7.45E-01	SULF2	-
279	rs72626592	20	55528415	intergenic	T	C	0.036	1.58E-09	3.84E-01	AL117380.2	1
280	rs6062264	20	61154871	intronic	T	C	0.279	2.12E-08	4.59E-01	C20orf166	13
281	rs3215453	22	29453442	UTR3	TA	T	0.482	3.90E-08	4.45E-01	ZNRF3	-
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	27
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	47
283	rs5750482	22	38117943	intronic	C	T	0.373	4.18E-09	4.13E-01	NOL12:TRIOBP	-
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	46
284	rs130653	22	39642740	intergenic	T	A	0.236	1.05E-15	4.71E-03	PDGFB	-
285	rs61748567	22	50685332	exonic	A	G	0.017	1.39E-22	3.23E-01	HDAC10:MAPK12	5

7.3 Functional annotations of genomic loci associated with language

To further link human language AC-associated SNPs to genes, the MAGMA software was used to perform genome-wide gene-based association analysis based on the results from mvGWAS (see chapter 2.7 for more details about the gene-based association analysis). A total of 981 genes attained the Bonferroni-corrected significance threshold $p = 2.61e-6(0.05/19,157)$ (Fig. 7.3). Overall, among the genes identified by positional mapping, eQTL mapping and/or chromatin interaction mapping, we found that:

- Fourteen of these genes were previously associated with language FC presented in chapter 6: *AC079807.4*, *CTD-2026C7.1*, *TBC1D5*, *EPHA3*, *RNU6-488P*, *ZIC4*, *FAM172A*, *PLCE1*, *DACT1*, *RP11-624L4.1*, *FURIN*, *KANSL1*, *PIPOX*, *SEZ6*.
- Twenty of these genes have been reported to be associate with left-right asymmetry [Sha et al., 2021]: *AL356458.1*, *MAPRE3*, *CGREF1*, *ZIC4*, *FAM13A*, *NR2F1*, *NCOA7*, *BBS9*, *GALNT12*, *ITIH5*, *DACT1*, *DAAM1*, *SPIRE2*, *TUBB3*, *CRHR1*, *MAPT*, *KANSL1*, *LRRC37A2*, *ARL17A*, *MACROD2*.
- *DCDC2*, *KIAA0319*, and *ACOT13* were associate with dyslexia [Deffenbacher et al., 2004, Eicher et al., 2016, Kaplan et al., 2002], *CNTNAP2* and *FOXP2* were associated with language impairment [Fisher et al., 1998, Vernes et al., 2008].
- Nine other gene were associated with reading and language traits: *FGF18* [Field et al., 2013], *COL4A2* [Skeide et al., 2016], *SETBP1* [Perdue et al., 2019], *ROBO2*, *INSC*, *CAND1*, and *DAPK3* [St Pourcain et al., 2014], .

Please refer to chapter 1.6 for details on these associations. For the proteins encoded by 944 genes (944 out of the 981 genes code for proteins), there were 2007 known pairwise interactions in the STRING database [Szklarczyk et al., 2017], compared with 1391 interactions expected for a random set of this size from the whole proteome ($p = 1.0e-16$). This observation supports the validity of our mvGWAS association findings, as random associations would not lead to such functional clustering. (Fig. 7.4).

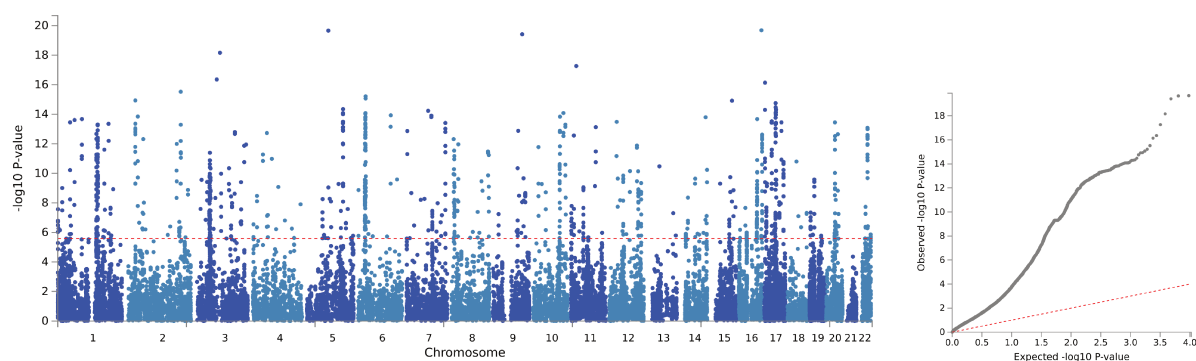


Figure 7.3 – **Genome-wide gene-based association analysis of the anatomical connectivity in 31,753 participants.** genome-wide gene-based association analysis of 313 ACs using mvGWAS results as input. Input SNPs were mapped to 19,157 protein coding genes. The red dashed line indicates the Bonferroni-corrected significance threshold for gene-based analysis $p = 2.61e-6(0.05/19,157)$. The Quantile-quantile plot is also shown.

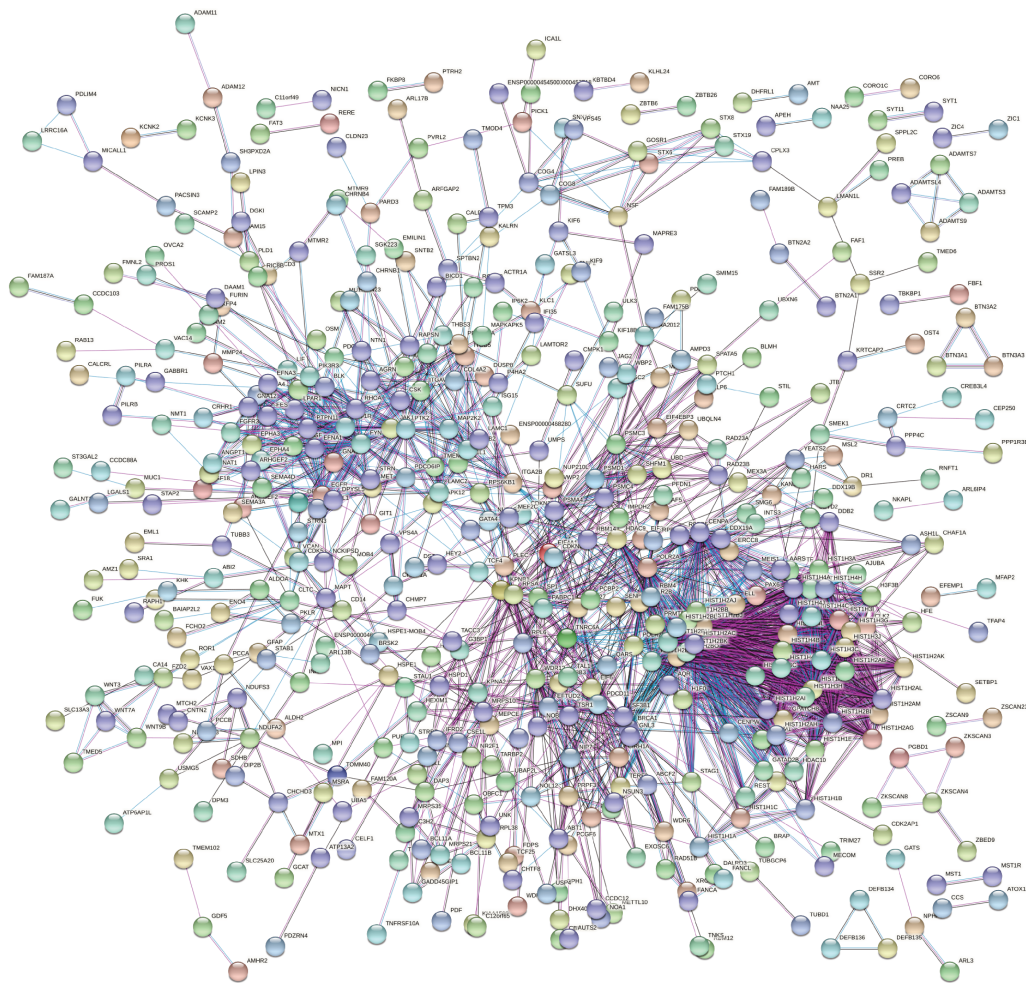


Figure 7.4 – **Human language AC-associated genes integrated into a protein–protein interaction network.** Proteins are represented by nodes. Edges between nodes represent different types of protein–protein interactions according to the STRING [Szklarczyk et al., 2017] database (chapter 2.8), including known interactions (turquoise and dark purple represent interactions identified by curated databases and biological experiments, respectively) and interactions determined by co-expression. Coloured nodes represent the queried proteins. Only at least medium-confidence (> 0.4) links were retained, and disconnected proteins are not shown.

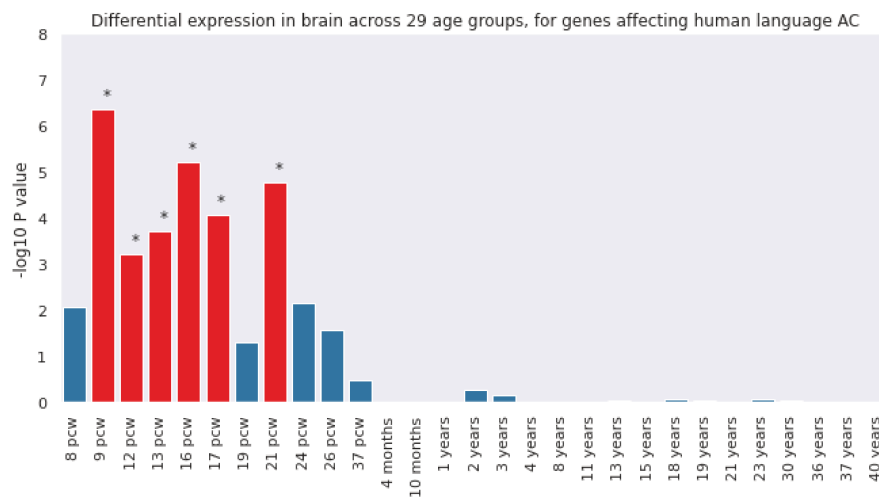
In order to obtain a broader biological context of these gene-level results, we conducted a gene-set analysis using MAGMA [de Leeuw et al., 2015] as implemented in FUMA platform [Watanabe et al., 2017]. We identified ninety-one biological system associated with human language AC differences (Table 7.2).

Table 7.2 – **MAGMA gene-set analysis.** Gene-set attaining statistical significance following Bonferroni control for multiple test.

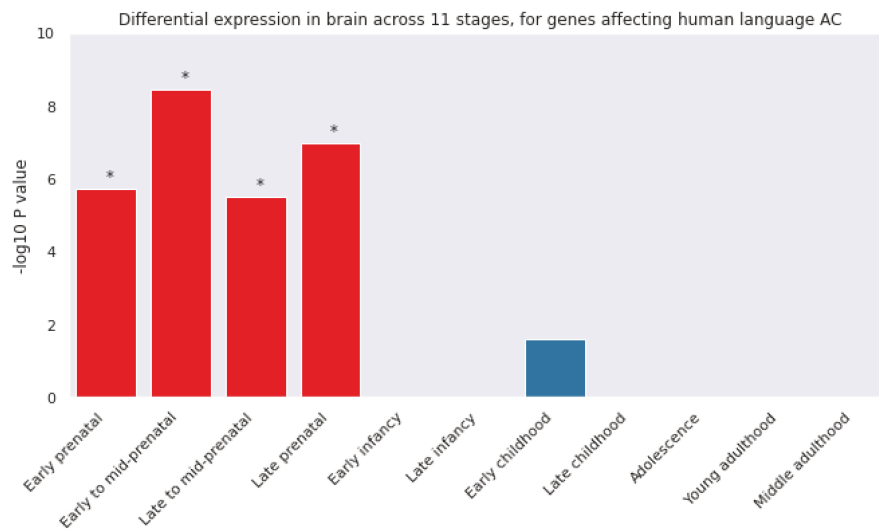
Gene-set Name	Number of genes in gene set	Beta	SE of Beta	P value
GO bp:go neurogenesis	1515	0.246	0.066	2.01E-16
GO bp:go neuron differentiation	1277	0.248	0.062	1.62E-14
GO bp:go cell part morphogenesis	639	0.343	0.062	5.16E-14
GO bp:go cellular component morphogenesis	1061	0.260	0.060	1.16E-13
GO bp:go cell morphogenesis involved in differentiation	688	0.323	0.060	1.25E-13
GO bp:go cell morphogenesis involved in neuron differentiation	546	0.356	0.059	7.06E-13
GO bp:go regulation of cell development	875	0.268	0.056	1.94E-12
GO bp:go regulation of cell differentiation	1723	0.197	0.056	2.28E-12
GO bp:go neuron projection guidance	264	0.481	0.056	8.21E-12
Curated gene sets:dacosta uv response via ercc3 dn	837	0.265	0.054	1.49E-11
GO bp:go regulation of anatomical structure morphogenesis	995	0.239	0.053	3.20E-11
GO bp:go regulation of nervous system development	857	0.254	0.053	4.33E-11
GO bp:go central nervous system development	918	0.250	0.053	4.49E-11
GO bp:go neuron development	1037	0.232	0.052	6.86E-11
GO bp:go cell projection organization	1433	0.196	0.052	9.73E-11
GO bp:go axon development	478	0.332	0.052	1.43E-10
GO bp:go cytoskeleton organization	1221	0.206	0.050	1.85E-10
GO bp:go regulation of neuron differentiation	613	0.287	0.050	2.34E-10
GO bp:go negative regulation of transcription by rna polymerase ii	777	0.253	0.050	4.27E-10
Curated gene sets:kegg axon guidance	124	0.624	0.050	6.39E-10
GO bp:go negative regulation of developmental process	896	0.236	0.050	6.64E-10
GO cc:go cell substrate junction	390	0.339	0.048	8.03E-10
GO bp:go taxis	610	0.291	0.051	1.27E-09
GO bp:go negative regulation of biosynthetic process	1452	0.182	0.048	1.69E-09
GO bp:go embryo development	949	0.225	0.049	1.89E-09
GO bp:go negative regulation of nervous system development	287	0.392	0.048	2.27E-09
Curated gene sets:reactome axon guidance	533	0.285	0.047	2.83E-09
GO bp:go protein phosphorylation	1830	0.163	0.048	3.06E-09
GO bp:go tube development	960	0.218	0.047	3.24E-09
GO bp:go negative regulation of multicellular organismal process	1118	0.200	0.047	3.39E-09
Curated gene sets:dacosta uv response via ercc3 common dn	453	0.309	0.047	5.04E-09
GO bp:go negative regulation of cell differentiation	675	0.253	0.047	5.07E-09
GO bp:go tube morphogenesis	781	0.235	0.046	5.79E-09
GO bp:go regulation of cell morphogenesis	448	0.304	0.046	6.82E-09
GO bp:go head development	726	0.244	0.047	7.51E-09
GO bp:go positive regulation of biosynthetic process	1858	0.157	0.046	8.25E-09
GO bp:go locomotion	1693	0.164	0.047	8.85E-09
GO cc:go anchoring junction	526	0.277	0.045	9.00E-09
GO bp:go negative regulation of rna biosynthetic process	1170	0.187	0.045	1.81E-08
GO bp:go regulation of phosphorus metabolic process	1563	0.166	0.045	1.82E-08
Curated gene sets:kegg pathways in cancer	318	0.324	0.041	3.03E-08
GO bp:go negative regulation of cell development	310	0.346	0.044	3.41E-08
GO bp:go regulation of intracellular signal transduction	1695	0.155	0.044	4.54E-08
Curated gene sets:charafe breast cancer luminal vs mesenchymal dn	442	0.290	0.044	5.68E-08
GO bp:go regulation of cell projection organization	628	0.231	0.041	7.01E-08
Curated gene sets:cui tcf21 targets 2 dn	791	0.214	0.043	7.97E-08
GO bp:go animal organ morphogenesis	1002	0.193	0.043	8.63E-08
GO bp:go positive regulation of signaling	1707	0.148	0.042	1.08E-07
GO bp:go positive regulation of developmental process	1269	0.167	0.041	1.35E-07
Curated gene sets:gobert oligodendrocyte differentiation dn	1023	0.186	0.042	1.42E-07
GO bp:go positive regulation of rna biosynthetic process	1518	0.156	0.042	1.51E-07
Curated gene sets:reactome semaphorin interactions	61	0.675	0.038	1.56E-07
GO bp:go regulation of protein modification process	1700	0.147	0.042	1.57E-07

Gene-set Name	Number of genes in gene set	Beta	SE of Beta	P value
GO bp:go cardiovascular system development	669	0.227	0.042	1.70E-07
GO bp:go positive regulation of gene expression	1837	0.142	0.042	1.70E-07
GO mf:go gtpase binding	502	0.261	0.042	2.06E-07
Curated gene sets:blalock alzheimers disease up	1569	0.146	0.040	2.68E-07
GO bp:go cell surface receptor signaling pathway involved in cell cell signaling	582	0.234	0.040	2.87E-07
GO bp:go regulation of neuron projection development	460	0.260	0.040	3.29E-07
Curated gene sets:pid fak pathway	57	0.708	0.039	3.63E-07
GO bp:go positive regulation of cellular component organization	1126	0.170	0.040	3.66E-07
Curated gene sets:martinez response to trabectedin dn	265	0.344	0.040	3.78E-07
GO bp:go cell cell signaling by wnt	486	0.249	0.039	4.05E-07
GO bp:go embryonic organ development	411	0.280	0.041	4.90E-07
GO bp:go developmental growth	604	0.223	0.039	5.10E-07
GO mf:go small gtpase binding	409	0.277	0.040	5.18E-07
GO bp:go embryo development ending in birth or egg hatching	603	0.226	0.040	6.05E-07
GO bp:go positive regulation of cell development	501	0.242	0.039	6.15E-07
GO bp:go negative regulation of neuron differentiation	210	0.374	0.039	7.23E-07
GO bp:go circulatory system development	1000	0.177	0.039	7.38E-07
GO bp:go regulation of cytoskeleton organization	490	0.242	0.038	8.39E-07
GO bp:go cell motility	1475	0.147	0.039	9.08E-07
GO bp:go pattern specification process	420	0.274	0.040	1.02E-06
GO mf:go cytoskeletal protein binding	905	0.177	0.038	1.07E-06
Curated gene sets:iked a mir30 targets up	110	0.542	0.041	1.08E-06
GO bp:go pulmonary valve morphogenesis	16	1.179	0.034	1.09E-06
GO bp:go small gtpase mediated signal transduction	521	0.235	0.038	1.27E-06
Curated gene sets:lim mammary luminal mature dn	93	0.570	0.040	1.37E-06
GO bp:go positive regulation of intracellular signal transduction	943	0.178	0.039	1.55E-06
Curated gene sets:reactome crmps in sema3a signaling	15	1.244	0.035	1.59E-06
Curated gene sets:reactome diseases of signal transduction	376	0.267	0.037	1.61E-06
GO bp:go negative regulation of cellular component organization	638	0.210	0.038	1.63E-06
Curated gene sets:vanharanta uterine fibroid up	40	0.852	0.039	1.71E-06
Curated gene sets:graessmann apoptosis by doxorubicin dn	1638	0.137	0.038	1.71E-06
Curated gene sets:reactome developmental biology	1073	0.177	0.041	1.82E-06
Curated gene sets:nikolsky breast cancer 16q24 amplicon	53	1.174	0.062	2.00E-06
GO bp:go artery development	84	0.554	0.037	2.26E-06
GO bp:go anatomical structure formation involved in morphogenesis	1034	0.165	0.037	2.38E-06
GO bp:go forebrain development	362	0.273	0.037	2.72E-06
GO bp:go regulation of cell morphogenesis involved in differentiation	276	0.312	0.037	2.76E-06
GO mf:go rho gtpase binding	159	0.403	0.037	2.97E-06

We also conducted a gene-expression analysis using data from the BrainSpan [Miller et al., 2014] database, from either 29 age groups or 11 defined general developmental stages of brain samples. We found relatively higher mRNA expression of human language AC associated genes during early-prenatal ($p = 1.94e-6$), early-mid-prenatal ($p = 3.59e-9$), Late mid-prenatal ($p = 3.11e-6$), Late prenatal ($p = 1.07e-7$), from 9 ($p = 4.32e-7$) to 21 ($p = 1.73e-5$) post-conceptual weeks (FDR-corrected p values < 0.05) (Fig. 7.5a, 7.5b).



A)



B)

Figure 7.5 – **MAGMA gene-property analysis.** A) Relation between gene-based association with human language AC and relatively higher mRNA expression in the human brain at particular ages, using BrainSpan data from 29 age groups. Asterisks indicate significant age groups meeting $p < 0.05$ with FDR correction. pcw: post-conceptual weeks. B) Relation between gene-based association with human language AC and relatively higher mRNA expression in the human brain at particular ages, using BrainSpan data from 11 developmental stages. Asterisks indicate significant groups meeting $p < 0.05$ with FDR correction.

7.4 Discussion

In this study, we extracted language endophenotypes, namely the anatomical connectivities, from diffusion-weighted acquisitions in the largest imaging-genetic cohort in general population to date. We targeted language white matter tract identified by [Rojkova et al., 2016]. After filtering on heritability significance, we performed a multivariate GWAS in order to take advantage of the correlation structure among AC endophenotypes and uncover potential pleiotropic loci. This analysis identified 285 distinct genetic loci associated with different aspects of adult human brain anatomical connectivity.

Thereby we highlighted potential key genes related to language processing including ones that were previously associated with language impairment through candidate genes and linkage analysis.

Altogether, these results provide a novel insight into the genetic architecture of the language in humans. Functional annotation implicated genes particularly involved in different brain organisation including neurogenesis and neuron differentiation as well as in prenatal brain development. Overall, this preliminary study, which highly calls for replication, brings interesting leads regarding the molecular foundation of human brain anatomical connectivity.

* * *
* *
*

General Conclusions

Human language is a complex trait. Different types of genetic variants are expected to affect its variability. This thesis aimed to bring a modest contribution by proposing the genetic underpinning of individual differences in image derived language-related endophenotypes.

Summary

In chapter 4, we provided an in-depth study of human language heritability of the phenotypes that we chose carefully. In particular, we evaluated the heritability of both psychometric and MRI derived endophenotypes on two large-scale resources currently available; the Human Connectome Project (HCP) and the UK Biobank (UKB) using state-of-the-art approaches. These endophenotypes include resting-state functional pairwise interaction, structural connectivity and, language-brain activation. From the Human Connectome Project, we observed a hierarchy of language heritability estimates with 19.4-71.7% for language-behavioural scores, 11.4-48.7% for resting-state FC, and 14.4-34.9% for language-task activations. In terms of heritability ranking, the resting-state functional connectivity edges show low levels of heritability compared to behavioural scores and structural features in both general population and twin based studies. This observations are consistent with heritabilities obtained on general population by [Elliott et al., 2018] who reported low levels of SNP heritability for the resting-state FC edges compared to structural and diffusion MRI endophenotypes. Regarding inter-cohort variability, we observed, in line with the literature, higher values of heritability in pedigree studies compared to SNP-based heritability study. The possible origins of these discrepancies were investigated according to an experimental design that takes into account dissimilarities between both cohort with no success.

We were particularly interested by the resting-state data and its relevance for studying specific brain functions such as language. This type of data offers multiple advantages as it is paradigm free and as such it is easier to implement in very large cohorts like the UK Biobank. It has certain interests over task-based design, specially when one knows that certain populations have difficulties to engage in specific tasks. Although resting state fMRI data are reported as heritable, little is known about their relation to the language-related cognitive process.

In chapter 5, we address this question regarding the shared genetics between these ones and direct measures of language as cognitive process. To do so, we estimated both phenotypic and genetic correlations and between cognitive abilities, neural activations and task-free data in language. To this end, we took advantage of the Human Connectome Project cohort which comprises language-related scores, language task-based fMRI data as well as resting-state fMRI data. The results show a shared aetiology between resting-state functional connectivity and neural activations. This work advocates for the existence of a link between resting-state and task-based data both phenotypically and genetically regarding human language brain organisation.

In chapter 6, we contributed to enlighten the genetic architecture of language functional connectivity by exhibiting potential key genes related to language processing. We extracted language endophenotypes from rsfMRI acquisitions in the largest imaging-genetic cohort on general population to date (UK Biobank). To make these endophenotypes as closer as possible to language function in the brain, we adopted an ROI-based approach for FC estimation with language ROIs derived from an comprehensive meta-analysis of tbfMRI studies on language. This approach makes the endophenotypes comparable across individuals and thus suitable for an association study. After filtering on heritability significance, we performed a multivariate GWAS technique in order to take advantage of the correlation structure among rsfMRI FC and uncover potential pleiotropic loci. Thereby, we highlighted potential key genes related to language processing including:

- *EPHA3* gene in the 3p11.1 locus with a role in the *fronto-temporal semantic* network,
- *THBS1* gene, modulated by the 15q14 locus, associated with the functional connectivity part of the *perceptual motor* interaction required for language, and
- *PLCE1* gene in the 10q23.33 locus with a potential role in the *bilateral fronto-temporal auditory-motor* network.

These genes could be prioritised to study language in suitable genetic model. A growing number of works claims that language studies should consider task-free fMRI data in general population in order to focus on the neurobiological organisation of language and how it supports natural language, i.e. as it is used in everyday life [Hasson et al., 2018]. On a clinical side, such research settings are highly desirable as there is a need to map language areas in patients unable to perform language tasks [Branco et al., 2016, Klingbeil et al., 2019, Park et al., 2020, Ramage et al., 2020]. By validating our findings in rsfMRI FC endophenotypes with dMRI-derived endophenotypes, we tried to provide both a functional and structural connectome point of view of the language organisation in the brain. Finally, recent imaging-genetic approaches such as multivariate GWAS, allowed us to increase in statistical power and circumvent the small effect sizes of these original endophenotypes. We believe that such approaches are promising and will broadly disseminate in the imaging-genetic field and beyond. With the presented approach, we tried to contribute to these innovative trends and to pave the way to other alternative task-free approaches to study natural language and its genetic underpinnings.

In chapter 7, we adopted a similar experimental design than the one presented in the chapter 6 to study the genetic architecture of anatomical connectivity. We extracted language endophenotypes from diffusion-weighted acquisitions using UK Biobank cohort. We adopted a ROI-based approach for AC estimation with well-known language white matter tracts. After filtering on heritability significance, we performed a multivariate GWAS technique in order to take advantage of the correlation structure among structural connectivity and uncover potential pleiotropic loci. Thereby, we highlighted potential key genes related to language processing. These genes could be prioritised to study language in suitable genetic model. The findings suggested new candidates genes which connect with biological pathways that are already suspected to be important for human brain-language. Overall, these results provide a novel insight into the genetic architecture of the language in humans.

Discussion, limitations

An integrative approach. We have followed an integrative approach in our study of language genetic architecture. First, by validating our findings in rsfMRI FC endophenotypes with dMRI-derived endophenotypes, we tried to provide both a functional and structural connectome point of view of the language organisation in the brain. Second, we used recent imaging-genetic association approaches such as multivariate GWAS, that allowed us to increase in statistical power and circumvent the small effect sizes of these original endophenotypes. We believe that such approaches are promising and will broadly disseminate in the imaging-genetic field and beyond. With the presented approach, we tried to contribute to these innovative trends and to pave the way to other alternative task-free approaches to study natural language and its genetic underpinnings.

A study in general population. We conducted our study on a general population cohort by selecting state-of-the-art language (endo)phenotypes. We believe it provides an interesting alternative framework to studies based on language impairment. General population cohorts provide a wealth of measurements that make possible to develop (endo)phenotypes of interest, including for complex cognitive processes. This approach offers a promising alternative to case-control paradigm.

Conclusion

Language is a central component of human life. Understanding how we are able to acquire and use this highly complex trait has required the development of numerous language descriptions at multiple levels. On the one hand, we leveraged well characterised language processing scores and intermediate phenotypes. On the other hand, we applied association methods to identify genes that relate to the behavioural differences, and understand how they impact the development and function of the brain by looking at the cellular function of the proteins they encode and the biological pathways in which they are involved. In order to build a comprehensive account

of the human language, it will be necessary to integrate the results from the different strategies undertaken and levels of description.

In this thesis, we approached the question of the genetic basis of human language using the state-of-art genomic strategies: by highlighting the brain inter-individual variability and evaluate the role of common SNPs regarding the whole genome with no *a priori* on potential candidate genes. We related these investigations to current knowledge in language-related cognitive performance and to molecular mechanisms known to occur in brain and development. The findings from this thesis provide new entry-points for future work bridging the gap between our genetic make-up and human language processing.

* * *
* *
*

Appendices

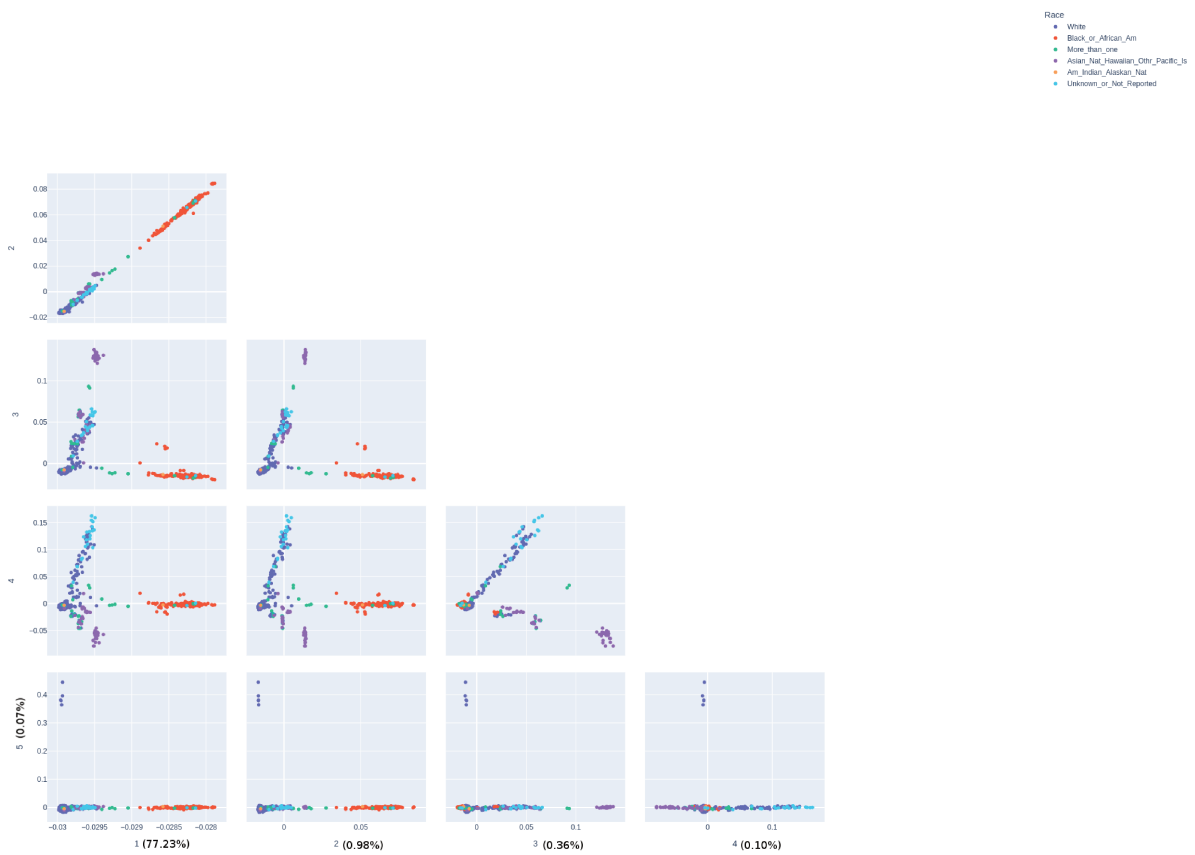
A.1 Population structure of Human connectome project data

Figure A.1 – Overview of the first five principal components of the principal component analysis of the genotype data. We can observe that the fifth principal component separate individual from the white ethnicity.

B.1 Pedigree-based heritability of behavioural and neuroimaging (endo)phenotypes: a twin study

Table B.1 – Heritabilities of brain functional connectivities extracted using SmSSPAs ROIs. Bold phenotypes represents FCs with significant heritability after false discovery rate correction at $p < 0.05$.

Endophenotype	heritability	P value	SE	FDR P value
F3tv↔F3orb	48.8%	1.61E-13	0.060	4.82E-11
Pole↔AG	42.9%	1.19E-10	0.066	1.79E-08
Prec↔F3td	43.2%	5.69E-10	0.066	5.69E-08
T2ml↔STSp	38.7%	2.79E-09	0.064	1.96E-07
AG↔F3orb	38.9%	3.93E-09	0.066	1.96E-07
STSp↔F2antR	37.3%	3.48E-09	0.063	1.96E-07
F3opd↔SMG	37.6%	4.68E-09	0.064	2.00E-07
Fusa↔AG	38.8%	5.36E-09	0.068	2.01E-07
T1a↔STSp	34.0%	8.89E-09	0.059	2.96E-07
T3p↔F2antR	36.2%	1.27E-08	0.063	3.82E-07
PT↔T1a	35.0%	1.79E-08	0.061	4.89E-07
AG↔T1a	34.9%	2.00E-08	0.063	5.00E-07
Pole↔F3opd	42.8%	3.29E-08	0.077	7.58E-07
Pole↔T1a	38.9%	4.57E-08	0.072	9.79E-07
Prec↔F3opd	33.4%	1.00E-07	0.063	1.87E-06
T2ml↔F3orb	36.6%	1.00E-07	0.069	1.87E-06
T2p↔RoIS	33.1%	2.00E-07	0.064	3.53E-06
PT↔T1a.HeschIR	33.2%	3.00E-07	0.066	5.00E-06
F3tv↔F3td	32.3%	4.00E-07	0.067	5.71E-06
PT↔T1	33.4%	4.00E-07	0.067	5.71E-06
AG↔F2antR	33.2%	4.00E-07	0.068	5.71E-06
RoIS↔T1R	35.0%	5.00E-07	0.069	6.52E-06
Pole↔F3tv	33.4%	5.00E-07	0.067	6.52E-06
Fusa↔T3p	34.4%	6.00E-07	0.070	7.50E-06
AG↔F3tv	34.3%	7.00E-07	0.073	8.08E-06
T1a↔T1a.HeschIR	34.4%	7.00E-07	0.072	8.08E-06
Pole↔T1	32.7%	8.00E-07	0.068	8.89E-06
F3td↔T3p	32.4%	1.00E-06	0.068	1.03E-05
T1a↔F3td	33.0%	1.00E-06	0.070	1.03E-05
F3tv↔F2antR	29.5%	1.20E-06	0.063	1.20E-05
PT↔F3tv	29.0%	1.80E-06	0.063	1.74E-05
F3opd↔T2pR	28.2%	2.50E-06	0.062	2.34E-05
Pole↔T2ml	30.2%	2.60E-06	0.067	2.36E-05
F3opd↔T3p	30.0%	2.70E-06	0.065	2.38E-05
T1↔F3tv	30.9%	2.90E-06	0.068	2.49E-05
T2ml↔SMG	28.6%	3.90E-06	0.065	3.24E-05
F3tv↔SMG	30.3%	4.00E-06	0.069	3.24E-05
F3opd↔AG	29.7%	4.20E-06	0.068	3.32E-05
Pole↔F3orb	31.7%	4.60E-06	0.071	3.45E-05
PrF3op↔SMG	28.9%	4.60E-06	0.065	3.45E-05
Prec↔AG	29.6%	6.10E-06	0.068	4.46E-05
F3td↔F3orb	29.5%	8.80E-06	0.071	6.29E-05

Endophenotype	heritability	P value	SE	FDR P value
F3tv↔STSp	29.0%	1.08E-05	0.069	7.53E-05
F3opdR↔T2pR	27.7%	1.11E-05	0.066	7.57E-05
T1↔T1a.HeschlR	29.4%	1.43E-05	0.071	9.53E-05
PrF3op↔F3tv	29.7%	1.57E-05	0.074	9.65E-05
T1a↔T2pR	28.5%	1.64E-05	0.070	9.65E-05
Prec↔F2antR	25.0%	1.61E-05	0.062	9.65E-05
Fusa↔T2p	28.8%	1.61E-05	0.070	9.65E-05
T2ml↔T1	29.1%	1.58E-05	0.071	9.65E-05
F2p↔STSp	26.8%	1.57E-05	0.065	9.65E-05
RoIS↔PrecR	27.4%	1.69E-05	0.066	9.75E-05
T1↔T2p	27.9%	1.78E-05	0.069	1.01E-04
T2p↔F2antR	27.4%	1.86E-05	0.070	1.03E-04
Pole↔F3td	27.3%	1.96E-05	0.068	1.07E-04
PrF3op↔T3p	27.3%	2.49E-05	0.069	1.33E-04
Pole↔T1R	28.1%	2.78E-05	0.070	1.46E-04
T2ml↔T1a.HeschlR	26.1%	2.88E-05	0.066	1.49E-04
F2p↔PrF3op	26.9%	3.29E-05	0.069	1.64E-04
T1↔T3p	26.1%	3.27E-05	0.066	1.64E-04
Fusa↔T1a	27.4%	3.72E-05	0.070	1.80E-04
F3opd↔F3orb	28.5%	3.70E-05	0.072	1.80E-04
F3td↔F2antR	26.1%	3.86E-05	0.067	1.84E-04
SMG↔F3opdR	25.0%	3.97E-05	0.065	1.84E-04
AG↔STSp	23.6%	3.99E-05	0.061	1.84E-04
STSp↔F3td	27.5%	4.19E-05	0.071	1.88E-04
F2p↔F3tv	26.8%	4.13E-05	0.070	1.88E-04
T2p↔T3p	24.9%	4.67E-05	0.065	2.06E-04
Pole↔Fusa	27.8%	4.76E-05	0.074	2.06E-04
F3opd↔F2antR	25.6%	4.81E-05	0.067	2.06E-04
F3opd↔T2ml	25.9%	4.97E-05	0.068	2.10E-04
STSp↔F3orb	28.7%	5.42E-05	0.077	2.26E-04
T3p↔F3orb	23.9%	5.75E-05	0.064	2.36E-04
STSp↔SMG	25.5%	6.14E-05	0.069	2.49E-04
F2p↔F3orb	26.0%	7.07E-05	0.069	2.83E-04
SMG↔F3orb	26.9%	7.45E-05	0.073	2.94E-04
T1a↔T3p	23.3%	8.05E-05	0.062	3.14E-04
T1↔RoIS	24.6%	8.51E-05	0.067	3.27E-04
PrF3op↔T1a	27.7%	9.91E-05	0.075	3.76E-04
F3opd↔F3tv	25.2%	1.03E-04	0.071	3.86E-04
F2antR↔T1a.HeschlR	24.4%	1.06E-04	0.067	3.92E-04
T1a↔T2p	23.7%	1.20E-04	0.066	4.39E-04

Endophenotype	heritability	P value	SE	FDR P value
T1↔T1a	23.7%	1.23E-04	0.066	4.40E-04
T2ml↔F3tv	26.4%	1.24E-04	0.073	4.40E-04
Pole↔STSp	25.4%	1.25E-04	0.070	4.40E-04
T1a↔RoIS	25.3%	1.29E-04	0.070	4.44E-04
SMG↔F3td	23.6%	1.28E-04	0.067	4.44E-04
Prec↔SMG	24.1%	2.08E-04	0.070	7.09E-04
F2p↔SMG	25.5%	2.27E-04	0.073	7.65E-04
F3opd↔F3td	24.4%	2.39E-04	0.070	7.97E-04
T1a↔T1R	24.5%	2.68E-04	0.072	8.85E-04
T3p↔T2pR	21.9%	2.83E-04	0.065	9.14E-04
T3p↔RoIS	23.6%	2.80E-04	0.070	9.14E-04
AG↔SMG	22.0%	2.93E-04	0.067	9.34E-04
PrF3op↔T1R	24.3%	3.16E-04	0.072	9.99E-04
STSp↔T2pR	24.6%	3.84E-04	0.076	1.20E-03
T1↔PrF3op	21.9%	3.91E-04	0.067	1.21E-03
F3opd↔T1	21.4%	4.07E-04	0.065	1.24E-03
T2pR↔T1R	23.1%	4.40E-04	0.071	1.33E-03
Prec↔F2p	21.2%	4.67E-04	0.066	1.40E-03
F3tv↔T1a.HeschlR	23.3%	5.03E-04	0.074	1.49E-03
T2ml↔AG	22.0%	5.08E-04	0.068	1.49E-03
F3opd↔STSp	24.1%	5.40E-04	0.075	1.57E-03
T1↔Fusa	23.0%	5.53E-04	0.072	1.58E-03
T2ml↔T1a	21.6%	5.53E-04	0.067	1.58E-03
T1↔T1R	23.2%	6.11E-04	0.073	1.73E-03
T2ml↔PrF3op	23.2%	6.30E-04	0.074	1.77E-03
Prec↔T1	20.7%	6.70E-04	0.066	1.86E-03
F3tv↔RoIS	21.8%	6.89E-04	0.070	1.90E-03
T2p↔F3td	21.1%	7.00E-04	0.068	1.91E-03
T1↔T2pR	22.3%	7.08E-04	0.071	1.91E-03
PrF3op↔RoIS	20.0%	7.18E-04	0.064	1.92E-03
F2p↔T2ml	24.1%	7.35E-04	0.078	1.94E-03
Prec↔T2p	21.5%	7.36E-04	0.069	1.94E-03
AG↔T3p	22.4%	7.73E-04	0.073	2.02E-03
PT↔T1R	21.8%	7.91E-04	0.070	2.05E-03
SMG↔T2p	20.5%	8.25E-04	0.068	2.10E-03
AG↔F3td	22.5%	8.38E-04	0.074	2.10E-03
T2ml↔RoIS	21.2%	8.42E-04	0.069	2.10E-03
T1↔STSp	21.8%	8.39E-04	0.070	2.10E-03
SMG↔T3p	19.6%	8.82E-04	0.064	2.19E-03
F3tv↔T1R	20.1%	9.19E-04	0.066	2.26E-03

Endophenotype	heritability	P value	SE	FDR P value
T2ml↔F3td	21.4%	9.77E-04	0.071	2.36E-03
Prec↔RoIS	22.3%	9.75E-04	0.073	2.36E-03
SMG↔F2antR	21.7%	1.01E-03	0.073	2.42E-03
F3orb↔F2antR	21.7%	1.02E-03	0.073	2.44E-03
PT↔F2antR	18.9%	1.05E-03	0.064	2.48E-03
T2ml↔T2pR	22.2%	1.11E-03	0.075	2.60E-03
F3orb↔F3opdR	18.4%	1.12E-03	0.061	2.61E-03
SMG↔T1a.HeschlR	19.7%	1.16E-03	0.067	2.64E-03
T1a↔SMG	19.5%	1.15E-03	0.066	2.64E-03
AG↔RoIS	20.5%	1.16E-03	0.070	2.64E-03
PrF3op↔T2pR	21.9%	1.26E-03	0.075	2.84E-03
F3opd↔T2p	21.3%	1.27E-03	0.073	2.84E-03
T2p↔T1a.HeschlR	19.5%	1.29E-03	0.067	2.87E-03
Prec↔T1R	20.1%	1.33E-03	0.068	2.94E-03
F3opd↔F2p	20.2%	1.35E-03	0.069	2.95E-03
F2p↔Fusa	19.1%	1.40E-03	0.066	3.03E-03
F2p↔T1a	22.8%	1.64E-03	0.079	3.55E-03
T2ml↔PrecR	21.1%	1.71E-03	0.074	3.66E-03
Fusa↔F3tv	20.9%	1.98E-03	0.076	4.22E-03
T1a↔F2antR	20.0%	2.01E-03	0.072	4.24E-03
Pole↔PrF3op	19.1%	2.03E-03	0.069	4.25E-03
F3opd↔T1R	18.5%	2.07E-03	0.067	4.32E-03
T1↔AG	19.0%	2.12E-03	0.068	4.38E-03
SMG↔RoIS	18.3%	2.17E-03	0.067	4.42E-03
Fusa↔SMG	19.3%	2.16E-03	0.069	4.42E-03
Pole↔T2pR	20.6%	2.27E-03	0.074	4.60E-03
AG↔T2p	18.9%	2.42E-03	0.069	4.88E-03
PrF3op↔F3orb	19.9%	2.56E-03	0.074	5.12E-03
F2p↔T1	18.6%	2.62E-03	0.069	5.21E-03
T1↔SMG	17.0%	2.68E-03	0.063	5.29E-03
Pole↔F2p	17.8%	3.01E-03	0.067	5.91E-03
Fusa↔F3td	18.8%	3.13E-03	0.070	6.05E-03
F3opd↔RoIS	19.7%	3.11E-03	0.074	6.05E-03
T2ml↔Fusa	17.7%	3.17E-03	0.066	6.10E-03
T1↔F3orb	17.8%	3.31E-03	0.067	6.32E-03
Pole↔T2p	18.3%	3.40E-03	0.069	6.45E-03
Fusa↔F3orb	18.5%	3.42E-03	0.072	6.45E-03
Pole↔T3p	19.3%	3.61E-03	0.074	6.77E-03
F3orb↔T1R	16.9%	3.79E-03	0.065	7.06E-03
Pole↔PT	16.7%	4.13E-03	0.065	7.64E-03
Prec↔STSp	19.6%	4.40E-03	0.076	8.10E-03
F2antR↔T1R	16.2%	4.61E-03	0.065	8.38E-03

Endophenotype	heritability	P value	SE	FDR P value
PT↔RoIS	17.1%	4.58E-03	0.068	8.38E-03
T2p↔T1R	19.5%	4.74E-03	0.077	8.57E-03
T2ml↔T3p	16.8%	4.78E-03	0.067	8.59E-03
T2ml↔F2antR	16.4%	4.90E-03	0.066	8.75E-03
RoIS↔T1a.HeschIR	17.5%	5.02E-03	0.069	8.90E-03
Prec↔T2ml	16.9%	5.06E-03	0.068	8.92E-03
STSp↔T3p	15.5%	5.25E-03	0.063	9.20E-03
Prec↔T1a	16.7%	5.36E-03	0.068	9.34E-03
F3opd↔PrF3op	18.2%	5.40E-03	0.074	9.37E-03
F3tv↔PrecR	16.9%	5.65E-03	0.068	9.74E-03
F3tv↔F3opdR	17.7%	6.12E-03	0.072	1.04E-02
F2p↔F2antR	16.2%	6.11E-03	0.067	1.04E-02
Prec↔PT	15.4%	6.42E-03	0.064	1.09E-02
RoIS↔T2pR	17.1%	6.71E-03	0.071	1.13E-02
T2ml↔F3opdR	15.8%	6.84E-03	0.067	1.15E-02
F3opd↔T1a	19.5%	7.00E-03	0.082	1.17E-02
Fusa↔STSp	17.0%	7.22E-03	0.071	1.19E-02
T1↔F3opdR	16.0%	7.19E-03	0.067	1.19E-02
T1a↔F3opdR	15.6%	7.85E-03	0.066	1.28E-02
STSp↔T1a.HeschIR	16.5%	7.82E-03	0.070	1.28E-02
F3opdR↔T1R	15.3%	8.58E-03	0.065	1.39E-02
AG↔T1R	15.5%	9.44E-03	0.069	1.52E-02
T1a.HeschIR↔PrecR	15.1%	9.74E-03	0.067	1.56E-02
T1↔F3td	15.2%	9.85E-03	0.066	1.57E-02
Prec↔T2pR	16.2%	1.02E-02	0.071	1.63E-02
Prec↔Fusa	16.0%	1.03E-02	0.071	1.63E-02
F3td↔PrecR	13.8%	1.29E-02	0.064	2.03E-02
STSp↔PrecR	15.3%	1.32E-02	0.072	2.06E-02
AG↔T1a.HeschIR	15.5%	1.32E-02	0.072	2.06E-02
F2p↔T1R	15.1%	1.40E-02	0.071	2.17E-02
F3opdR↔F2antR	14.7%	1.49E-02	0.069	2.27E-02
Fusa↔PrF3op	13.4%	1.48E-02	0.063	2.27E-02
PrF3op↔STSp	14.5%	1.49E-02	0.069	2.27E-02
Pole↔F2antR	15.5%	1.54E-02	0.074	2.33E-02
F3orb↔PrecR	14.5%	1.56E-02	0.069	2.36E-02
SMG↔T2pR	14.0%	1.65E-02	0.068	2.48E-02
PrF3op↔F2antR	15.0%	1.67E-02	0.072	2.49E-02
PrF3op↔F3opdR	15.9%	1.72E-02	0.077	2.56E-02
F3td↔T2pR	14.3%	1.73E-02	0.070	2.56E-02
PT↔Fusa	13.4%	1.84E-02	0.066	2.71E-02

Endophenotype	heritability	P value	SE	FDR P value
AG↔T2pR	14.7%	1.94E-02	0.073	2.83E-02
Pole↔T1a.HeschlR	14.4%	2.12E-02	0.073	3.08E-02
Fusa↔F3opdR	13.0%	2.30E-02	0.067	3.33E-02
SMG↔T1R	12.9%	2.60E-02	0.068	3.75E-02
PT↔STSp	12.5%	2.66E-02	0.066	3.82E-02
Prec↔T3p	12.8%	2.69E-02	0.069	3.85E-02
RoIS↔F2antR	13.7%	2.92E-02	0.074	4.15E-02
AG↔F3opdR	11.6%	2.98E-02	0.063	4.21E-02
F3tv↔T2pR	12.9%	3.00E-02	0.071	4.23E-02
F2p↔F3td	11.5%	3.08E-02	0.063	4.32E-02
Prec↔F3orb	11.7%	3.10E-02	0.064	4.33E-02
T2ml↔T2p	11.6%	3.21E-02	0.065	4.46E-02
STSp↔RoIS	11.8%	3.25E-02	0.066	4.49E-02
PrF3op↔F3td	11.9%	3.95E-02	0.070	5.44E-02
PT↔AG	10.7%	4.27E-02	0.066	5.85E-02
F2p↔AG	11.6%	4.39E-02	0.070	5.99E-02
PT↔T2p	11.7%	4.47E-02	0.071	6.04E-02
F3tv↔T2p	11.9%	4.45E-02	0.072	6.04E-02
T1a.HeschlR↔T1R	12.6%	4.61E-02	0.077	6.18E-02
T1↔PrecR	11.9%	4.61E-02	0.072	6.18E-02
Fusa↔T1a.HeschlR	10.6%	4.66E-02	0.065	6.22E-02
F2p↔RoIS	11.3%	4.98E-02	0.070	6.62E-02
STSp↔T2p	11.3%	5.01E-02	0.071	6.62E-02
T3p↔T1R	10.7%	5.07E-02	0.066	6.67E-02
Pole↔SMG	11.4%	5.14E-02	0.072	6.74E-02
T2pR↔T1a.HeschlR	9.4%	5.25E-02	0.059	6.85E-02
T2ml↔T1R	10.2%	6.83E-02	0.070	8.87E-02
F3tv↔T1a	10.4%	7.18E-02	0.073	9.29E-02
Prec↔T1a.HeschlR	9.5%	7.62E-02	0.068	9.81E-02
Prec↔PrF3op	9.8%	7.72E-02	0.071	9.90E-02
T1↔F2antR	9.2%	8.23E-02	0.068	1.05E-01
F2p↔T2pR	9.3%	8.61E-02	0.070	1.09E-01
PrF3op↔AG	10.1%	8.67E-02	0.076	1.10E-01
Fusa↔T2pR	9.0%	8.77E-02	0.067	1.11E-01
Prec↔Pole	8.9%	9.24E-02	0.069	1.16E-01
Fusa↔RoIS	10.1%	9.31E-02	0.079	1.16E-01
F3orb↔T1a.HeschlR	8.4%	9.77E-02	0.067	1.22E-01
STSp↔T1R	9.3%	1.02E-01	0.074	1.25E-01
T2p↔F3orb	7.6%	1.01E-01	0.062	1.25E-01

Endophenotype	heritability	P value	SE	FDR P value
T3p↔F3opdR	7.7%	1.02E-01	0.062	1.26E-01
F2antR↔T2pR	7.7%	1.08E-01	0.063	1.32E-01
F2p↔T1a.HeschIR	7.9%	1.17E-01	0.068	1.43E-01
T1a↔F3orb	7.4%	1.20E-01	0.065	1.46E-01
F3tv↔T3p	8.4%	1.22E-01	0.074	1.47E-01
AG↔PrecR	7.7%	1.26E-01	0.068	1.51E-01
PT↔F3td	7.1%	1.35E-01	0.066	1.62E-01
PT↔PrF3op	8.2%	1.37E-01	0.076	1.63E-01
PrF3op↔T2p	7.4%	1.37E-01	0.069	1.63E-01
Pole↔F3opdR	7.5%	1.36E-01	0.070	1.63E-01
Fusa↔T1R	7.0%	1.50E-01	0.069	1.77E-01
F3orb↔RolS	7.4%	1.55E-01	0.074	1.82E-01
STSp↔F3opdR	6.2%	1.57E-01	0.062	1.83E-01
F3opd↔T1a.HeschIR	5.8%	1.76E-01	0.064	2.06E-01
F3td↔RolS	5.7%	1.81E-01	0.064	2.11E-01
PT↔T3p	4.7%	2.29E-01	0.065	2.65E-01
Prec↔F3tv	5.0%	2.30E-01	0.068	2.66E-01
PT↔F3opdR	4.6%	2.36E-01	0.065	2.71E-01
RolS↔F3opdR	4.9%	2.42E-01	0.071	2.77E-01
F3td↔T1R	4.5%	2.49E-01	0.067	2.84E-01
Prec↔PrecR	4.6%	2.52E-01	0.069	2.87E-01
F3orb↔T2pR	4.4%	2.62E-01	0.070	2.96E-01
Fusa↔F2antR	3.9%	2.73E-01	0.065	3.08E-01
Prec↔F3opdR	3.6%	2.95E-01	0.068	3.30E-01
F3opd↔Fusa	3.7%	2.95E-01	0.069	3.30E-01
T2p↔F3opdR	3.3%	2.96E-01	0.063	3.30E-01
PT↔T2ml	3.5%	3.01E-01	0.068	3.34E-01
T1a↔PrecR	3.2%	3.09E-01	0.065	3.42E-01
T2p↔T2pR	3.0%	3.26E-01	0.068	3.60E-01
T3p↔T1a.HeschIR	2.8%	3.30E-01	0.065	3.63E-01
F3opdR↔T1a.HeschIR	2.9%	3.35E-01	0.068	3.67E-01
F3opd↔F3opdR	2.9%	3.38E-01	0.071	3.68E-01
PT↔F3orb	2.1%	3.79E-01	0.069	4.12E-01
PT↔F3opd	1.9%	3.86E-01	0.066	4.18E-01
PT↔SMG	1.7%	3.89E-01	0.059	4.20E-01
T1R↔PrecR	1.6%	4.03E-01	0.064	4.33E-01
F2p↔F3opdR	1.0%	4.34E-01	0.063	4.65E-01
PrF3op↔PrecR	1.0%	4.37E-01	0.064	4.66E-01
F2antR↔PrecR	0.0%	5.00E-01	0.000	5.00E-01
Fusa↔PrecR	0.0%	5.00E-01	0.000	5.00E-01

Endophenotype	heritability	P value	SE	FDR P value
F2p↔PrecR	0.0%	5.00E-01	0.000	5.00E-01
PrF3op↔T1a.HeschIR	0.0%	5.00E-01	0.000	5.00E-01
PT↔F2p	0.0%	5.00E-01	0.000	5.00E-01
Pole↔RolS	0.0%	5.00E-01	0.000	5.00E-01
F3td↔T1a.HeschIR	0.0%	5.00E-01	0.000	5.00E-01
F3opdR↔PrecR	0.0%	5.00E-01	0.000	5.00E-01
Pole↔PrecR	0.0%	5.00E-01	0.000	5.00E-01
F3opd↔PrecR	0.0%	5.00E-01	0.000	5.00E-01
SMG↔PrecR	0.0%	5.00E-01	0.000	5.00E-01
F3td↔F3opdR	0.0%	5.00E-01	0.000	5.00E-01
T2p↔PrecR	0.0%	5.00E-01	0.000	5.00E-01
T2pR↔PrecR	0.0%	5.00E-01	0.000	5.00E-01
F2p↔T2p	0.0%	5.00E-01	0.000	5.00E-01
T3p↔PrecR	0.0%	5.00E-01	0.000	5.00E-01
PT↔T2pR	0.0%	5.00E-01	0.000	5.00E-01
PT↔PrecR	0.0%	5.00E-01	0.000	5.00E-01
F2p↔T3p	0.0%	5.00E-01	0.000	5.00E-01

B.2 SNP-based heritability of neuroimaging endophenotypes: a general population study

Table B.2 – Heritabilities of brain functional connectivities extracted using OcSS ROIs. Bold phenotypes represents FCs with significant heritability after false discovery rate correction at $p < 0.05$.

Endophenotype	heritability	P value	SE	FDR P value
IFGtri↔IFGorb	12.0%	1.01E-12	0.018	1.52E-11
aSTS↔TPJ	11.4%	3.07E-11	0.018	2.30E-10
TP↔aSTS	11.0%	1.04E-10	0.018	5.20E-10
IFGorb↔pSTS	8.9%	1.65E-07	0.018	6.18E-07
TP↔TPJ	8.3%	1.02E-06	0.018	2.19E-06
IFGorb↔TPJ	8.3%	8.84E-07	0.018	2.19E-06
aSTS↔pSTS	8.3%	1.00E-06	0.018	2.19E-06
IFGorb↔TP	8.0%	2.41E-06	0.018	4.51E-06
pSTS↔TPJ	7.9%	2.83E-06	0.018	4.72E-06
IFGorb↔aSTS	7.2%	8.42E-06	0.017	1.26E-05
IFGtri↔aSTS	7.1%	1.69E-05	0.017	2.31E-05
IFGtri↔TP	6.3%	1.16E-04	0.017	1.45E-04
TP↔pSTS	5.3%	1.06E-03	0.018	1.22E-03
IFGtri↔pSTS	4.9%	2.05E-03	0.017	2.20E-03
IFGtri↔TPJ	2.0%	1.22E-01	0.017	1.22E-01

Table B.3 – **Heritabilities of brain functional connectivities extracted using SmPSS ROIs.** Bold phenotypes represents FCs with significant heritability after false discovery rate correction at $p < 0.05$.

Endophenotype	heritability	P value	SE	FDR P value
F3opd↔SMG	14.3%	3.89E-16	0.018	1.17E-13
AG↔F3tv	13.3%	9.99E-15	0.018	1.50E-12
T2ml↔STSp	12.4%	5.04E-13	0.018	5.04E-11
STSp↔T1R	12.3%	7.60E-13	0.018	5.70E-11
Pole↔STSp	12.2%	1.05E-12	0.018	6.27E-11
Pole↔T2ml	11.9%	2.14E-12	0.018	1.07E-10
PT↔T1	11.9%	4.90E-12	0.018	2.10E-10
F2p↔STSp	11.6%	1.32E-11	0.018	4.95E-10
T2ml↔SMG	11.2%	2.75E-11	0.018	8.26E-10
T2ml↔AG	11.5%	2.62E-11	0.018	8.26E-10
AG↔F2antR	11.3%	3.87E-11	0.018	1.06E-09
PrF3op↔SMG	11.4%	5.13E-11	0.018	1.28E-09
F3tv↔STSp	11.2%	9.27E-11	0.018	2.14E-09
RoIS↔PrecR	11.1%	2.27E-10	0.018	4.87E-09
PrF3op↔F3tv	10.7%	9.50E-10	0.018	1.90E-08
F3tv↔SMG	10.4%	1.93E-09	0.018	3.61E-08
F2p↔SMG	9.9%	2.84E-09	0.018	5.01E-08
STSp↔F3orb	10.1%	3.71E-09	0.018	6.18E-08
SMG↔F2antR	9.7%	1.32E-08	0.018	2.08E-07
T1↔F3tv	9.1%	1.40E-07	0.018	2.10E-06
Fusa↔AG	8.7%	1.79E-07	0.018	2.56E-06
F3opd↔F3td	8.9%	1.89E-07	0.018	2.58E-06
T2p↔T3p	8.4%	3.86E-07	0.018	5.04E-06
Prec↔F3opd	8.5%	4.82E-07	0.018	6.03E-06
T2ml↔F3tv	8.3%	5.64E-07	0.018	6.76E-06
SMG↔F3orb	8.4%	6.25E-07	0.018	7.21E-06
Prec↔SMG	8.4%	6.54E-07	0.018	7.27E-06
F3tv↔F3orb	8.3%	1.04E-06	0.018	1.11E-05
SMG↔F3td	8.2%	1.80E-06	0.018	1.80E-05
T2ml↔F3td	8.1%	1.79E-06	0.018	1.80E-05
AG↔T1a	7.8%	2.55E-06	0.018	2.47E-05
F3opd↔T2ml	7.8%	2.66E-06	0.018	2.50E-05
AG↔SMG	8.0%	2.88E-06	0.018	2.54E-05
F3tv↔F2antR	7.8%	2.80E-06	0.018	2.54E-05
Pole↔T1a	7.8%	3.00E-06	0.018	2.58E-05

Endophenotype	heritability	P value	SE	FDR P value
T1↔PrF3op	7.7%	4.74E-06	0.018	3.95E-05
AG↔STSp	7.7%	4.98E-06	0.018	4.04E-05
F3opd↔STSp	7.7%	5.28E-06	0.018	4.17E-05
PrF3op↔T3p	7.4%	6.48E-06	0.018	4.98E-05
PrF3op↔F2antR	7.5%	7.05E-06	0.018	5.28E-05
T2ml↔F3orb	7.5%	9.05E-06	0.018	6.62E-05
T1a↔STSp	7.3%	1.07E-05	0.018	7.62E-05
F3opdR↔F2antR	7.3%	1.13E-05	0.018	7.87E-05
Pole↔F3opd	7.3%	1.55E-05	0.018	1.06E-04
AG↔T3p	7.1%	1.86E-05	0.018	1.24E-04
STSp↔T2pR	7.0%	2.38E-05	0.018	1.55E-04
F3opd↔PrF3op	6.9%	3.25E-05	0.018	2.08E-04
T1↔T1a/HeschlR	6.9%	3.60E-05	0.018	2.25E-04
Prec↔AG	6.9%	4.14E-05	0.018	2.54E-04
F3tv↔T2pR	6.7%	4.25E-05	0.018	2.54E-04
T1a↔T1a/HeschlR	6.8%	4.32E-05	0.018	2.54E-04
PrF3op↔AG	6.7%	4.52E-05	0.018	2.60E-04
T1↔T1a	6.8%	5.09E-05	0.018	2.88E-04
Prec↔F3tv	6.7%	5.59E-05	0.018	3.10E-04
Pole↔AG	6.6%	5.71E-05	0.018	3.11E-04
F3opd↔AG	6.5%	6.10E-05	0.017	3.27E-04
F3tv↔F3opdR	6.6%	6.29E-05	0.018	3.31E-04
F3orb↔F2antR	6.6%	7.82E-05	0.018	3.85E-04
PT↔T1a/HeschlR	6.6%	7.74E-05	0.018	3.85E-04
SMG↔T3p	6.7%	7.78E-05	0.018	3.85E-04
T1a↔SMG	6.5%	7.59E-05	0.018	3.85E-04
Prec↔T1	6.5%	8.27E-05	0.018	4.00E-04
Prec↔T2p	6.5%	8.48E-05	0.018	4.04E-04
PrF3op↔STSp	6.5%	9.15E-05	0.018	4.29E-04
T2ml↔PrF3op	6.4%	1.32E-04	0.018	6.11E-04
Fusa↔T2p	6.3%	1.39E-04	0.018	6.30E-04
F3opd↔T3p	6.3%	1.74E-04	0.018	7.78E-04
T2ml↔T1R	6.1%	1.93E-04	0.018	8.53E-04
F2p↔F3tv	6.1%	2.04E-04	0.018	8.84E-04
T2ml↔F3opdR	6.2%	2.06E-04	0.018	8.84E-04
F3opd↔F2p	6.0%	2.13E-04	0.017	8.99E-04
T2ml↔T1a	6.2%	2.18E-04	0.018	9.07E-04
Pole↔T1R	6.2%	2.33E-04	0.018	9.59E-04
SMG↔T2pR	5.9%	2.42E-04	0.017	9.70E-04
AG↔F3orb	6.0%	2.43E-04	0.018	9.70E-04
AG↔T2p	6.0%	2.63E-04	0.018	1.04E-03
STSp↔F3opdR	5.5%	6.15E-04	0.017	2.40E-03
F2p↔T2ml	5.7%	6.92E-04	0.018	2.66E-03
T2p↔RoIS	5.6%	7.09E-04	0.018	2.69E-03
T3p↔F3orb	5.5%	7.21E-04	0.017	2.70E-03
T1a.HeschlR↔T1R	5.5%	8.11E-04	0.018	3.01E-03
F3tv↔RoIS	5.5%	1.01E-03	0.018	3.65E-03
F3orb↔T2pR	5.2%	1.01E-03	0.017	3.65E-03
T2ml↔T2pR	5.3%	1.08E-03	0.018	3.84E-03

Endophenotype	heritability	P value	SE	FDR P value
F3tv↔T1R	5.3%	1.14E-03	0.018	4.03E-03
F3td↔T3p	5.3%	1.18E-03	0.018	4.07E-03
T2p↔F3opdR	5.0%	1.18E-03	0.017	4.07E-03
F3opd↔F3orb	5.2%	1.34E-03	0.017	4.58E-03
F2p↔PrF3op	5.1%	1.46E-03	0.017	4.92E-03
F3tv↔T2p	5.2%	1.52E-03	0.018	5.02E-03
T1↔RolS	5.2%	1.51E-03	0.018	5.02E-03
Prec↔F3opdR	5.2%	1.59E-03	0.018	5.18E-03
T1↔AG	5.1%	1.62E-03	0.017	5.23E-03
STSp↔F3td	5.2%	1.72E-03	0.018	5.37E-03
F3tv↔T1a/HeschIR	5.1%	1.70E-03	0.018	5.37E-03
STSp↔F2antR	5.1%	1.71E-03	0.018	5.37E-03
F3tv↔T3p	5.0%	2.11E-03	0.018	6.53E-03
Fusa↔T3p	4.9%	2.37E-03	0.017	7.26E-03
Prec↔Fusa	5.0%	2.61E-03	0.018	7.90E-03
F3opdR↔T2pR	4.7%	2.96E-03	0.017	8.87E-03
F3tv↔F3td	4.7%	3.13E-03	0.018	9.22E-03
F3opdR↔T1R	4.7%	3.11E-03	0.017	9.22E-03
SMG↔F3opdR	4.6%	3.45E-03	0.017	1.01E-02
F3orb↔F3opdR	4.4%	3.88E-03	0.017	1.12E-02
Fusa↔STSp	4.6%	4.18E-03	0.018	1.18E-02
T1↔T1R	4.6%	4.15E-03	0.017	1.18E-02
F3td↔F2antR	4.4%	4.60E-03	0.017	1.29E-02
PrF3op↔T1R	4.6%	4.70E-03	0.018	1.30E-02
F2antR↔T2pR	4.5%	4.88E-03	0.018	1.34E-02
T1↔F3orb	4.5%	4.97E-03	0.017	1.36E-02
Prec↔RolS	4.5%	5.65E-03	0.018	1.53E-02
F2p↔AG	4.5%	5.93E-03	0.018	1.59E-02
T1a↔F3orb	4.4%	6.00E-03	0.018	1.59E-02
T1a↔T1R	4.3%	6.51E-03	0.017	1.71E-02
PT↔T1R	4.2%	6.84E-03	0.017	1.78E-02
T2pR↔PrecR	4.2%	7.83E-03	0.018	2.02E-02
Pole↔F2p	4.1%	8.14E-03	0.017	2.09E-02
F3opd↔F2antR	4.1%	8.40E-03	0.017	2.13E-02
Prec↔F3td	4.2%	9.12E-03	0.018	2.30E-02
SMG↔T2p	4.0%	9.78E-03	0.017	2.44E-02
PT↔AG	4.1%	9.95E-03	0.018	2.47E-02
PrF3op↔F3opdR	4.0%	1.04E-02	0.018	2.55E-02
PrF3op↔F3td	3.9%	1.15E-02	0.017	2.77E-02
Fusa↔T1a	4.0%	1.14E-02	0.018	2.77E-02
T2pR↔T1R	3.8%	1.35E-02	0.017	3.24E-02
T2p↔F2antR	3.9%	1.39E-02	0.018	3.32E-02
T2ml↔T2p	3.6%	1.41E-02	0.017	3.34E-02
T2ml↔T1	3.8%	1.45E-02	0.017	3.39E-02
F2antR↔T1a/HeschIR	3.9%	1.50E-02	0.018	3.48E-02
F3opd↔T1R	3.8%	1.60E-02	0.018	3.69E-02
F3td↔T1R	3.7%	1.70E-02	0.017	3.88E-02
F3opd↔T1	3.7%	1.74E-02	0.018	3.92E-02

Endophenotype	heritability	P value	SE	FDR P value
F3opd↔T1a	3.6%	1.73E-02	0.017	3.92E-02
F2p↔T2p	3.6%	1.79E-02	0.017	4.01E-02
Prec↔F2antR	3.6%	1.85E-02	0.017	4.11E-02
Fusa↔SMG	3.6%	1.93E-02	0.017	4.24E-02
AG↔PrecR	3.7%	1.94E-02	0.018	4.24E-02
SMG↔T1R	3.5%	1.95E-02	0.017	4.24E-02
F2antR↔T1R	3.6%	2.02E-02	0.018	4.36E-02
F2p↔T2pR	3.5%	2.17E-02	0.018	4.65E-02
T1↔SMG	3.5%	2.30E-02	0.017	4.89E-02
F2p↔T1	3.5%	2.33E-02	0.018	4.93E-02
STSp↔SMG	3.4%	2.50E-02	0.018	5.25E-02
F2p↔T1a	3.4%	2.54E-02	0.017	5.29E-02
F3tv↔T1a	3.3%	2.57E-02	0.017	5.32E-02
Prec↔T3p	3.4%	2.63E-02	0.018	5.40E-02
T2ml↔F2antR	3.3%	2.81E-02	0.017	5.73E-02
Pole↔F3orb	3.2%	2.90E-02	0.017	5.89E-02
F3opd↔T2pR	3.2%	3.22E-02	0.017	6.49E-02
Prec↔F3orb	3.0%	3.68E-02	0.017	7.35E-02
PrF3op↔T2pR	3.1%	3.70E-02	0.017	7.35E-02
RolS↔F3opdR	3.1%	3.77E-02	0.017	7.44E-02
PT↔RolS	3.1%	3.82E-02	0.017	7.47E-02
AG↔T1R	3.1%	3.86E-02	0.017	7.47E-02
T1a↔RolS	3.1%	3.85E-02	0.018	7.47E-02
PrF3op↔T2p	3.0%	4.08E-02	0.017	7.85E-02
Pole↔PT	2.9%	4.16E-02	0.017	7.95E-02
PT↔F2antR	2.9%	4.30E-02	0.017	8.17E-02
T1↔PrecR	3.0%	4.34E-02	0.017	8.18E-02
Pole↔PrF3op	2.9%	4.44E-02	0.017	8.32E-02
RolS↔F2antR	3.0%	4.48E-02	0.017	8.36E-02
T1↔T2p	2.9%	4.76E-02	0.017	8.82E-02
Pole↔T1	2.9%	4.83E-02	0.017	8.90E-02
PrF3op↔PrecR	2.8%	5.01E-02	0.017	9.17E-02
F3opd↔RolS	2.8%	5.17E-02	0.017	9.39E-02
Pole↔F3tv	2.8%	5.21E-02	0.017	9.42E-02
T1↔T2pR	2.8%	5.50E-02	0.017	9.88E-02
Pole↔T2p	2.7%	5.84E-02	0.017	1.04E-01
F2p↔F2antR	2.7%	5.99E-02	0.017	1.06E-01
T1↔F2antR	2.6%	6.18E-02	0.017	1.09E-01
PT↔T2ml	2.7%	6.21E-02	0.018	1.09E-01
F3tv↔PrecR	2.7%	6.39E-02	0.017	1.10E-01
T2ml↔T1a.HeschlR	2.6%	6.36E-02	0.017	1.10E-01
AG↔F3td	2.6%	6.35E-02	0.017	1.10E-01
T2p↔T1R	2.6%	6.86E-02	0.017	1.18E-01
PT↔T2pR	2.5%	7.04E-02	0.017	1.20E-01
F3opd↔F3tv	2.5%	7.24E-02	0.017	1.23E-01

Endophenotype	heritability	P value	SE	FDR P value
F2p↔T1R	2.5%	7.45E-02	0.017	1.26E-01
AG↔RolS	2.5%	7.69E-02	0.017	1.29E-01
STSp↔T3p	2.5%	7.91E-02	0.018	1.31E-01
T2ml↔T3p	2.5%	7.91E-02	0.017	1.31E-01
F2p↔F3opdR	2.4%	7.97E-02	0.017	1.31E-01
T1↔STSp	2.4%	8.26E-02	0.017	1.35E-01
T3p↔T2pR	2.4%	8.40E-02	0.017	1.37E-01
Fusa↔PrF3op	2.3%	8.76E-02	0.017	1.42E-01
STSp↔T2p	2.3%	8.88E-02	0.017	1.43E-01
T1↔Fusa	2.4%	8.93E-02	0.018	1.43E-01
T3p↔RolS	2.3%	9.24E-02	0.017	1.46E-01
AG↔T2pR	2.3%	9.26E-02	0.017	1.46E-01
Prec↔PrF3op	2.3%	9.17E-02	0.017	1.46E-01
Fusa↔T1a.HeschIR	2.3%	9.17E-02	0.018	1.46E-01
T3p↔F3opdR	2.2%	9.38E-02	0.017	1.47E-01
Prec↔STSp	2.2%	9.65E-02	0.017	1.50E-01
T2ml↔RolS	2.3%	9.88E-02	0.018	1.53E-01
T1a↔F2antR	2.2%	1.03E-01	0.017	1.59E-01
SMG↔PrecR	2.1%	1.04E-01	0.017	1.60E-01
Prec↔T1a	2.1%	1.05E-01	0.017	1.60E-01
PT↔STSp	2.2%	1.05E-01	0.018	1.60E-01
PT↔T2p	2.1%	1.06E-01	0.017	1.60E-01
T3p↔T1R	2.1%	1.07E-01	0.017	1.60E-01
T1↔F3opdR	2.1%	1.12E-01	0.017	1.67E-01
F3td↔T2pR	2.1%	1.14E-01	0.017	1.67E-01
PT↔T1a	2.1%	1.13E-01	0.018	1.67E-01
F3opdR↔PrecR	2.1%	1.14E-01	0.017	1.67E-01
T1a↔F3opdR	2.1%	1.14E-01	0.017	1.67E-01
Fusa↔F3tv	2.0%	1.19E-01	0.017	1.71E-01
F3orb↔T1R	2.1%	1.19E-01	0.018	1.71E-01
PT↔F3opd	2.0%	1.18E-01	0.017	1.71E-01
PrF3op↔F3orb	2.1%	1.18E-01	0.017	1.71E-01
F3orb↔RolS	2.0%	1.24E-01	0.017	1.78E-01
Fusa↔F2antR	2.0%	1.25E-01	0.017	1.78E-01
F3opd↔F3opdR	2.0%	1.27E-01	0.018	1.79E-01
F2p↔T3p	2.0%	1.27E-01	0.017	1.79E-01
Fusa↔T1R	2.0%	1.28E-01	0.017	1.79E-01
F3td↔F3orb	2.0%	1.29E-01	0.018	1.80E-01
RolS↔T1R	1.9%	1.30E-01	0.017	1.81E-01
F3td↔F3opdR	2.0%	1.31E-01	0.018	1.81E-01
F2p↔F3orb	1.9%	1.32E-01	0.017	1.82E-01
Fusa↔RolS	1.9%	1.33E-01	0.017	1.83E-01
F3orb↔T1a.HeschIR	1.9%	1.35E-01	0.017	1.84E-01
T1a↔T2p	1.9%	1.39E-01	0.017	1.88E-01
T1R↔PrecR	1.8%	1.42E-01	0.017	1.92E-01
Prec↔T1a.HeschIR	1.8%	1.50E-01	0.017	2.01E-01
Pole↔SMG	1.8%	1.53E-01	0.017	2.03E-01
SMG↔T1a.HeschIR	1.8%	1.53E-01	0.017	2.03E-01
Prec↔T2ml	1.8%	1.52E-01	0.017	2.03E-01

Endophenotype	heritability	P value	SE	FDR P value
Pole↔T1a.HeschIR	1.8%	1.55E-01	0.017	2.05E-01
RoIS↔T1a.HeschIR	1.8%	1.57E-01	0.017	2.07E-01
T2p↔F3orb	1.7%	1.68E-01	0.017	2.20E-01
T2p↔T2pR	1.6%	1.71E-01	0.017	2.22E-01
T3p↔F2antR	1.7%	1.72E-01	0.017	2.22E-01
F3opd↔T2p	1.7%	1.72E-01	0.017	2.22E-01
Pole↔RoIS	1.6%	1.77E-01	0.017	2.28E-01
F2antR↔PrecR	1.5%	1.99E-01	0.017	2.55E-01
STSp↔T1a.HeschIR	1.3%	2.35E-01	0.017	2.99E-01
STSp↔RoIS	1.2%	2.37E-01	0.017	3.02E-01
Prec↔PrecR	1.3%	2.38E-01	0.017	3.02E-01
Fusa↔F3orb	1.2%	2.41E-01	0.017	3.04E-01
Pole↔T2pR	1.1%	2.52E-01	0.017	3.17E-01
T3p↔PrecR	1.0%	2.76E-01	0.017	3.45E-01
Fusa↔F3td	1.0%	2.82E-01	0.017	3.51E-01
T2pR↔T1a.HeschIR	1.0%	2.86E-01	0.017	3.55E-01
T1a↔T3p	1.0%	2.89E-01	0.017	3.56E-01
Pole↔PrecR	1.0%	2.89E-01	0.017	3.56E-01
Pole↔F3td	0.9%	3.04E-01	0.017	3.72E-01
T2p↔F3td	0.9%	3.06E-01	0.017	3.73E-01
Fusa↔F3opdR	0.9%	3.08E-01	0.017	3.74E-01
Pole↔Fusa	0.8%	3.29E-01	0.017	3.98E-01
T1↔T3p	0.7%	3.34E-01	0.017	4.02E-01
Fusa↔PrecR	0.7%	3.45E-01	0.017	4.15E-01
Pole↔F2antR	0.6%	3.60E-01	0.017	4.30E-01
Prec↔Pole	0.6%	3.64E-01	0.017	4.31E-01
PrF3op↔T1a	0.6%	3.62E-01	0.017	4.31E-01
PT↔SMG	0.6%	3.70E-01	0.017	4.37E-01
T2p↔T1a.HeschIR	0.5%	3.91E-01	0.017	4.60E-01
PT↔F3tv	0.3%	4.26E-01	0.017	4.99E-01
Pole↔F3opdR	0.0%	5.00E-01	0.017	5.00E-01
STSp↔PrecR	0.0%	5.00E-01	0.017	5.00E-01
T1a↔PrecR	0.0%	5.00E-01	0.017	5.00E-01
F2p↔T1a.HeschIR	0.0%	5.00E-01	0.017	5.00E-01
Fusa↔T2pR	0.2%	4.46E-01	0.017	5.00E-01
F3td↔PrecR	0.0%	5.00E-01	0.017	5.00E-01
T2ml↔Fusa	0.0%	5.00E-01	0.017	5.00E-01
PT↔F2p	0.0%	5.00E-01	0.017	5.00E-01
PT↔F3opdR	0.0%	5.00E-01	0.017	5.00E-01
PT↔F3orb	0.0%	5.00E-01	0.017	5.00E-01
F3td↔RoIS	0.0%	5.00E-01	0.017	5.00E-01
Prec↔T2pR	0.2%	4.50E-01	0.017	5.00E-01
PrF3op↔RoIS	0.0%	5.00E-01	0.017	5.00E-01
F3orb↔PrecR	0.0%	5.00E-01	0.017	5.00E-01
AG↔F3opdR	0.0%	5.00E-01	0.017	5.00E-01
T1a↔T2pR	0.2%	4.65E-01	0.017	5.00E-01
F2p↔PrecR	0.0%	5.00E-01	0.017	5.00E-01
PrF3op↔T1a.HeschIR	0.0%	5.00E-01	0.017	5.00E-01

Endophenotype	heritability	P value	SE	FDR P value
F2p↔RoIS	0.0%	5.00E-01	0.017	5.00E-01
F2p↔Fusa	0.0%	5.00E-01	0.017	5.00E-01
F3opdR↔T1a.HeschIR	0.3%	4.35E-01	0.017	5.00E-01
T1a↔F3td	0.0%	5.00E-01	0.017	5.00E-01
Prec↔T1R	0.0%	4.99E-01	0.017	5.00E-01
T2p↔PrecR	0.0%	5.00E-01	0.017	5.00E-01
T1a.HeschIR↔PrecR	0.0%	5.00E-01	0.017	5.00E-01
T2ml↔PrecR	0.0%	5.00E-01	0.017	5.00E-01
F3opd↔T1a.HeschIR	0.0%	5.00E-01	0.017	5.00E-01
PT↔Fusa	0.0%	5.00E-01	0.017	5.00E-01
SMG↔RoIS	0.0%	5.00E-01	0.017	5.00E-01
Pole↔T3p	0.0%	5.00E-01	0.017	5.00E-01
AG↔T1a.HeschIR	0.0%	4.99E-01	0.017	5.00E-01
F3opd↔PrecR	0.0%	5.00E-01	0.017	5.00E-01
PT↔T3p	0.0%	5.00E-01	0.017	5.00E-01
PT↔F3td	0.0%	5.00E-01	0.018	5.00E-01
RoIS↔T2pR	0.0%	5.00E-01	0.017	5.00E-01
PT↔PrecR	0.0%	5.00E-01	0.017	5.00E-01
T1↔F3td	0.0%	5.00E-01	0.017	5.00E-01
F2p↔F3td	0.0%	5.00E-01	0.017	5.00E-01
Prec↔F2p	0.0%	5.00E-01	0.017	5.00E-01
T3p↔T1a.HeschIR	0.0%	5.00E-01	0.017	5.00E-01
F3opd↔Fusa	0.0%	5.00E-01	0.017	5.00E-01
PT↔PrF3op	0.0%	5.00E-01	0.017	5.00E-01
Prec↔PT	0.0%	5.00E-01	0.017	5.00E-01
F3td↔T1a.HeschIR	0.0%	5.00E-01	0.017	5.00E-01

Table B.4 – Heritabilities of brain anatomical connectivities. Bold phenotypes represents ACs with significant heritability after false discovery rate correction at $p < 0.05$.

Category	Endophenotype	$h^2(\%)$	P value	SE	FDR P value
ICVF	Corpus callosum	62.7%	1.98E-272	0.018	6.22E-270
ICVF	Arcuate Long Segment Left	60.5%	2.43E-251	0.018	3.82E-249
ICVF	Superior Longitudinal Fasciculus I Right	60.2%	8.83E-251	0.018	9.27E-249
ICVF	Superior Longitudinal Fasciculus I Left	60.0%	3.88E-250	0.018	3.06E-248
ICVF	Inferior Fronto Occipital fasciculus Left	59.9%	1.27E-247	0.018	8.00E-246
ICVF	Superior Longitudinal Fasciculus II Left	59.7%	3.42E-243	0.018	1.80E-241
ICVF	Superior Longitudinal Fasciculus III Left	58.5%	9.02E-243	0.018	4.06E-241
ICVF	Inferior Longitudinal Left	59.1%	4.71E-241	0.018	1.86E-239
ICVF	Superior Longitudinal Fasciculus III Right	58.2%	2.31E-240	0.018	8.08E-239
ICVF	Superior Longitudinal Fasciculus II Right	58.4%	3.33E-235	0.018	1.05E-233
ICVF	Arcuate Long Segment Right	58.1%	8.32E-233	0.018	2.38E-231
ICVF	Arcuate Anterior Segment Right	57.6%	2.22E-232	0.018	5.82E-231
ICVF	Arcuate Posterior Segment Left	57.5%	2.49E-232	0.018	6.03E-231
ICVF	Inferior Fronto Occipital fasciculus Right	58.0%	9.51E-231	0.018	2.14E-229
ICVF	Frontal Aslant Tract Right	57.2%	3.18E-230	0.018	6.69E-229
ICVF	Frontal Aslant Tract Left	57.2%	3.65E-230	0.018	7.19E-229
ICVF	Inferior Longitudinal Right	57.5%	2.17E-226	0.018	4.03E-225
ICVF	Arcuate Posterior Segment Right	56.9%	3.10E-224	0.018	5.43E-223
ICVF	Uncinate Left	54.1%	7.06E-208	0.018	1.17E-206
ICVF	Arcuate Anterior Segment Left	53.2%	1.08E-201	0.018	1.70E-200
ICVF	Uncinate Right	51.3%	2.79E-187	0.018	4.19E-186
FA	Corpus callosum	50.8%	2.19E-181	0.018	3.14E-180
L3	Inferior Fronto Occipital fasciculus Left	50.3%	1.17E-177	0.018	1.60E-176
FA	Arcuate Long Segment Left	49.9%	1.68E-174	0.018	2.20E-173
L3	Arcuate Long Segment Left	49.9%	7.05E-174	0.018	8.89E-173
MD	Arcuate Long Segment Left	49.7%	1.10E-172	0.018	1.33E-171
FA	Inferior Fronto Occipital fasciculus Left	49.3%	3.77E-172	0.018	4.40E-171
ICVF	Frontal Inferior longitudinal Right	49.0%	2.44E-171	0.018	2.75E-170
L3	Inferior Fronto Occipital fasciculus Right	48.5%	3.99E-166	0.018	4.33E-165
FA	Inferior Fronto Occipital fasciculus Right	48.4%	4.71E-166	0.018	4.95E-165
FA	Arcuate Anterior Segment Right	47.5%	2.21E-164	0.018	2.24E-163
FA	Superior Longitudinal Fasciculus III Right	47.7%	2.96E-164	0.018	2.91E-163
MD	Inferior Fronto Occipital fasciculus Left	48.4%	3.67E-164	0.018	3.51E-163

Category	Endophenotype	h2(%)	P value	SE	FDR P value
L3	Corpus callosum	48.0%	1.06E-163	0.018	9.83E-163
MD	Superior Longitudinal Fasciculus III Left	47.1%	4.72E-158	0.018	4.25E-157
L3	Superior Longitudinal Fasciculus II Left	47.7%	6.38E-157	0.018	5.58E-156
L2	Superior Longitudinal Fasciculus III Left	46.7%	7.14E-156	0.018	6.08E-155
FA	Superior Longitudinal Fasciculus II Left	47.8%	3.25E-155	0.018	2.69E-154
L2	Inferior Fronto Occipital fasciculus Left	46.7%	2.02E-154	0.018	1.63E-153
L2	Arcuate Long Segment Left	46.6%	2.47E-153	0.018	1.94E-152
L3	Superior Longitudinal Fasciculus III Right	45.9%	8.16E-153	0.018	6.27E-152
L3	Superior Longitudinal Fasciculus III Left	45.9%	1.30E-151	0.018	9.76E-151
L2	Superior Longitudinal Fasciculus III Right	45.7%	1.98E-151	0.018	1.45E-150
L3	Inferior Longitudinal Right	46.3%	4.84E-151	0.018	3.47E-150
FA	Superior Longitudinal Fasciculus III Left	45.5%	4.01E-150	0.018	2.80E-149
MD	Superior Longitudinal Fasciculus II Left	46.1%	1.44E-149	0.018	9.85E-149
MD	Superior Longitudinal Fasciculus III Right	45.4%	5.65E-149	0.018	3.79E-148
L3	Arcuate Anterior Segment Right	45.2%	1.35E-147	0.018	8.86E-147
L2	Corpus callosum	45.2%	1.69E-146	0.018	1.08E-145
FA	Superior Longitudinal Fasciculus II Right	45.1%	1.39E-142	0.018	8.74E-142
L2	Arcuate Anterior Segment Right	44.3%	1.72E-142	0.018	1.06E-141
MD	Arcuate Anterior Segment Right	44.6%	8.39E-142	0.018	5.08E-141
MD	Corpus callosum	44.5%	5.43E-141	0.018	3.23E-140
FA	Inferior Longitudinal Right	45.0%	7.60E-141	0.018	4.43E-140
L2	Superior Longitudinal Fasciculus II Left	44.9%	1.61E-140	0.018	9.22E-140
L3	Inferior Longitudinal Left	44.3%	2.83E-138	0.018	1.59E-137
OD	Fronto Insular tract3 Right	43.6%	1.19E-136	0.018	6.59E-136
L3	Arcuate Long Segment Right	43.8%	1.49E-136	0.018	8.10E-136
ICVF	Fronto Insular tract2 Right	43.2%	4.63E-136	0.018	2.47E-135
MD	Arcuate Long Segment Right	43.7%	1.73E-135	0.018	9.06E-135
ICVF	Frontal Superior Longitudinal Right	44.2%	8.28E-135	0.018	4.28E-134
FA	Superior Longitudinal Fasciculus I Left	43.7%	5.33E-134	0.018	2.71E-133
ICVF	Fronto Insular tract4 Right	43.1%	1.26E-132	0.018	6.30E-132
L3	Superior Longitudinal Fasciculus II Right	43.3%	2.01E-132	0.018	9.87E-132
L1	Inferior Fronto Occipital fasciculus Left	43.4%	3.46E-132	0.018	1.68E-131
L1	Superior Longitudinal Fasciculus III Left	43.5%	3.68E-132	0.018	1.76E-131
MD	Inferior Fronto Occipital fasciculus Right	43.1%	2.28E-131	0.018	1.07E-130
L2	Inferior Fronto Occipital fasciculus Right	43.0%	7.82E-131	0.018	3.62E-130
FA	Uncinate Left	43.4%	1.23E-130	0.018	5.61E-130
ICVF	Fronto Insular tract3 Right	41.9%	1.94E-128	0.018	8.74E-128
L3	Superior Longitudinal Fasciculus I Right	42.9%	3.34E-128	0.018	1.48E-127
L2	Superior Longitudinal Fasciculus II Right	42.5%	6.81E-128	0.018	2.98E-127
L1	Superior Longitudinal Fasciculus III Right	42.2%	1.50E-126	0.018	6.47E-126
L3	Superior Longitudinal Fasciculus I Left	42.1%	2.40E-126	0.018	1.02E-125
L3	Frontal Aslant Tract Left	41.7%	3.00E-125	0.018	1.26E-124
FA	Superior Longitudinal Fasciculus I Right	42.5%	1.11E-123	0.018	4.61E-123
MD	Superior Longitudinal Fasciculus II Right	41.3%	8.72E-122	0.018	3.57E-121
MD	Arcuate Posterior Segment Right	41.6%	9.44E-122	0.018	3.81E-121
MD	Frontal Aslant Tract Left	41.1%	1.15E-121	0.018	4.60E-121
FA	Frontal Aslant Tract Left	41.4%	1.67E-121	0.018	6.57E-121
FA	Arcuate Long Segment Right	41.3%	3.32E-121	0.018	1.29E-120
MD	Inferior Longitudinal Left	41.6%	1.24E-120	0.018	4.76E-120
L3	Uncinate Left	41.5%	2.92E-120	0.018	1.11E-119
MD	Arcuate Posterior Segment Left	41.2%	4.10E-120	0.018	1.54E-119

Category	Endophenotype	h2(%)	P value	SE	FDR P value
FA	Inferior Longitudinal Left	41.4%	1.20E-118	0.018	4.46E-118
ICVF	Frontal Inferior longitudinal Left	41.6%	5.12E-118	0.018	1.87E-117
L2	Arcuate Long Segment Right	40.6%	1.19E-117	0.018	4.32E-117
ICVF	Fronto Insular tract5 Right	40.6%	3.48E-117	0.018	1.24E-116
FA	Uncinate Right	40.5%	2.22E-116	0.018	7.86E-116
MD	Inferior Longitudinal Right	40.5%	3.73E-116	0.018	1.30E-115
L2	Uncinate Left	40.9%	2.62E-115	0.018	9.09E-115
L2	Frontal Aslant Tract Left	40.4%	5.63E-115	0.018	1.93E-114
L3	Uncinate Right	40.2%	7.38E-115	0.018	2.50E-114
L1	Arcuate Anterior Segment Right	40.6%	8.19E-114	0.019	2.74E-113
L2	Inferior Longitudinal Left	40.7%	1.22E-113	0.018	4.03E-113
OD	Arcuate Anterior Segment Right	40.6%	1.91E-113	0.019	6.28E-113
L2	Inferior Longitudinal Right	40.3%	2.63E-113	0.018	8.53E-113
ICVF	Frontal Superior Longitudinal Left	40.3%	3.28E-113	0.018	1.06E-112
L1	Corpus callosum	40.2%	8.26E-113	0.018	2.63E-112
L1	Arcuate Long Segment Left	40.1%	5.00E-112	0.019	1.57E-111
L1	Superior Longitudinal Fasciculus II Left	39.0%	8.17E-112	0.018	2.55E-111
L1	Frontal Aslant Tract Left	39.2%	1.17E-111	0.018	3.62E-111
L2	Superior Longitudinal Fasciculus I Left	39.2%	6.85E-111	0.018	2.10E-110
L2	Arcuate Posterior Segment Left	39.5%	3.09E-110	0.018	9.36E-110
L3	Arcuate Posterior Segment Right	39.5%	1.04E-109	0.018	3.11E-109
OD	Inferior Fronto Occipital fasciculus Left	39.7%	1.26E-109	0.018	3.75E-109
ICVF	Fronto Insular tract1 Right	38.9%	1.41E-109	0.018	4.15E-109
MD	Arcuate Anterior Segment Left	38.9%	4.95E-109	0.018	1.44E-108
MD	Frontal Aslant Tract Right	38.4%	4.90E-106	0.018	1.42E-105
MD	Superior Longitudinal Fasciculus I Left	38.4%	8.92E-106	0.018	2.55E-105
L2	Arcuate Anterior Segment Left	38.3%	8.60E-105	0.018	2.44E-104
MO	Frontal Aslant Tract Left	38.2%	1.31E-104	0.018	3.68E-104
L1	Frontal Aslant Tract Right	37.9%	8.94E-104	0.018	2.49E-103
MD	Superior Longitudinal Fasciculus I Right	38.2%	1.09E-103	0.018	3.02E-103
L3	Frontal Aslant Tract Right	37.5%	1.15E-101	0.018	3.14E-101
FA	Frontal Aslant Tract Right	37.7%	1.58E-101	0.018	4.30E-101
ISOVF	Superior Longitudinal Fasciculus III Right	37.4%	4.45E-101	0.018	1.20E-100
L2	Frontal Aslant Tract Right	37.5%	7.79E-100	0.018	2.08E-99
OD	Superior Longitudinal Fasciculus II Left	37.0%	3.27E-99	0.018	8.65E-99
ISOVF	Superior Longitudinal Fasciculus III Left	37.4%	4.02E-99	0.018	1.06E-98
ISOVF	Inferior Fronto Occipital fasciculus Left	37.2%	3.35E-98	0.018	8.73E-98
MD	Uncinate Left	37.4%	3.92E-98	0.018	1.01E-97
L2	Superior Longitudinal Fasciculus I Right	36.8%	1.46E-97	0.018	3.73E-97
MO	Frontal Aslant Tract Right	37.2%	6.74E-97	0.018	1.71E-96
L2	Uncinate Right	36.2%	8.87E-96	0.018	2.23E-95
MD	Frontal Inferior longitudinal Right	36.4%	1.53E-95	0.018	3.84E-95
MO	Arcuate Anterior Segment Right	36.8%	7.11E-95	0.018	1.76E-94
L1	Inferior Fronto Occipital fasciculus Right	36.1%	2.09E-94	0.018	5.14E-94
MO	Fronto Insular tract3 Right	35.7%	2.48E-92	0.018	6.05E-92

Category	Endophenotype	h2(%)	P value	SE	FDR P value
OD	Fronto Insular tract4 Right	33.9%	3.86E-91	0.018	9.36E-91
MO	Corpus callosum	36.3%	5.82E-91	0.019	1.40E-90
L3	Arcuate Posterior Segment Left	35.5%	7.21E-91	0.018	1.72E-90
L1	Superior Longitudinal Fasciculus II Right	35.5%	1.59E-90	0.018	3.77E-90
MO	Superior Longitudinal Fasciculus I Left	36.1%	5.44E-90	0.019	1.28E-89
OD	Inferior Fronto Occipital fasciculus Right	35.5%	7.26E-90	0.018	1.69E-89
MD	Uncinate Right	35.1%	4.39E-89	0.018	1.02E-88
L2	Arcuate Posterior Segment Right	35.3%	6.57E-89	0.018	1.51E-88
L3	Arcuate Anterior Segment Left	35.0%	6.94E-89	0.018	1.58E-88
FA	Arcuate Anterior Segment Left	34.9%	1.50E-88	0.018	3.39E-88
OD	Fronto Insular tract5 Right	34.9%	1.76E-88	0.018	3.97E-88
OD	Corpus callosum	35.7%	7.24E-88	0.019	1.62E-87
ISOVF	Superior Longitudinal Fasciculus II Left	34.8%	3.75E-87	0.018	8.32E-87
ICVF	Fronto Insular tract3 Left	34.6%	3.27E-86	0.018	7.20E-86
OD	Arcuate Long Segment Left	35.2%	8.03E-86	0.019	1.76E-85
OD	Superior Longitudinal Fasciculus II Right	34.5%	8.32E-86	0.018	1.81E-85
ISOVF	Inferior Fronto Occipital fasciculus Right	34.5%	5.23E-85	0.018	1.13E-84
MO	Inferior Fronto Occipital fasciculus Left	35.2%	1.90E-84	0.019	4.07E-84
ISOVF	Arcuate Long Segment Left	33.9%	1.45E-83	0.018	3.09E-83
OD	Superior Longitudinal Fasciculus III Right	34.4%	1.40E-82	0.019	2.96E-82
FA	Arcuate Posterior Segment Right	34.0%	1.24E-81	0.018	2.61E-81
MO	Inferior Fronto Occipital fasciculus Right	33.6%	2.08E-81	0.018	4.34E-81
L1	Arcuate Anterior Segment Left	33.6%	4.18E-81	0.018	8.67E-81
FA	Fronto Insular tract3 Right	33.8%	4.51E-81	0.018	9.28E-81
FA	Arcuate Posterior Segment Left	33.4%	3.82E-80	0.018	7.81E-80
L3	Frontal Inferior longitudinal Right	33.2%	4.60E-80	0.018	9.35E-80
OD	Superior Longitudinal Fasciculus III Left	33.5%	6.95E-80	0.018	1.40E-79
L1	Arcuate Long Segment Right	33.2%	4.86E-79	0.018	9.74E-79
MO	Superior Longitudinal Fasciculus I Right	33.7%	1.38E-78	0.019	2.75E-78
L1	Superior Longitudinal Fasciculus I Left	33.3%	1.55E-77	0.019	3.06E-77
MO	Fronto Insular tract4 Right	32.0%	2.35E-77	0.018	4.63E-77
MO	Arcuate Long Segment Left	33.5%	4.15E-77	0.019	8.11E-77
L1	Inferior Longitudinal Left	32.7%	2.66E-76	0.019	5.17E-76
L2	Frontal Inferior longitudinal Right	32.6%	3.57E-76	0.018	6.89E-76
FA	Fronto Insular tract1 Right	32.9%	6.36E-76	0.019	1.22E-75
L1	Superior Longitudinal Fasciculus I Right	32.8%	2.06E-75	0.019	3.93E-75
ISOVF	Inferior Longitudinal Right	32.6%	4.96E-75	0.018	9.41E-75
L1	Inferior Longitudinal Right	32.0%	6.02E-75	0.018	1.14E-74
L1	Arcuate Posterior Segment Right	32.1%	2.07E-74	0.018	3.89E-74
OD	Frontal Aslant Tract Left	31.6%	6.70E-74	0.018	1.25E-73
L1	Fronto Insular tract4 Right	30.9%	3.72E-73	0.018	6.90E-73
FA	Fronto Insular tract5 Right	32.0%	2.34E-72	0.019	4.31E-72
L2	Fronto Insular tract2 Right	31.3%	7.82E-71	0.018	1.43E-70
ISOVF	Corpus callosum	31.1%	1.04E-70	0.018	1.89E-70
L1	Fronto Insular tract3 Right	30.9%	1.12E-70	0.018	2.02E-70
ISOVF	Arcuate Anterior Segment Right	30.7%	7.18E-70	0.018	1.29E-69
MO	Superior Longitudinal Fasciculus III Left	31.0%	2.41E-69	0.018	4.32E-69
OD	Superior Longitudinal Fasciculus I Left	31.7%	1.60E-68	0.019	2.85E-68
FA	Fronto Insular tract2 Right	30.7%	2.05E-68	0.018	3.64E-68

Category	Endophenotype	h2(%)	P value	SE	FDR P value
L1	Arcuate Posterior Segment Left	30.8%	1.63E-67	0.019	2.86E-67
FA	Fronto Insular tract4 Right	30.3%	4.18E-67	0.018	7.31E-67
L1	Uncinate Left	30.1%	2.55E-65	0.018	4.43E-65
L1	Fronto Insular tract5 Right	29.7%	2.63E-65	0.018	4.55E-65
FA	Frontal Inferior longitudinal Right	29.9%	1.05E-64	0.018	1.80E-64
L3	Fronto Insular tract2 Right	29.8%	8.16E-64	0.018	1.40E-63
ISOVF	Superior Longitudinal Fasciculus II Right	29.4%	3.10E-63	0.018	5.27E-63
MD	Fronto Insular tract2 Right	29.7%	4.33E-63	0.018	7.34E-63
L1	Frontal Inferior longitudinal Right	29.1%	6.48E-61	0.018	1.09E-60
ISOVF	Inferior Longitudinal Left	29.6%	9.60E-61	0.019	1.61E-60
ISOVF	Frontal Aslant Tract Right	28.7%	8.42E-60	0.018	1.40E-59
MO	Superior Longitudinal Fasciculus III Right	28.6%	5.78E-59	0.018	9.58E-59
ICVF	Fronto Insular tract2 Left	28.3%	1.18E-58	0.018	1.94E-58
OD	Frontal Aslant Tract Right	28.1%	1.78E-58	0.018	2.93E-58
ICVF	Fronto Insular tract5 Left	28.3%	1.97E-58	0.018	3.21E-58
L1	Fronto Insular tract2 Right	28.4%	1.72E-57	0.018	2.79E-57
L2	Fronto Insular tract1 Right	27.7%	1.11E-56	0.018	1.79E-56
OD	Superior Longitudinal Fasciculus I Right	28.8%	1.23E-56	0.019	1.97E-56
L3	Fronto Insular tract5 Right	28.1%	1.64E-56	0.018	2.62E-56
L1	Uncinate Right	27.9%	2.26E-56	0.018	3.59E-56
ISOVF	Frontal Aslant Tract Left	27.6%	1.20E-55	0.018	1.89E-55
ISOVF	Superior Longitudinal Fasciculus I Right	27.3%	6.06E-55	0.018	9.54E-55
L3	Fronto Insular tract1 Left	27.2%	8.22E-55	0.018	1.29E-54
OD	Uncinate Left	27.3%	3.24E-54	0.018	5.06E-54
ISOVF	Arcuate Anterior Segment Left	27.6%	4.03E-54	0.019	6.26E-54
MD	Fronto Insular tract5 Right	26.7%	1.07E-52	0.018	1.65E-52
ISOVF	Arcuate Long Segment Right	26.4%	2.03E-52	0.018	3.12E-52
ISOVF	Arcuate Posterior Segment Right	26.6%	4.34E-52	0.018	6.64E-52
MO	Arcuate Long Segment Right	26.6%	5.05E-52	0.018	7.68E-52
L2	Fronto Insular tract3 Right	26.5%	4.07E-51	0.018	6.16E-51
ISOVF	Frontal Inferior longitudinal Right	26.5%	8.12E-51	0.018	1.22E-50
MD	Frontal Inferior longitudinal Left	26.3%	2.22E-50	0.018	3.33E-50
L1	Fronto Insular tract1 Right	25.8%	7.51E-50	0.018	1.12E-49
OD	Arcuate Long Segment Right	26.1%	8.86E-50	0.018	1.32E-49
ISOVF	Arcuate Posterior Segment Left	25.8%	1.28E-49	0.018	1.90E-49
MO	Fronto Insular tract5 Right	26.0%	2.02E-49	0.018	2.98E-49
L3	Fronto Insular tract3 Right	26.1%	2.24E-49	0.018	3.29E-49
MD	Fronto Insular tract1 Left	25.9%	4.53E-49	0.018	6.61E-49
L2	Fronto Insular tract1 Left	25.8%	2.61E-47	0.019	3.79E-47
OD	Fronto Insular tract1 Right	25.6%	2.64E-47	0.018	3.82E-47
MD	Fronto Insular tract3 Right	25.2%	4.80E-47	0.018	6.91E-47
ISOVF	Uncinate Left	25.3%	1.29E-46	0.018	1.85E-46
FA	Fronto Insular tract2 Left	25.2%	2.10E-46	0.018	2.99E-46
L2	Fronto Insular tract5 Right	24.9%	6.80E-46	0.018	9.64E-46
ISOVF	Superior Longitudinal Fasciculus I Left	24.6%	2.37E-45	0.018	3.35E-45
MO	Inferior Longitudinal Left	24.6%	7.19E-45	0.018	1.01E-44
L3	Fronto Insular tract1 Right	24.6%	7.60E-45	0.018	1.06E-44
MO	Inferior Longitudinal Right	24.5%	1.27E-44	0.018	1.77E-44

Category	Endophenotype	h2(%)	P value	SE	FDR P value
OD	Fronto Insular tract2 Left	24.7%	1.01E-43	0.018	1.40E-43
MD	Fronto Insular tract1 Right	24.1%	1.29E-43	0.018	1.79E-43
ISOVF	Fronto Insular tract3 Right	24.4%	2.15E-43	0.018	2.96E-43
L3	Fronto Insular tract4 Right	24.2%	2.09E-42	0.018	2.86E-42
MD	Fronto Insular tract4 Right	23.8%	5.62E-42	0.018	7.66E-42
ICVF	Fronto Insular tract1 Left	23.4%	1.34E-41	0.018	1.83E-41
MO	Superior Longitudinal Fasciculus II Right	23.8%	1.46E-40	0.019	1.98E-40
OD	Uncinate Right	23.9%	2.10E-40	0.019	2.83E-40
L2	Fronto Insular tract4 Right	23.4%	7.97E-40	0.018	1.07E-39
ISOVF	Fronto Insular tract5 Right	22.6%	9.59E-39	0.018	1.28E-38
OD	Inferior Longitudinal Right	22.9%	1.08E-38	0.018	1.44E-38
MO	Uncinate Left	23.1%	6.46E-38	0.019	8.54E-38
L3	Frontal Inferior longitudinal Left	22.6%	7.22E-38	0.018	9.51E-38
MO	Fronto Insular tract1 Right	22.1%	4.62E-37	0.018	6.07E-37
MO	Arcuate Anterior Segment Left	22.5%	2.62E-36	0.018	3.43E-36
ISOVF	Fronto Insular tract2 Right	22.1%	8.41E-36	0.018	1.10E-35
L1	Fronto Insular tract2 Left	22.0%	3.28E-35	0.018	4.26E-35
MD	Fronto Insular tract3 Left	21.7%	3.94E-35	0.018	5.08E-35
MO	Fronto Insular tract2 Right	21.6%	3.66E-34	0.018	4.70E-34
L2	Fronto Insular tract2 Left	21.0%	1.03E-33	0.018	1.31E-33
OD	Fronto Insular tract2 Right	21.3%	2.35E-33	0.018	2.99E-33
OD	Arcuate Anterior Segment Left	20.7%	8.73E-32	0.018	1.11E-31
MO	Fronto Insular tract2 Left	20.5%	9.99E-32	0.018	1.26E-31
L1	Fronto Insular tract1 Left	20.4%	1.68E-31	0.018	2.12E-31
L2	Fronto Insular tract3 Left	20.2%	1.83E-31	0.018	2.30E-31
OD	Inferior Longitudinal Left	20.5%	6.13E-31	0.018	7.67E-31
L1	Frontal Superior Longitudinal Right	20.3%	1.60E-30	0.018	1.99E-30
MO	Uncinate Right	20.4%	2.28E-30	0.018	2.82E-30
ICVF	Fronto Insular tract4 Left	19.3%	1.52E-28	0.018	1.88E-28
L2	Frontal Inferior longitudinal Left	19.1%	8.21E-28	0.018	1.01E-27
ISOVF	Fronto Insular tract1 Right	19.0%	2.23E-26	0.018	2.74E-26
MO	Superior Longitudinal Fasciculus II Left	18.8%	3.85E-26	0.018	4.71E-26
L1	Frontal Inferior longitudinal Left	18.8%	4.64E-26	0.018	5.64E-26
FA	Fronto Insular tract1 Left	18.3%	1.26E-25	0.018	1.52E-25
OD	Arcuate Posterior Segment Right	17.7%	3.83E-24	0.018	4.63E-24
ISOVF	Uncinate Right	17.4%	1.69E-23	0.018	2.03E-23
L1	Fronto Insular tract5 Left	17.2%	4.10E-23	0.018	4.91E-23
MD	Fronto Insular tract5 Left	16.8%	2.27E-22	0.018	2.71E-22
L1	Frontal Superior Longitudinal Left	16.8%	4.69E-22	0.018	5.58E-22
FA	Frontal Superior Longitudinal Left	16.6%	5.85E-21	0.018	6.92E-21
MD	Frontal Superior Longitudinal Left	16.2%	1.14E-20	0.018	1.34E-20
L2	Frontal Superior Longitudinal Left	16.3%	1.15E-20	0.018	1.35E-20
MO	Frontal Superior Longitudinal Right	16.5%	1.94E-20	0.018	2.27E-20
L3	Fronto Insular tract3 Left	16.0%	3.22E-20	0.018	3.76E-20
ISOVF	Fronto Insular tract4 Right	15.9%	3.61E-20	0.018	4.19E-20
OD	Arcuate Posterior Segment Left	16.0%	7.37E-20	0.018	8.54E-20

Category	Endophenotype	h2(%)	P value	SE	FDR P value
FA	Frontal Superior Longitudinal Right	16.1%	9.29E-20	0.018	1.07E-19
MD	Frontal Superior Longitudinal Right	16.0%	1.29E-19	0.018	1.48E-19
FA	Frontal Inferior longitudinal Left	15.9%	1.48E-19	0.018	1.70E-19
MO	Frontal Inferior longitudinal Right	16.2%	2.13E-19	0.018	2.43E-19
L3	Fronto Insular tract5 Left	15.5%	3.38E-19	0.018	3.84E-19
OD	Fronto Insular tract5 Left	15.4%	4.01E-19	0.018	4.55E-19
L3	Fronto Insular tract2 Left	15.6%	8.86E-19	0.018	1.00E-18
ISOVF	Fronto Insular tract1 Left	16.6%	2.63E-18	0.020	2.96E-18
MO	Fronto Insular tract5 Left	15.2%	3.72E-18	0.018	4.17E-18
ISOVF	Frontal Inferior longitudinal Left	14.8%	1.05E-17	0.018	1.18E-17
L2	Fronto Insular tract5 Left	14.5%	2.77E-17	0.018	3.08E-17
L3	Frontal Superior Longitudinal Left	14.4%	1.91E-16	0.018	2.12E-16
OD	Frontal Inferior longitudinal Right	14.6%	2.52E-16	0.018	2.79E-16
L3	Frontal Superior Longitudinal Right	14.3%	4.46E-16	0.018	4.91E-16
FA	Fronto Insular tract5 Left	14.0%	9.77E-16	0.018	1.07E-15
L2	Frontal Superior Longitudinal Right	13.9%	3.32E-15	0.018	3.63E-15
MD	Fronto Insular tract2 Left	13.7%	4.41E-15	0.018	4.81E-15
ISOVF	Fronto Insular tract3 Left	12.7%	3.71E-13	0.018	4.03E-13
ISOVF	Frontal Superior Longitudinal Left	12.5%	4.07E-13	0.018	4.40E-13
L1	Fronto Insular tract3 Left	12.2%	2.93E-12	0.018	3.16E-12
MO	Frontal Superior Longitudinal Left	11.9%	1.06E-11	0.018	1.14E-11
MO	Arcuate Posterior Segment Right	11.5%	2.07E-11	0.018	2.21E-11
FA	Fronto Insular tract3 Left	11.3%	6.24E-11	0.018	6.67E-11
ISOVF	Frontal Superior Longitudinal Right	11.2%	1.26E-10	0.018	1.35E-10
MD	Fronto Insular tract4 Left	11.0%	1.58E-10	0.018	1.68E-10
MO	Frontal Inferior longitudinal Left	10.7%	7.56E-10	0.018	7.99E-10
OD	Fronto Insular tract1 Left	10.0%	3.73E-09	0.018	3.93E-09
L3	Fronto Insular tract4 Left	10.0%	4.41E-09	0.018	4.63E-09
OD	Frontal Superior Longitudinal Left	10.0%	4.77E-09	0.018	4.99E-09
OD	Frontal Superior Longitudinal Right	9.9%	5.58E-09	0.018	5.82E-09
L2	Fronto Insular tract4 Left	9.7%	1.30E-08	0.018	1.35E-08
ISOVF	Fronto Insular tract2 Left	11.2%	2.77E-08	0.022	2.87E-08
MO	Arcuate Posterior Segment Left	9.3%	7.79E-08	0.018	8.05E-08
ISOVF	Fronto Insular tract5 Left	11.1%	8.78E-08	0.022	9.04E-08
MO	Fronto Insular tract3 Left	8.0%	2.24E-06	0.018	2.30E-06
FA	Fronto Insular tract4 Left	6.6%	7.15E-05	0.018	7.31E-05
OD	Frontal Inferior longitudinal Left	6.7%	7.78E-05	0.018	7.93E-05
L1	Fronto Insular tract4 Left	6.6%	8.26E-05	0.018	8.39E-05
OD	Fronto Insular tract3 Left	6.6%	8.39E-05	0.018	8.50E-05
MO	Fronto Insular tract1 Left	6.4%	1.47E-04	0.018	1.49E-04
OD	Fronto Insular tract4 Left	3.8%	1.32E-02	0.018	1.33E-02
MO	Fronto Insular tract4 Left	2.9%	5.05E-02	0.018	5.07E-02
ISOVF	Fronto Insular tract4 Left	2.4%	8.67E-02	0.018	8.67E-02

C.1 (Endo)phenotypic correlation between both resting-state functional connectivity and neural activations

Table C.1 – Pearson Correlation between both resting-state functional connectivity and neural activations. Bold values represents pairs of endophenotypes with significant correlation after false discovery rate correction at $p < 0.05$.

Endophenotype	PrF3op	Tzml	F3td	TiR	F3opd	F2p	T1a	FzantR	T1a.HeschlR	AG	STSp	T1	PT	F3orb	Pole	Prec	F3opdR	F3tv	T2pR	SMG	T2p	RoIS	Fusa	PreCR	T3p
PrF3op+SMG	-0.211*	0.047	0.055	0.046	0.099	-0.034	0.028	0.042	-0.026	-0.046	0.014	0.024	-0.001	0.0	-0.008	-0.005	0.019	0.009	-0.013	0.004	0.003	-0.031	-0.013	0.003	-0.004
Pole+PrF3op	0.313*	-0.031	-0.027	-0.024	-0.096	0.008	-0.032	-0.018	0.002	0.034	-0.019	-0.041	-0.011	0.01	-0.017	-0.005	-0.064	0.025	0.002	0.015	-0.006	0.044	0.037	-0.018	-0.032
Pole+Tzml	-0.012	0.303*	0.013	0.116*	-0.007	0.048	0.066	-0.018	0.079	0.026	-0.024	0.114*	0.1	0.009	0.16*	0.026	0.093	0.009	0.085	-0.021	0.0	0.012	0.049	-0.013	0.004
Tzml+F3td	-0.024	-0.028	0.28*	0.013	0.143*	-0.01	-0.001	0.101	-0.003	0.032	0.058	0.005	-0.018	0.058	0.024	0.042	0.082	-0.01	-0.007	0.056	0.042	0.031	-0.001	0.019	0.043
PrF3op+STSp	0.261*	-0.012	-0.008	0.007	-0.128*	-0.038	-0.012	-0.021	0.014	0.017	-0.039	-0.017	0.023	0.014	-0.025	0.024	-0.066	0.054	0.021	0.027	-0.025	0.074	0.019	0.064	-0.019
T1a+TiR	0.063	0.058	0.075	0.256*	0.07	0.01	0.165*	0.064	0.115*	0.001	-0.001	-0.034	-0.005	-0.015	-0.023	-0.013	-0.111*	0.101	0.037	0.028	0.035	0.065	0.009	0.025	0.01
Pole+F3op	-0.041	0.032	0.153*	-0.005	0.254*	-0.11	-0.042	0.077	-0.055	-0.024	-0.032	0.015	-0.061	0.005	-0.057	0.057	0.087	-0.102	-0.025	-0.002	0.009	0.038	-0.019	0.04	0.048
Pole+F2p	-0.005	-0.026	-0.001	-0.005	-0.026	0.251*	0.017	0.058	0.087	-0.009	-0.022	-0.011	-0.01	0.031	-0.027	-0.038	-0.011	0.022	-0.011	-0.003	0.053	-0.021	0.07	0.013	0.03
SMG+F3td	0.064	-0.014	-0.249*	-0.039	-0.111	-0.037	0.031	-0.126*	-0.012	0.021	-0.029	-0.042	0.017	0.012	0.023	-0.068	-0.037	0.033	-0.012	-0.008	-0.048	-0.009	-0.041	0.011	-0.047
F3opd+SMG	0.062	-0.011	-0.187	-0.093	-0.218*	0.121*	0.021	-0.056	0.031	0.001	-0.001	-0.034	-0.005	-0.015	-0.023	-0.013	-0.111*	0.101	0.037	0.028	0.035	0.065	0.009	0.025	0.01
PrF3op+F3tv	0.181*	0.105	0.029	0.02	0.018	0.234*	0.033	0.034	0.035	0.088	0.105	0.038	0.042	0.086	0.097	0.001	0.052	0.051	0.071	0.051	0.092	0.04	0.071	0.088	0.004
F2p+SMG	0.038	0.006	0.128*	0.095	0.173*	-0.229*	-0.06	0.087	-0.005	-0.042	0.043	0.014	-0.02	0.004	-0.012	0.092	0.087	-0.014	0.033	-0.062	-0.052	-0.026	-0.042	-0.056	-0.001
T1a+T1a.HeschlR	0.015	0.036	0.001	0.061	0.049	0.03	0.228*	-0.05	0.219*	0.062	0.038	0.025	-0.032	0.014	0.144*	0.06	0.028	0.022	0.035	0.002	0.016	0.013	0.045	0.015	0.042
PrF3op+T1a	0.228*	-0.045	-0.04	-0.008	-0.055	-0.069	0.047	-0.026	0.04	0.014	-0.048	0.008	0.014	0.032	-0.034	0.058	-0.036	0.035	0.027	-0.017	-0.041	0.006	0.021	-0.005	-0.016
SMG+FzantR	0.042	0.064	-0.077	-0.058	-0.019	0.056	0.041	-0.221*	0.04	0.027	0.025	-0.022	0.023	0.015	0.088	-0.059	-0.02	0.028	-0.028	-0.089	-0.0	0.026	-0.008	-0.003	-0.004
PrF3op+T2pR	-0.226*	-0.039	0.044	0.01	0.053	-0.095	-0.045	-0.008	-0.02	-0.045	-0.022	0.004	0.025	-0.04	-0.014	0.001	0.025	-0.035	-0.046	-0.016	0.019	-0.011	0.003	0.042	0.025
F3opd+Tzml	0.077	0.012	0.132*	0.046	0.22*	-0.05	-0.009	0.035	-0.007	-0.058	0.009	0.034	0.009	0.016	-0.001	-0.002	0.085	-0.087	-0.049	-0.02	-0.007	-0.017	-0.007	0.035	-0.002
PT+T1a	-0.082	0.049	0.034	0.03	0.095	0.026	0.21*	0.008	0.075	0.04	0.041	-0.017	0.143*	0.011	0.138*	0.086	0.057	-0.033	0.056	-0.021	0.033	0.011	0.034	0.025	0.02
Tzml+RoIS	0.054	0.21*	0.01	0.033	0.044	0.04	-0.005	0.031	0.016	0.082	0.018	-0.066	-0.035	0.069	0.021	0.05	0.029	0.042	0.054	0.078	0.115*	0.035	0.068	0.047	0.001
AG+F3tv	-0.029	-0.103	0.047	-0.006	0.087	0.026	-0.069	0.063	-0.063	-0.209*	0.047	0.004	-0.015	-0.058	-0.007	-0.009	0.048	-0.044	-0.007	-0.052	-0.02	-0.056	-0.023	-0.016	0.009
F3opd+T1a	0.008	-0.004	0.116*	0.109	0.209*	-0.052	-0.001	0.093	0.068	0.049	0.044	0.096	0.026	0.012	-0.001	0.041	0.113	-0.028	0.009	-0.037	0.116*	0.077	0.071	0.088	0.008
Tzml+PrF3op	0.207*	-0.029	-0.104	-0.118*	-0.158*	-0.036	-0.041	-0.031	-0.042	-0.01	-0.059	-0.076	-0.054	-0.011	-0.038	-0.027	-0.073	0.018	-0.036	-0.025	-0.004	0.028	-0.019	-0.035	-0.012
Prec+F3tv	-0.051	0.065	0.153*	0.03	0.207*	-0.017	-0.012	0.041	0.0	-0.015	0.047	0.051	0.033	0.042	0.032	0.052	0.048	-0.025	-0.034	0.018	-0.007	0.002	0.009	-0.013	0.002
Pole+FzantR	-0.018	0.056	0.112	0.059	0.086	0.16	0.044	0.201*	0.018	0.062	0.071	0.031	0.047	0.066	-0.005	0.106	0.045	-0.093	0.06	0.095	0.073	0.047	0.064	0.041	0.05
Pole+STSp	-0.036	-0.069	0.026	0.008	0.01	0.002	0.051	-0.025	0.047	-0.081	0.201*	-0.026	0.124*	0.01	-0.056	0.004	-0.035	0.001	0.002	0.005	-0.036	0.017	-0.016	0.028	0.012
T1a+T1a.HeschlR	0.002	0.118*	-0.057	0.1	0.013	0.086	0.083	-0.072	0.08	0.05	0.047	0.197*	0.134*	0.01	0.134*	0.005	-0.009	-0.002	0.001	0.004	0.021	0.016	0.053	0.06	-0.005
Tzml+T1a	0.011	0.196*	0.007	0.017	-0.056	0.056	0.115*	-0.002	0.067	-0.05	-0.031	0.006	-0.001	-0.009	0.032	-0.009	-0.013	0.061	0.013	-0.047	-0.018	-0.01	0.021	-0.002	-0.004
RoIS+TiR	-0.079	0.095	0.038	0.195*	0.041	0.002	-0.024	0.063	0.034	0.057	0.01	-0.027	0.045	-0.008	-0.027	0.032	0.069	0.003	0.076	0.040	-0.005	0.041	0.032	0.011	0.021
Pole+F3td	-0.071	0.017	0.192*	0.056	0.04	-0.039	0.002	0.079	-0.019	0.021	0.025	0.045	0.012	-0.045	-0.006	0.022	0.024	-0.004	-0.006	-0.044	0.019	-0.03	0.057	-0.041	-0.002
PrF3op+TiR	0.192*	-0.002	-0.006	-0.041	-0.072	0.081	-0.017	-0.035	-0.003	0.045	-0.021	-0.041	-0.037	0.013	-0.032	-0.042	-0.067	-0.008	0.017	0.012	-0.012	0.023	0.059	-0.037	-0.005
Tzml+STSp	0.0	0.077	-0.012	0.062	0.105	0.008	0.01	-0.032	0.032	0.017	0.180*	0.139*	-0.045	0.022	0.013	-0.0	-0.003	0.005	-0.056	0.021	0.033	0.038	-0.013	0.015	0.015
F2p+STSp	-0.021	0.081	-0.018	0.009	0.004	0.078	0.185*	-0.052	0.13*	-0.026	0.052	0.035	0.068	-0.048	0.078	0.015	0.03	0.041	0.029	0.02	0.095	0.099	0.059	0.055	0.014
PrF3op+Tzml	-0.065	0.021	0.004	-0.03	0.032	0.185*	0.02	0.009	-0.042	-0.048	-0.045	0.043	0.07	0.055	0.086	-0.068	-0.006	0.022	0.032	0.004	-0.065	0.028	-0.022	-0.025	-0.038
PT+T1a.HeschlR	0.038	0.074	0.053	0.067	0.091	0.063	0.145*	0.01	0.185*	0.032	0.111	0.041	0.132*	0.027	0.132*	0.06	0.093	0.022	0.068	0.022	0.017	0.003	0.025	0.023	0.018
Tzml+TiR	-0.018	0.037	0.022	0.184*	-0.026	0.014	-0.035	0.073	0.047	-0.057	-0.02	-0.013	-0.005	0.003	0.038	0.013	0.1	0.004	0.099	0.017	-0.036	0.003	0.044	0.028	-0.017
Prec+F3tv	-0.051	0.065	0.153*	0.03	0.207*	-0.017	-0.012	0.041	0.0	-0.015	0.047	0.051	0.033	0.042	0.032	0.052	0.048	-0.025	-0.034	0.018	-0.007	0.002	0.009	-0.013	0.002
Pole+FzantR	-0.018	0.056	0.112	0.059	0.086	0.16	0.044	0.201*	0.018	0.062	0.071	0.031	0.047	0.066	-0.005	0.106	0.045	-0.093	0.06	0.095	0.073	0.047	0.064	0.041	0.05
Pole+STSp	-0.036	-0.069	0.026	0.008	0.01	0.002	0.051	-0.025	0.047	-0.081	0.201*	-0.026	0.124*	0.01	-0.056	0.004	-0.035	0.001	0.002	0.005	-0.036	0.017	-0.016	0.028	0.012
T1a+T1a.HeschlR	0.002	0.118*	-0.057	0.1	0.013	0.086	0.083	-0.072	0.08	0.05	0.047	0.197*	0.134*	0.01	0.134*	0.005	-0.009	-0.002	0.001	0.004	0.021	0.016	0.053	0.06	-0.005
Tzml+T1a	0.011	0.196*	0.007	0.017	-0.056	0.056	0.115*	-0.002	0.067	-0.05	-0.031	0.006	-0.001	-0.009	0.032	-0.009	-0.013	0.061	0.013						

Endophenotype	PfF3op	Taml	F3td	T1R	F3opd	F3p	T1a	F2antR	T1a_HeschR	AG	STSp	PT	F3orb	Pole	Prec	F3opdR	F3tv	T2pR	SMG	T2p	RoIS	Fusa	PreCR	T3p
T2p++F2antR	0.009	0.009	-0.053	-0.075	-0.058	0.03	-0.014	-0.147*	-0.079	-0.066	-0.032	0.006	0.023	0.008	0.003	-0.077	0.015	-0.048	-0.016	-0.055	0.01	-0.001	-0.072	-0.046
Precc+SMG	0.049	0.028	-0.054	0.037	0.001	0.025	0.026	-0.037	-0.005	0.04	-0.014	0.003	0.018	-0.022	0.024	-0.146*	0.029	0.03	0.444	0.105	-0.033	0.014	-0.038	0.051
SMG++T1R	-0.011	0.0	-0.01	-0.146*	0.004	-0.001	-0.027	-0.008	0.028	0.005	0.077	-0.082	-0.006	-0.021	0.037	-0.026	-0.03	-0.06	-0.032	0.071	-0.019	0.019	-0.001	-0.039
Precc+T1	-0.05	-0.02	0.055	0.05	0.063	-0.074	-0.026	0.065	-0.021	-0.047	0.013	-0.006	0.044	0.017	-0.012	0.143*	0.049	0.022	-0.026	0.007	0.049	0.011	-0.005	0.001
AG++T3p	0.031	-0.017	-0.033	-0.031	0.037	0.017	-0.045	0.029	0.093	-0.114*	-0.047	-0.045	0.005	0.004	-0.028	0.015	0.041	0.007	-0.013	-0.004	-0.007	-0.005	-0.013	-0.041
Pole++T1R	-0.11	-0.041	-0.045	0.141*	-0.061	-0.027	-0.07	0.029	0.034	-0.089	-0.048	-0.019	-0.019	-0.015	0.055	-0.049	0.049	0.027	0.055	-0.007	-0.053	-0.06	-0.017	-0.003
Precc+PF3op	-0.14*	0.047	0.022	0.054	0.056	-0.03	0.034	-0.016	0.03	-0.042	0.042	0.036	0.005	-0.032	0.038	-0.044	-0.022	-0.007	0.014	-0.026	-0.065	-0.07	-0.038	-0.007
STSp++SMG	-0.056	0.033	-0.031	-0.066	-0.055	-0.015	0.05	-0.077	-0.025	0.041	-0.14*	-0.054	-0.001	-0.014	0.03	-0.071	-0.105	-0.052	-0.074	-0.097	-0.049	-0.038	0.0	-0.01
Precc+Pole	-0.007	0.005	0.006	0.052	-0.015	-0.021	0.041	0.018	-0.046	0.056	0.052	0.002	0.007	0.023	0.101	0.18*	-0.017	-0.068	0.017	-0.122	0.004	-0.01	0.028	0.012
Taml++F3opdR	-0.0	-0.137*	-0.051	-0.035	-0.046	-0.051	-0.038	0.079	-0.037	-0.011	-0.042	-0.102	-0.106	-0.021	-0.065	-0.0	0.127*	-0.009	0.023	0.004	0.01	0.001	0.018	0.011
Fusa++T2p	0.065	0.045	-0.024	0.026	0.052	0.022	0.136*	0.008	0.07	0.077	0.008	0.089	0.055	-0.005	0.077	0.05	0.048	0.013	-0.031	0.039	0.02	0.053	0.057	0.06
T1a++STSp	-0.071	-0.058	0.054	-0.063	0.038	0.052	-0.064	-0.003	-0.059	-0.135*	0.114*	-0.036	-0.031	-0.025	-0.014	-0.016	-0.022	0.023	-0.02	0.011	0.02	-0.032	-0.021	0.01
Taml++T1a_HeschR	-0.017	0.041	0.012	-0.025	-0.024	0.01	-0.063	0.023	0.043	0.021	-0.053	-0.134*	-0.031	-0.037	-0.054	0.036	0.059	0.007	0.035	0.047	0.022	0.002	-0.025	0.008
F2p++AG	0.046	0.039	-0.047	0.047	-0.027	-0.117*	-0.001	-0.007	-0.037	0.134*	-0.01	-0.038	-0.002	0.011	0.004	-0.019	0.008	-0.048	0.009	0.019	-0.029	-0.008	0.044	-0.003
Precc+T2ml	0.033	-0.038	-0.067	-0.081	-0.132*	-0.095	-0.051	-0.026	-0.069	-0.043	-0.058	-0.117*	-0.048	-0.022	-0.102	0.131*	-0.035	-0.039	-0.042	-0.012	-0.066	-0.023	-0.063	-0.019
F3opd++F3p	0.13*	0.029	0.046	-0.018	0.06	0.076	0.007	-0.064	-0.012	0.053	0.031	0.04	0.029	0.038	0.023	0.023	0.026	-0.029	-0.025	0.013	-0.047	0.013	0.031	0.021
Pole++F3v	-0.022	0.043	-0.058	-0.056	-0.019	0.039	0.042	-0.097	0.018	0.036	-0.01	-0.034	-0.003	0.044	0.017	-0.013	-0.049	0.13*	-0.019	0.05	0.08	0.013	0.013	0.006
Taml++SMG	-0.021	-0.115*	-0.024	-0.048	-0.038	-0.017	-0.1	0.005	-0.018	-0.079	-0.123*	-0.041	-0.025	-0.033	-0.059	-0.008	-0.035	0.018	-0.05	-0.018	-0.045	0.005	-0.051	-0.029
F3tv++F2antR	0.083	0.067	0.067	0.06	0.015*	0.037	0.058	-0.13*	0.029	0.045	0.054	0.088	0.078	0.059	0.037	0.031	0.015	0.023	0.05	0.078	0.059	0.011	0.052	0.022
F2p++F2antR	-0.019	-0.004	-0.002	0.001	-0.027	0.019	0.011	0.099	0.0	-0.037	-0.06	-0.04	-0.028	-0.039	-0.003	0.042	0.027	-0.025	0.015	-0.072	-0.033	-0.004	-0.046	0.005
Pole++T3p	-0.019	-0.001	-0.005	0.054	0.002	0.019	0.129*	-0.026	0.063	0.023	0.084	0.05	0.125*	0.059	0.078	0.035	0.05	0.076	0.052	0.022	0.022	0.081	-0.008	0.098
F3tv++STSp	-0.006	0.057	-0.049	-0.068	-0.024	0.002	-0.045	0.03	0.112	-0.123*	0.014	-0.013	0.011	0.024	-0.004	-0.065	0.022	-0.066	0.036	-0.021	-0.011	-0.015	0.015	-0.058
T1++F3orb	-0.061	-0.028	-0.060	-0.083	-0.022	-0.05	-0.0	-0.056	-0.078	-0.068	0.019	-0.029	-0.084	-0.079	-0.042	-0.056	0.044	-0.038	-0.058	-0.067	-0.048	-0.076	-0.049	-0.046
AG++T2p	0.04	0.128*	-0.02	0.091	0.048	0.06	0.06	0.037	0.071	0.12*	-0.0	0.056	0.051	0.055	0.086	0.018	0.095	0.037	0.07	0.042	0.035	-0.016	0.038	-0.009
F3tv++SMG	0.081	0.068	0.113	0.075	0.127*	-0.035	0.008	0.104	-0.04	0.066	0.083	0.033	0.005	0.033	0.084	0.054	0.079	-0.118*	0.055	-0.017	0.007	-0.024	0.047	-0.049
Precc+T2ml	-0.009	0.074	-0.018	0.049	-0.053	-0.015	0.052	-0.02	0.042	0.02	0.024	-0.002	0.01	-0.004	0.049	-0.028	-0.011	-0.021	0.126*	-0.017	0.016	-0.026	-0.037	-0.018
PF3op++RoIS	0.015	-0.023	0.017	-0.023	0.038	0.03	-0.064	-0.019	-0.041	0.044	0.025	0.015	0.005	0.014	0.023	0.014	0.049	0.014	0.054	0.024	0.014	0.054	0.022	0.018
Precc+T1	0.017	0.017	0.039	-0.03	-0.019	0.004	0.051	0.019	-0.018	0.08	0.047	0.043	0.022	0.048	0.068	0.124*	-0.039	-0.004	-0.01	-0.005	0.005	0.026	0.031	0.004
Taml++F3v	-0.043	0.014	-0.124*	-0.027	-0.104	0.032	0.061	-0.12*	0.063	0.029	0.011	-0.052	-0.019	0.004	0.025	-0.049	-0.091	0.111	-0.084	-0.024	-0.005	-0.033	0.028	-0.024
Precc+F2antR	-0.122*	0.015	-0.045	0.015	0.023	-0.014	0.011	-0.031	0.024	-0.028	-0.033	-0.028	-0.009	0.013	-0.073	-0.1	0.036	-0.035	0.022	0.026	0.019	-0.047	-0.009	-0.027
Fusa++STSp	-0.021	-0.115*	-0.024	-0.048	-0.038	-0.017	-0.1	0.005	-0.018	-0.079	-0.123*	-0.041	-0.025	-0.033	-0.059	-0.008	-0.035	0.018	-0.05	-0.018	-0.045	0.005	-0.051	-0.029
F3opd++AG	0.024	0.055	-0.07	0.063	0.123*	0.058	0.102	-0.048	0.025	0.068	0.027	-0.024	0.025	-0.009	-0.019	-0.039	0.052	-0.005	0.035	0.002	0.019	0.026	-0.008	-0.041
PT++T2pR	0.015	-0.025	0.018	-0.123*	0.005	0.056	-0.066	-0.057	-0.065	-0.016	0.008	-0.027	-0.029	0.016	-0.026	-0.004	-0.017	0.03	-0.086	0.008	-0.024	0.005	-0.047	-0.038
F3tv++T1a_HeschR	-0.021	-0.044	-0.033	0.007	-0.053	-0.026	-0.078	0.005	-0.122*	-0.017	-0.071	-0.061	-0.039	-0.051	0.013	-0.037	0.025	-0.026	0.026	-0.002	0.015	-0.017	0.037	-0.014
T1a++F3opdR	-0.012	-0.05	-0.024	-0.035	-0.048	-0.014	0.014	-0.02	-0.02	-0.031	-0.017	-0.07	-0.004	-0.02	0.022	0.024	-0.122*	0.005	0.03	-0.033	-0.029	-0.036	0.033	0.01
SMG++F3opdR	0.041	0.018	0.017	-0.023	0.038	0.01	0.024	-0.017	-0.013	0.041	-0.048	0.021	0.012	0.043	-0.051	-0.035	0.112	0.043	0.121*	0.002	0.039	0.045	0.023	-0.021
F3orb++T2pR	-0.042	-0.023	-0.036	0.038	-0.033	-0.024	-0.017	-0.028	-0.033	-0.04	-0.052	-0.002	-0.029	-0.084	-0.049	-0.011	-0.015	-0.035	0.121*	-0.027	-0.022	-0.003	-0.011	0.006
F3opd++F3orb	0.002	0.093	0.073	0.072	0.12*	0.059	0.047	0.022	0.036	0.052	0.08	0.063	0.01	0.044	0.077	0.029	0.04	-0.033	0.024	0.011	-0.002	-0.002	0.072	0.016
AG++T1a_HeschR	-0.017	0.001	-0.066	-0.005	-0.058	-0.013	-0.006	-0.017	-0.053	-0.051	-0.048	-0.042	-0.046	-0.042	0.003	-0.096	-0.007	-0.074	-0.12*	-0.107	-0.038	-0.035	0.003	-0.104
F3opd++T1R	-0.021	-0.025	-0.002	-0.098	0.05	-0.022	-0.008	-0.024	-0.02	-0.023	0.223	-0.038	-0.005	0.052	0.068	0.019	0.023	0.022	-0.073	0.036	0.007	-0.012	-0.026	0.017
F2p++T1a	-0.05	-0.055	-0.046	-0.073	-0.118*	0.084	-0.061	0.008	-0.009	-0.043	-0.075	-0.018	-0.048	-0.009	-0.069	-0.053	-0.037	-0.047	-0.012	0.008	-0.031	-0.013	-0.014	-0.013
PF3op++F3td	-0.116*	0.053	0.107	0.019	0.017*	-0.052	0.013	0.021	-0.019	-0.002	-0.018	0.039	0.036	0.047	0.01	0.058	0.023	0.033	-0.009	0.019	-0.035	0.039	0.013	-0.002
F2p++T2pR	0.004	-0.039	0.019	0.004	-0.03	-0.117*	-0.027	0.005	-0.051	-0.067	-0.081	-0.038	-0.013	-0.038	-0.036	-0.008	0.011	-0.033	-0.063	-0.024	-0.039	-0.036	-0.024	-0.058
T1a++T2pR	-0.003																							

C.1. (Endo)phenotypic correlation with other resting-state functional connectivity and neural activations

Endophenotype	PrF3op	Tm	F3td	TR	F3opd	F2p	T1a	FzantR	T1a_HeschlR	AG	ST5p	T1	PT	F3orb	Pole	Prec	F3opdR	F3tv	T2pR	SMG	T2p	RoLS	Fusa	PrecR	T3p
PT++F3orb	0.0	0.004	0.333	0.035	0.03	-0.022	-0.003	0.058	0.015	0.009	0.01	-0.071	0.002	0.026	0.035	0.055	0.038	-0.028	0.058	0.093	0.01	-0.001	0.044	0.049	0.011
SMG++F3orb	-0.01	-0.093	-0.019	-0.021	-0.022	-0.028	-0.05	-0.008	0.014	-0.006	-0.042	-0.064	-0.02	-0.066	-0.023	0.103	0.024	0.082	-0.023	0.047	0.003	0.009	-0.031	0.038	0.01
ST5p++T1R	-0.07	-0.071	-0.055	-0.066	-0.031	-0.029	-0.009	-0.054	-0.092	-0.058	0.019	0.031	-0.007	0.006	0.038	-0.056	-0.042	-0.013	-0.089	-0.02	-0.079	-0.034	-0.034	-0.062	-0.071
Fusa++SMG	0.078	0.005	0.07	0.063	0.084	0.039	-0.027	0.015	0.02	0.071	0.008	-0.001	0.013	0.059	0.051	0.092	0.041	0.066	0.021	0.058	0.011	0.067	-0.002	0.082	-0.009
Prec++AG	0.015	-0.022	0.03	-0.044	0.076	0.091	-0.031	0.053	-0.037	-0.036	0.023	-0.024	-0.002	0.01	0.027	-0.038	0.038	0.035	0.042	0.063	0.092	0.03	0.021	-0.001	0.081
F3opdR++T2pR	-0.033	0.059	-0.027	-0.005	-0.04	0.059	0.087	-0.006	0.012	0.042	0.056	0.021	-0.007	0.039	0.078	-0.057	-0.091	0.048	-0.049	0.015	0.037	0.015	0.044	-0.013	-0.007
T1a++PrecR	0.051	0.003	-0.025	-0.001	0.002	-0.062	0.023	-0.008	-0.073	-0.047	-0.015	0.04	-0.051	0.007	-0.021	-0.006	-0.038	0.054	-0.051	-0.011	-0.016	0.009	-0.01	0.012	-0.014
T1a++F3tv	-0.013	-0.037	-0.012	0.017	-0.016	-0.041	0.033	0.052	0.049	0.033	0.015	-0.001	0.024	0.012	0.022	-0.002	0.065	0.013	0.059	0.02	-0.003	0.045	0.017	-0.018	0.017
ST5p++T2p	-0.003	-0.012	-0.068	-0.054	-0.091	0.03	-0.075	-0.015	-0.03	0.01	-0.056	-0.032	-0.046	-0.022	-0.012	-0.048	-0.069	-0.013	0.027	0.02	-0.08	-0.076	-0.065	-0.044	-0.03
FzantR++T1a_HeschlR	-0.044	0.015	-0.007	0.007	-0.007	0.041	-0.024	-0.083	-0.091	-0.013	-0.033	0.004	-0.005	0.012	0.005	0.008	-0.024	-0.008	0.008	0.011	0.034	-0.005	0.02	0.031	0.038
Pole++PT	0.001	-0.081	-0.028	-0.054	-0.027	-0.049	-0.082	0.003	-0.058	-0.021	-0.046	-0.049	0.091	0.022	-0.003	-0.011	-0.018	-0.021	-0.063	-0.019	-0.006	-0.007	-0.052	-0.025	-0.056
Fusa++FzantR	-0.022	0.007	-0.031	0.053	-0.024	-0.032	-0.001	0.022	-0.009	-0.007	-0.069	-0.042	-0.01	0.039	-0.059	-0.003	-0.014	-0.038	-0.023	-0.027	-0.05	-0.012	-0.061	-0.01	-0.073
RoLS++T2pR	-0.004	-0.070	-0.06	-0.035	-0.062	-0.081	-0.051	-0.078	-0.047	-0.082	-0.09	0.03	-0.047	-0.233	0.018	-0.028	-0.003	-0.029	-0.051	-0.064	-0.06	-0.045	-0.031	-0.046	0.003
F2p++T2p	-0.002	-0.017	0.002	-0.029	-0.021	-0.059	-0.026	0.011	-0.042	-0.016	-0.069	-0.005	-0.054	-0.057	-0.063	-0.019	-0.005	-0.09	0.025	-0.065	-0.01	0.035	-0.028	-0.037	0.011
PrF3op++F3orb	0.09	0.005	-0.036	-0.0	-0.025	-0.029	-0.026	0.023	0.012	0.001	-0.02	0.06	-0.029	0.054	-0.012	0.017	-0.017	0.03	0.036	-0.007	0.016	0.055	-0.02	0.011	-0.013
T1a++SMG	0.009	-0.006	-0.089	-0.025	-0.08	0.014	-0.011	-0.062	0.036	-0.03	-0.045	0.004	-0.021	-0.022	-0.045	-0.012	-0.068	0.011	-0.061	0.019	-0.026	-0.0	-0.047	0.02	-0.019
Prec++F3orb	-0.028	-0.007	0.032	-0.017	-0.042	-0.068	-0.046	0.017	-0.068	-0.038	-0.052	-0.089	-0.068	0.035	-0.026	0.019	-0.03	-0.047	-0.002	-0.009	-0.005	-0.057	-0.074	-0.023	-0.042
F3tv++PrecR	-0.037	-0.027	0.018	-0.039	-0.012	0.014	0.029	-0.002	-0.036	0.007	0.027	0.01	-0.032	0.001	-0.063	-0.019	-0.006	-0.03	0.041	-0.052	0.012	-0.017	0.001	-0.032	0.021
F3opd++T2p	0.002	0.014	0.008	0.038	-0.032	0.013	-0.042	0.02	0.011	-0.026	-0.028	0.089	0.072	-0.033	-0.006	0.085	-0.031	0.039	0.004	0.028	0.052	0.095	0.053	-0.005	0.026
ST5p++PrecR	-0.046	0.008	-0.043	0.016	-0.003	0.002	0.033	-0.043	0.046	-0.049	0.088	-0.002	0.026	-0.038	0.029	-0.038	-0.02	-0.012	0.004	0.007	0.032	0.009	-0.006	-0.015	0.006
ST5p++RoLS	0.033	-0.065	0.002	-0.012	-0.009	0.037	-0.04	0.061	0.006	-0.023	0.088	0.031	-0.007	0.001	0.009	-0.032	0.076	0.03	-0.016	-0.045	0.014	-0.022	0.003	0.023	0.002
Fusa++PrecR	-0.035	-0.054	0.012	0.036	-0.005	-0.049	-0.043	-0.011	-0.014	-0.066	-0.087	0.017	-0.02	-0.009	-0.047	0.009	0.037	0.005	-0.003	-0.035	-0.0	0.02	0.048	0.026	0.036
Fusa++F3opdR	-0.007	-0.006	-0.025	-0.003	-0.065	0.059	0.015	-0.016	-0.037	0.034	0.025	-0.043	-0.048	0.019	0.008	-0.049	-0.087	0.033	0.076	-0.031	-0.021	-0.01	-0.038	0.014	0.0
Prec++F2p	0.086	0.016	-0.031	0.004	-0.071	0.026	-0.027	0.002	0.024	-0.012	0.006	0.027	0.025	-0.011	-0.024	-0.013	0.03	-0.034	-0.005	0.006	-0.022	-0.008	-0.026	0.007	0.018
F3orb++F3orb	-0.016	-0.002	-0.014	-0.0	-0.015	0.026	0.014	-0.003	0.016	0.015	-0.026	-0.011	-0.027	0.086	0.0	0.026	-0.008	0.026	0.011	0.017	-0.027	-0.005	-0.051	-0.046	-0.004
F3orb++FzantR	0.045	-0.004	0.073	0.071	0.042	-0.085	-0.031	0.023	0.054	-0.046	-0.039	0.026	-0.003	-0.084	-0.023	0.024	0.074	0.001	0.037	0.004	-0.027	-0.011	-0.003	0.033	-0.018
AG++T2pR	-0.03	0.068	0.047	0.064	0.049	-0.027	0.017	0.045	0.04	0.034	0.035	0.085	0.068	0.003	0.006	0.02	0.05	-0.009	-0.04	-0.023	0.039	0.051	0.01	0.046	0.026
T1a++F3td	-0.055	0.024	0.084	0.057	0.023	0.009	-0.037	0.041	-0.011	-0.059	-0.034	-0.034	-0.018	-0.034	-0.014	0.034	-0.004	-0.056	0.052	-0.008	-0.012	-0.02	0.02	0.014	-0.021
F3tv++RoLS	0.001	0.041	-0.052	0.041	-0.055	-0.02	0.084	-0.058	0.037	0.005	0.004	0.020	0.014	0.003	0.031	-0.034	0.004	0.012	0.051	-0.04	0.011	0.043	-0.015	0.079	-0.0
T2p++T1a_HeschlR	-0.021	-0.027	0.018	-0.039	-0.012	0.014	0.029	-0.002	-0.036	-0.011	0.002	-0.015	-0.021	0.001	-0.063	-0.019	-0.006	-0.03	0.041	-0.052	0.012	-0.017	0.001	-0.032	0.021
T1a++T2p	0.01	-0.004	0.024	0.015	0.014	-0.042	-0.039	-0.025	0.022	-0.037	0.002	0.04	0.083	0.013	0.015	0.015	0.01	0.031	0.007	-0.012	-0.022	0.046	0.002	0.03	-0.013
F3opd++T2pR	0.026	-0.034	-0.049	-0.002	-0.073	0.082	-0.019	-0.018	0.007	0.013	-0.019	-0.028	-0.005	-0.033	0.009	0.042	-0.002	0.045	0.057	0.027	0.015	0.068	-0.013	0.065	-0.009
FzantR++T1a_HeschlR	0.022	0.055	-0.017	0.034	0.0	0.053	0.073	-0.009	0.047	-0.049	0.025	0.044	0.061	-0.005	0.027	-0.062	0.03	-0.032	-0.082	-0.014	-0.021	-0.046	-0.051	-0.052	-0.023
F3td++T1a_HeschlR	0.016	-0.033	-0.057	-0.008	-0.038	-0.025	-0.004	-0.051	0.02	-0.017	-0.029	-0.015	-0.026	-0.014	-0.046	-0.028	-0.009	0.008	-0.038	-0.005	0.004	-0.005	-0.082	-0.003	-0.038
F3opd++T2p	-0.018	-0.077	-0.019	-0.028	-0.031	-0.059	-0.071	0.039	-0.027	-0.02	-0.08	-0.073	-0.082	-0.017	-0.026	-0.005	-0.027	-0.072	0.001	-0.015	-0.01	-0.025	-0.008	-0.03	0.038
F2p++PrF3op	-0.021	0.019	0.017	0.015	0.044	0.052	0.017	0.027	0.006	-0.015	-0.008	0.008	0.022	0.036	0.057	0.012	0.024	0.036	-0.019	0.024	-0.015	0.001	-0.016	-0.024	0.015
T2p++PrecR	-0.005	-0.082	-0.046	-0.033	-0.071	-0.018	-0.025	-0.035	0.016	-0.081	-0.046	-0.042	-0.032	-0.024	-0.061	-0.039	-0.018	-0.009	-0.031	-0.08	0.01	-0.007	-0.023	0.015	0.062
Fusa++T1R	-0.026	0.048	0.001	-0.022	-0.008	0.023	0.036	-0.029	-0.013	0.061	0.051	0.014	0.033	0.02	0.003	0.027	0.002	0.024	0.003	0.036	0.015	0.012	-0.014	0.053	0.002
Prec++T2p	0.057	0.001	0.037	-0.012	0.004	0.039	-0.039	0.012	-0.023	0.081	0.046	0.02	0.001	0.022	-0.02	-0.001	0.02	-0.009	0.021	-0.014	0.064	0.032	0.059	-0.05	-0.081
T1a++PrecR	-0.03	-0.041	0.003	0.022	-0.06	-0.025	-0.066	0.014	-0.019	-0.061	-0.062	0.02	-0.031	-0.024	-0.045	-0.011	-0.013	-0.033	-0.052	-0.028	-0.081	-0.036	-0.027	0.004	-0.07
Tamli++T2p	0.001	-0.029	0.013	0.014	-0.045	0.047	0.017	-0.044	-0.044	-0.011	0.011	0.029	0.015	-0.005	-0.021	0.019	0.029	0.046	0.01	0.068	-0.012	0.024	0.024	0.053	0.003
T2p++T2pR	0.032	0.049	0.028	-0.002	0.064	0.015	0.073	-0.007	0.053	0.014	0.065	0.051	0.027	0.02	0.075	0.059	0.078	-0.026	-0.043	0.027	0.047	-0.001	0.034	0.028	0.081
PrF3op++PrecR	-0.0	-0.002	0.029	0.038	0.007	-0.013	-0.005	0.031	0.072	-0.045	0.039	0.064	0.053	0.014	-0.02	0.031	0.032	0.081	-0.025	0.021	0.01	0.031	0.005	0.016	0.002
Pole++T1	-0.045	-0.08	0.079	-0.058	0.046	-0.05	0.009	0.016	-0.076	0.064	0.039	0.062	-0.027	-0.019	-0.044	-0.005	-0.024	-0.011	-0.037	-0.025	0.027	0.011	0.001	-0.021	0.029
T1a++F3orb	-0.08	-0.065	-0.012	-0.063	-0.046	0.002	-0.038	0.063	-0.048	-0.027	-0.038	-0.066	-0.021	-0.028	0.003	-0.066	-0.041	-0.044	-0.056	0.008	-0.032	-0.028	-0.035	-0.06	-0.059
F3opdR++T1a_HeschlR	-0.003	-0.01	0.026	-0.013	0.048	0.017	-0.043	0.01	-0.065	-0.023	-0.013	-0.001	0.001												

C.2 (Endo)phenotypic correlation between both resting-state functional connectivity and language-related scores

Table C.2 – Pearson Correlation between both resting-state functional connectivity and language-related scores. No significant correlation was observed after false discovery rate correction at $p < 0.05$.

Endophenotype	Language task overall accuracy			Working Memory			Oral reading recognition			Vocabulary Comprehension		
	r	p value	FDR p value	r	p value	FDR p value	r	p value	FDR p value	r	p value	FDR p value
PT↔PrF3op	0.111	1.77E-03	1.76E-01	0.051	1.49E-01	5.74E-01	0.137	1.13E-04	1.36E-01	0.123	5.34E-04	1.49E-01
F3td↔RolS	-0.122	6.22E-04	1.49E-01	0.018	6.17E-01	8.95E-01	-0.032	3.63E-01	7.77E-01	-0.001	9.85E-01	9.93E-01
T2ml↔T1a.HeschlR	-0.080	2.44E-02	3.29E-01	-0.040	2.59E-01	6.96E-01	-0.126	4.07E-04	1.49E-01	-0.049	1.69E-01	6.00E-01
PrF3op↔T3p	0.055	1.21E-01	5.45E-01	0.024	4.98E-01	8.69E-01	0.129	2.75E-04	1.49E-01	0.035	3.22E-01	7.57E-01
F3td↔F3orb	0.108	2.31E-03	1.76E-01	0.033	3.50E-01	7.74E-01	0.089	1.22E-02	2.98E-01	0.088	1.35E-02	2.98E-01
F3opdR↔T1a.HeschlR	0.009	7.92E-01	9.49E-01	0.032	3.73E-01	7.82E-01	0.109	2.29E-03	1.76E-01	0.059	9.69E-02	4.78E-01
T1↔T3p	-0.040	2.66E-01	6.96E-01	-0.009	8.11E-01	9.49E-01	-0.100	5.09E-03	2.61E-01	-0.111	1.81E-03	1.76E-01
SMG↔F3orb	-0.044	2.16E-01	6.44E-01	-0.013	7.20E-01	9.31E-01	-0.109	2.19E-03	1.76E-01	-0.069	5.43E-02	4.03E-01
F2p↔T2ml	0.102	4.07E-03	2.44E-01	0.023	5.25E-01	8.80E-01	0.108	2.35E-03	1.76E-01	0.038	2.84E-01	7.12E-01
T3p↔F3orb	0.090	1.18E-02	2.98E-01	0.041	2.51E-01	6.91E-01	0.109	2.25E-03	1.76E-01	0.083	1.92E-02	3.03E-01
Fusa↔F3orb	0.022	5.44E-01	8.80E-01	0.034	3.46E-01	7.73E-01	0.111	1.84E-03	1.76E-01	0.066	6.56E-02	4.15E-01
Fusa↔RolS	-0.048	1.79E-01	6.13E-01	-0.047	1.86E-01	6.13E-01	-0.111	1.78E-03	1.76E-01	-0.088	1.33E-02	2.98E-01
T1↔T1a.HeschlR	0.111	1.80E-03	1.76E-01	0.033	3.52E-01	7.74E-01	0.100	5.04E-03	2.61E-01	0.040	2.59E-01	6.96E-01
F2p↔AG	0.043	2.31E-01	6.67E-01	0.110	2.07E-03	1.76E-01	0.024	5.00E-01	8.69E-01	0.006	8.63E-01	9.69E-01
T1↔T1a	0.106	2.87E-03	2.03E-01	0.073	4.07E-02	3.89E-01	0.068	5.81E-02	4.03E-01	0.057	1.13E-01	5.27E-01
F3opdR↔F2antR	0.103	3.96E-03	2.44E-01	0.053	1.34E-01	5.64E-01	0.047	1.83E-01	6.13E-01	0.030	3.96E-01	7.97E-01
Pole↔T2ml	0.102	4.00E-03	2.44E-01	0.039	2.76E-01	7.09E-01	0.059	9.93E-02	4.85E-01	-0.054	1.29E-01	5.52E-01
Pole↔PrecR	-0.061	8.70E-02	4.56E-01	-0.100	4.91E-03	2.61E-01	-0.087	1.40E-02	2.98E-01	-0.050	1.58E-01	5.85E-01
T3p↔F2antR	-0.053	1.39E-01	5.67E-01	-0.047	1.85E-01	6.13E-01	-0.099	5.29E-03	2.61E-01	-0.097	6.22E-03	2.66E-01
T1↔F3orb	-0.048	1.76E-01	6.09E-01	0.003	9.42E-01	9.77E-01	-0.080	2.39E-02	3.29E-01	-0.099	5.43E-03	2.61E-01
T2ml↔F3tv	0.098	5.94E-03	2.64E-01	0.009	8.03E-01	9.49E-01	0.018	6.14E-01	8.95E-01	0.057	1.12E-01	5.27E-01
AG↔T1R	-0.098	5.91E-03	2.64E-01	-0.072	4.36E-02	3.89E-01	-0.087	1.40E-02	2.98E-01	-0.078	2.86E-02	3.35E-01
T2ml↔STSp	0.077	3.00E-02	3.36E-01	-0.014	6.87E-01	9.15E-01	0.097	6.61E-03	2.73E-01	0.052	1.46E-01	5.71E-01
T1↔AG	-0.063	7.64E-02	4.36E-01	-0.089	1.26E-02	2.98E-01	-0.085	1.67E-02	3.00E-01	-0.026	4.71E-01	8.59E-01
Pole↔F3td	-0.012	7.44E-01	9.43E-01	0.026	4.65E-01	8.55E-01	0.028	4.36E-01	8.33E-01	0.089	1.20E-02	2.98E-01
F3tv↔STSp	0.044	2.20E-01	6.51E-01	0.090	1.14E-02	2.98E-01	0.022	5.36E-01	8.80E-01	0.033	3.57E-01	7.76E-01
F3tv↔F3opdR	-0.088	1.31E-02	2.98E-01	-0.079	2.75E-02	3.35E-01	-0.061	8.48E-02	4.51E-01	-0.033	3.50E-01	7.74E-01
SMG↔F3opdR	0.093	9.18E-03	2.98E-01	0.062	8.18E-02	4.45E-01	0.020	5.75E-01	8.83E-01	0.002	9.62E-01	9.81E-01
Pole↔T1a.HeschlR	-0.043	2.28E-01	6.64E-01	-0.043	2.23E-01	6.54E-01	-0.090	1.18E-02	2.98E-01	-0.011	7.57E-01	9.46E-01
Fusa↔SMG	-0.024	4.99E-01	8.69E-01	0.016	6.54E-01	8.99E-01	0.072	4.34E-02	3.89E-01	0.090	1.10E-02	2.98E-01
PT↔T1a	0.053	1.39E-01	5.67E-01	0.092	1.01E-02	2.98E-01	0.033	3.59E-01	7.76E-01	-0.005	8.83E-01	9.71E-01
AG↔F3orb	0.020	5.68E-01	8.83E-01	0.020	5.76E-01	8.83E-01	0.090	1.14E-02	2.98E-01	0.028	4.37E-01	8.33E-01
Prec↔SMG	0.003	9.33E-01	9.77E-01	0.037	3.03E-01	7.36E-01	0.091	1.06E-02	2.98E-01	0.089	1.28E-02	2.98E-01
T3p↔T1a.HeschlR	-0.039	3.50E-01	7.74E-01	0.021	5.62E-01	8.83E-01	0.087	1.41E-02	2.98E-01	0.050	1.63E-01	5.94E-01
F2p↔T1	0.033	2.73E-01	7.09E-01	0.039	2.70E-01	7.04E-01	-0.095	7.53E-03	2.98E-01	-0.014	6.97E-01	9.22E-01
F3opd↔F2antR	-0.079	2.71E-02	3.35E-01	0.006	8.75E-01	9.70E-01	-0.052	1.45E-01	5.71E-01	-0.091	1.02E-02	2.98E-01
F3opd↔F3orb	0.061	8.50E-02	4.51E-01	0.024	5.07E-01	8.73E-01	0.087	1.46E-02	2.99E-01	0.089	1.27E-02	2.98E-01
F3opd↔F2p	0.040	2.63E-01	6.96E-01	-0.090	1.17E-02	2.98E-01	0.035	3.20E-01	7.57E-01	0.036	3.13E-01	7.52E-01
PrF3op↔RolS	-0.050	1.59E-01	5.86E-01	-0.087	1.41E-02	2.98E-01	-0.034	3.34E-01	7.64E-01	-0.027	4.53E-01	8.48E-01
RolS↔T2pR	0.035	3.32E-01	7.63E-01	0.007	8.46E-01	9.64E-01	0.093	8.80E-03	2.98E-01	0.043	2.31E-01	6.67E-01
T2ml↔T1a	-0.010	7.69E-01	9.49E-01	0.009	7.93E-01	9.49E-01	-0.064	7.20E-02	4.29E-01	-0.092	9.45E-03	2.98E-01
T2ml↔F3td	-0.015	6.76E-01	9.09E-01	0.011	7.54E-01	9.46E-01	0.048	1.78E-01	6.12E-01	0.088	1.30E-02	2.98E-01
Prec↔F3orb	-0.074	3.72E-02	3.83E-01	-0.007	8.38E-01	9.62E-01	-0.090	1.11E-02	2.98E-01	-0.047	1.89E-01	6.14E-01
T1a↔T1R	0.033	3.53E-01	7.74E-01	0.089	1.28E-02	2.98E-01	0.068	5.68E-02	4.03E-01	0.051	1.54E-01	5.82E-01
F3opd↔T3p	0.065	6.80E-02	4.23E-01	0.073	3.94E-02	3.89E-01	0.087	1.47E-02	2.99E-01	0.020	5.84E-01	8.84E-01
T1a.HeschlR↔T1R	0.037	2.93E-01	7.27E-01	0.086	1.59E-02	3.00E-01	0.056	1.13E-01	5.27E-01	0.049	1.69E-01	6.00E-01
F3opd↔T2pR	0.010	7.73E-01	9.49E-01	-0.012	7.42E-01	9.41E-01	-0.062	8.17E-02	4.45E-01	-0.085	1.69E-02	3.00E-01
F3opd↔T1a	0.033	3.51E-01	7.74E-01	0.007	8.53E-01	9.64E-01	0.085	1.76E-02	3.00E-01	0.062	8.39E-02	4.51E-01
Prec↔PrF3op	0.085	1.74E-02	3.00E-01	0.042	2.41E-01	6.80E-01	0.036	3.16E-01	7.54E-01	0.022	5.36E-01	8.80E-01
Prec↔F2p	-0.014	7.05E-01	9.26E-01	-0.070	4.82E-02	4.03E-01	-0.023	5.14E-01	8.77E-01	-0.084	1.81E-02	3.00E-01
T2ml↔PrecR	-0.033	3.58E-01	7.76E-01	0.084	1.84E-02	3.00E-01	-0.026	4.63E-01	8.55E-01	-0.053	1.38E-01	5.67E-01
T2ml↔F2antR	-0.011	7.59E-01	9.46E-01	-0.085	1.74E-02	3.00E-01	0.004	9.16E-01	9.74E-01	0.035	3.32E-01	7.63E-01
AG↔F3opdR	-0.085	1.65E-02	3.00E-01	-0.057	1.07E-01	5.12E-01	-0.056	1.19E-01	5.40E-01	-0.035	3.28E-01	7.63E-01
AG↔F3tv	0.027	4.51E-01	8.48E-01	-0.032	3.75E-01	7.82E-01	0.086	1.55E-02	3.00E-01	-0.006	8.71E-01	9.69E-01
AG↔F3td	0.004	9.12E-01	9.74E-01	-0.056	1.13E-01	5.27E-01	-0.083	1.99E-02	3.11E-01	-0.084	1.84E-02	3.00E-01
T1a↔T2pR	0.086	1.55E-02	3.00E-01	0.026	4.60E-01	8.54E-01	0.066	6.31E-02	4.11E-01	0.045	2.04E-01	6.31E-01
SMG↔PrecR	0.077	3.15E-02	3.47E-01	0.038	2.84E-01	7.12E-01	0.015	6.75E-01	9.09E-01	0.084	1.81E-02	3.00E-01
F2antR↔T1R	0.052	1.48E-01	5.73E-01	-0.013	7.23E-01	9.33E-01	0.085	1.73E-02	3.00E-01	0.066	6.47E-02	4.13E-01

C.2. (Endo)phenotypic correlation between both resting-state functional connectivity and language-related scores

173

Endophenotype	Language task overall accuracy			Working Memory			Oral reading recognition			Vocabulary Comprehension		
	r	p value	FDR p value	r	p value	FDR p value	r	p value	FDR p value	r	p value	FDR p value
Prec↔T2pR	-0.084	1.85E-02	3.00E-01	-0.061	8.48E-02	4.51E-01	-0.002	9.51E-01	9.79E-01	-0.006	8.71E-01	9.69E-01
T2p↔T2pR	-0.026	4.73E-01	8.59E-01	-0.040	2.62E-01	6.96E-01	-0.084	1.89E-02	3.03E-01	-0.017	6.28E-01	8.95E-01
T1↔F2antR	-0.017	6.38E-01	8.97E-01	-0.001	9.78E-01	9.90E-01	0.031	3.89E-01	7.95E-01	0.081	2.37E-02	3.29E-01
T1a↔PrecR	-0.009	8.11E-01	9.49E-01	-0.018	6.14E-01	8.95E-01	0.045	2.06E-01	6.31E-01	0.081	2.22E-02	3.29E-01
F3opdR↔T1R	0.031	3.88E-01	7.95E-01	-0.080	2.43E-02	3.29E-01	0.009	7.93E-01	9.49E-01	0.029	4.18E-01	8.16E-01
PrF3op↔T2p	-0.039	2.74E-01	7.09E-01	-0.069	5.31E-02	4.03E-01	0.082	2.17E-02	3.29E-01	0.047	1.86E-01	6.13E-01
T1a↔STSp	-0.073	4.12E-02	3.89E-01	-0.081	2.27E-02	3.29E-01	-0.025	4.76E-01	8.62E-01	-0.035	3.20E-01	7.57E-01
T1↔PrecR	-0.022	5.40E-01	8.80E-01	-0.082	2.17E-02	3.29E-01	-0.043	2.32E-01	6.68E-01	-0.078	2.92E-02	3.35E-01
F2p↔T1a	-0.066	6.35E-02	4.11E-01	-0.038	2.89E-01	7.20E-01	-0.080	2.39E-02	3.29E-01	-0.020	5.69E-01	8.83E-01
SMG↔T1a.HeschlR	0.022	5.33E-01	8.80E-01	0.080	2.44E-02	3.29E-01	0.044	2.13E-01	6.40E-01	0.016	6.62E-01	9.01E-01
Fusa↔F3opdR	-0.051	1.55E-01	5.82E-01	-0.002	9.58E-01	9.79E-01	-0.080	2.47E-02	3.29E-01	-0.068	5.71E-02	4.03E-01
T1↔F3tv	-0.081	2.28E-02	3.29E-01	-0.069	5.34E-02	4.03E-01	-0.001	9.83E-01	9.93E-01	-0.060	9.32E-02	4.70E-01
PT↔RolS	-0.082	2.20E-02	3.29E-01	0.006	8.70E-01	9.69E-01	-0.043	2.26E-01	6.61E-01	0.009	8.07E-01	9.49E-01
PT↔F3tv	-0.018	6.08E-01	8.95E-01	-0.030	4.05E-01	8.03E-01	-0.078	2.88E-02	3.35E-01	-0.080	2.56E-02	3.35E-01
T1a↔F3td	0.078	2.90E-02	3.35E-01	0.051	1.56E-01	5.82E-01	0.017	6.24E-01	8.95E-01	0.058	1.02E-01	4.94E-01
Fusa↔F3td	0.044	2.18E-01	6.49E-01	-0.009	7.94E-01	9.49E-01	0.078	2.77E-02	3.35E-01	0.009	7.97E-01	9.49E-01
Fusa↔F2antR	-0.079	2.72E-02	3.35E-01	-0.024	5.08E-01	8.74E-01	-0.043	2.31E-01	6.67E-01	-0.046	1.99E-01	6.28E-01
T1↔STSp	0.015	6.76E-01	9.09E-01	0.079	2.75E-02	3.35E-01	0.068	5.69E-02	4.03E-01	0.068	5.56E-02	4.03E-01
F2p↔RolS	0.005	8.98E-01	9.72E-01	-0.030	4.03E-01	8.02E-01	-0.078	2.92E-02	3.35E-01	-0.064	7.46E-02	4.33E-01
Pole↔F3opdR	-0.078	2.89E-02	3.35E-01	-0.021	5.49E-01	8.80E-01	-0.048	1.74E-01	6.08E-01	0.010	7.89E-01	9.49E-01
RolS↔F2antR	0.036	3.19E-01	7.57E-01	0.035	3.30E-01	7.63E-01	-0.040	2.63E-01	6.96E-01	0.078	2.90E-02	3.35E-01
F3opd↔SMG	-0.025	4.90E-01	8.68E-01	-0.048	1.74E-01	6.08E-01	-0.078	2.96E-02	3.35E-01	-0.053	1.38E-01	5.67E-01
T1a↔T1a.HeschlR	0.079	2.75E-02	3.35E-01	0.066	6.50E-02	4.13E-01	0.031	3.81E-01	7.88E-01	0.011	7.88E-01	9.49E-01
F2p↔F3tv	-0.009	7.94E-01	9.49E-01	0.078	2.96E-02	3.35E-01	0.001	9.74E-01	9.89E-01	0.060	9.29E-02	4.70E-01
Prec↔PrecR	-0.020	5.81E-01	8.83E-01	-0.030	3.98E-01	7.99E-01	0.076	3.24E-02	3.47E-01	0.073	4.04E-02	3.89E-01
PT↔F2antR	-0.028	4.35E-01	8.33E-01	-0.022	5.36E-01	8.80E-01	-0.074	3.74E-02	3.83E-01	-0.076	3.21E-02	3.47E-01
F3opd↔PrecR	0.007	8.47E-01	9.64E-01	-0.004	9.08E-01	9.74E-01	-0.076	3.21E-02	3.47E-01	-0.017	6.30E-01	8.95E-01
PrF3op↔F3opdR	0.076	3.23E-02	3.47E-01	0.052	1.46E-01	5.71E-01	0.027	4.43E-01	8.41E-01	-0.021	5.59E-01	8.83E-01
Fusa↔STSp	-0.052	1.44E-01	5.71E-01	-0.066	6.50E-02	4.13E-01	-0.065	7.03E-02	4.27E-01	-0.075	3.48E-02	3.65E-01
PT↔F3td	-0.019	6.03E-01	8.94E-01	-0.075	3.50E-02	3.65E-01	-0.074	3.77E-02	3.83E-01	-0.022	5.30E-01	8.80E-01
F3opd↔AG	-0.013	7.07E-01	9.27E-01	-0.013	7.17E-01	9.31E-01	-0.075	3.49E-02	3.65E-01	-0.024	5.02E-01	8.70E-01
STSp↔F3opdR	-0.074	3.91E-02	3.89E-01	-0.024	4.94E-01	8.69E-01	-0.011	7.60E-01	9.46E-01	-0.011	7.68E-01	9.49E-01
F3opd↔T2ml	-0.073	4.15E-02	3.89E-01	0.020	5.71E-01	8.83E-01	0.034	3.36E-01	7.65E-01	0.034	3.43E-01	7.70E-01
F3opd↔STSp	-0.073	4.02E-02	3.89E-01	-0.055	1.23E-01	5.47E-01	-0.022	5.38E-01	8.80E-01	-0.040	2.64E-01	6.96E-01
Pole↔PrF3op	-0.073	4.06E-02	3.89E-01	-0.025	4.88E-01	8.67E-01	-0.004	9.11E-01	9.74E-01	-0.025	4.82E-01	8.66E-01
PrF3op↔F3orb	-0.055	1.24E-01	5.47E-01	-0.025	4.86E-01	8.67E-01	-0.073	4.02E-02	3.89E-01	-0.021	5.52E-01	8.82E-01
Pole↔F3tv	0.006	8.68E-01	9.69E-01	0.073	4.13E-02	3.89E-01	0.017	6.25E-01	8.95E-01	0.010	7.78E-01	9.49E-01
F3orb↔T1R	-0.026	4.63E-01	8.55E-01	-0.063	7.78E-02	4.36E-01	-0.021	5.65E-01	8.83E-01	-0.072	4.28E-02	3.89E-01
RolS↔F3opdR	0.054	1.28E-01	5.49E-01	0.072	4.31E-02	3.89E-01	-0.020	5.66E-01	8.83E-01	0.014	6.93E-01	9.20E-01
T2ml↔F3orb	0.020	5.81E-01	8.83E-01	0.063	7.88E-02	4.40E-01	0.072	4.34E-02	3.89E-01	0.041	2.59E-01	6.95E-01
PT↔STSp	-0.072	4.32E-02	3.89E-01	0.021	5.65E-01	8.83E-01	-0.049	1.68E-01	6.00E-01	-0.016	6.53E-01	8.99E-01
T1↔T1R	-0.006	8.66E-01	9.69E-01	0.072	4.41E-02	3.89E-01	-0.026	4.64E-01	8.55E-01	0.005	8.86E-01	9.71E-01
T1a↔SMG	0.005	8.79E-01	9.71E-01	0.013	7.09E-01	9.27E-01	-0.072	4.40E-02	3.89E-01	-0.044	2.64E-01	6.49E-01
F3opd↔F3tv	-0.037	2.98E-01	7.28E-01	0.017	6.26E-01	8.95E-01	0.071	4.60E-02	4.02E-01	0.059	9.92E-02	4.85E-01
SMG↔T3p	0.036	3.12E-01	7.52E-01	0.063	7.57E-02	4.35E-01	0.071	4.66E-02	4.02E-01	0.061	8.74E-02	4.56E-01
PT↔T1a.HeschlR	0.064	7.37E-02	4.29E-01	0.031	3.89E-01	7.95E-01	0.067	6.20E-02	4.11E-01	0.071	4.66E-02	4.02E-01
Prec↔F2antR	-0.019	5.96E-01	8.89E-01	0.046	1.97E-01	6.28E-01	-0.069	5.27E-02	4.03E-01	-0.025	4.84E-01	8.66E-01
F3opd↔PrF3op	0.071	4.77E-02	4.03E-01	0.007	7.38E-01	9.62E-01	0.016	6.44E-01	8.98E-01	0.009	7.93E-01	9.49E-01
Pole↔SMG	-0.008	8.23E-01	9.54E-01	0.009	8.91E-01	9.49E-01	-0.068	5.74E-02	4.03E-01	-0.069	5.26E-02	4.03E-01
F3opdR↔T2pR	0.070	4.99E-02	4.03E-01	0.003	9.29E-01	9.77E-01	0.043	2.28E-01	6.64E-01	0.039	2.73E-01	7.09E-01
T2ml↔T1R	0.022	5.45E-01	8.80E-01	-0.007	8.51E-01	9.64E-01	-0.009	8.09E-01	9.49E-01	-0.068	5.59E-02	4.03E-01
PT↔PrecR	-0.047	1.91E-01	6.15E-01	-0.068	5.61E-02	4.03E-01	-0.041	2.54E-01	6.92E-01	0.000	9.90E-01	9.95E-01
F2p↔F3td	0.069	5.42E-02	4.03E-01	0.003	9.40E-01	9.77E-01	-0.033	3.50E-01	7.74E-01	-0.002	9.56E-01	9.79E-01
T2p↔F3orb	-0.054	1.27E-01	5.49E-01	-0.038	2.83E-01	7.12E-01	-0.069	5.38E-02	4.03E-01	-0.052	1.41E-01	5.71E-01
PrF3op↔STSp	-0.068	5.64E-02	4.03E-01	-0.027	4.45E-01	8.42E-01	-0.042	2.42E-01	6.82E-01	-0.020	5.81E-01	8.83E-01
Pole↔T1	-0.050	1.64E-01	5.94E-01	-0.068	5.54E-02	4.03E-01	-0.003	9.44E-01	9.77E-01	0.020	5.73E-01	8.83E-01
Pole↔AG	0.011	7.49E-01	9.45E-01	0.015	6.81E-01	9.13E-01	0.001	9.89E-01	9.95E-01	0.069	5.41E-02	4.03E-01
T2ml↔T2pR	0.014	6.96E-01	9.22E-01	0.023	5.12E-01	8.75E-01	0.040	2.63E-01	6.96E-01	0.069	5.40E-02	4.03E-01
PrF3op↔T1R	-0.020	5.71E-01	8.83E-01	-0.070	5.11E-02	4.03E-01	-0.067	6.19E-02	4.11E-01	-0.057	1.08E-01	5.12E-01
T1↔F3opdR	-0.068	5.53E-02	4.03E-01	0.013	7.09E-01	9.27E-01	-0.011	7.52E-01	9.46E-01	0.039	2.80E-01	7.12E-01
T2p↔T1a.HeschlR	-0.020	5.77E-01	8.83E-01	-0.069	5.23E-02	4.03E-01	-0.035	3.32E-01	7.63E-01	-0.064	7.08E-02	4.27E-01
T1a↔RolS	0.049	1.68E-01	6.00E-01	0.020	5.72E-01	8.83E-01	0.070	4.95E-02	4.03E-01	0.055	1.23E-01	5.47E-01
T2ml↔T2p	-0.017	6.30E-01	8.95E-01	-0.012	7.35E-01	9.39E-01	-0.069	5.38E-02	4.03E-01	-0.020	5.68E-01	8.83E-01
Fusa↔F3tv	-0.071	4.76E-02	4.03E-01	0.025	4.81E-01	8.66E-01	-0.061	8.64E-02	4.56E-01	-0.062	8.21E-02	4.45E-01
PrF3op↔F2antR	-0.066	6.37E-02	4.11E-01	-0.068	5.67E-02	4.03E-01	-0.047	1.90E-01	6.14E-01	-0.010	7.76E-01	9.49E-01
F3td↔F3opdR	-0.070	4.94E-02	4.03E-01	-0.025	4.87E-01	8.67E-01	0.034	3.41E-01	7.67E-01	0.002	9.48E-01	9.77E-01
Fusa↔T1a.HeschlR	-0.039	2.75E-01	7.09E-01	-0.070	5.01E-02	4.03E-01	-0.045	2.06E-01	6.31E-01	-0.035	3.27E-01	7.63E-01
F3orb↔RolS	-0.010	7.79E-01	9.49E-01	0.014	6.98E-01	9.22E-01	-0.070	4.99E-02	4.03E-01	-0.040	2.64E-01	6.96E-01
F3tv↔T2pR	-0.064	7.27E-02	4.29E-01	0.028	4.36E-01	8.33E-01	-0.068	5.68E-02	4.03E-01	0.015	6.64E-01	9.01E-01
T2p↔F3opdR	0.011	7.66E-01	9.49E-01	0.068	5.80E-02	4.03E-01	0.003	9.29E-01	9.77E-01	-0.008	8.30E-01	9.57E-01
STSp↔T3p	0.017	6.40E-01	8.97E-01	-0.018	6.22E-01	8.95E-01	-0.067	5.92E-02	4.08E-01	-0.021	5.64E-01	8.83E-01

Endophenotype	Language task overall accuracy			Working Memory			Oral reading recognition			Vocabulary Comprehension		
	r	p value	FDR p value	r	p value	FDR p value	r	p value	FDR p value	r	p value	FDR p value
AG↔T1a.HeschlR	0.006	8.76E-01	9.71E-01	0.048	1.82E-01	6.13E-01	0.067	5.96E-02	4.09E-01	0.050	1.63E-01	5.94E-01
T2ml↔PrF3op	0.001	9.70E-01	9.85E-01	-0.006	8.63E-01	9.69E-01	-0.067	6.00E-02	4.09E-01	-0.034	3.34E-01	7.64E-01
F3tv↔SMG	-0.030	4.00E-01	7.99E-01	-0.035	3.20E-01	7.57E-01	0.046	2.01E-01	6.31E-01	0.067	6.08E-02	4.11E-01
T1a↔F3orb	0.013	7.06E-01	9.27E-01	-0.066	6.28E-02	4.11E-01	0.042	2.39E-01	6.78E-01	0.045	2.09E-01	6.34E-01
F3tv↔T1R	0.066	6.22E-02	4.11E-01	-0.013	7.20E-01	9.31E-01	-0.002	9.58E-01	9.79E-01	-0.010	7.89E-01	9.49E-01
SMG↔RolS	0.011	7.60E-01	9.46E-01	0.010	7.86E-01	9.49E-01	-0.067	6.17E-02	4.11E-01	-0.064	7.16E-02	4.29E-01
T2ml↔AG	-0.066	6.32E-02	4.11E-01	0.028	4.27E-01	8.24E-01	-0.009	7.99E-01	9.49E-01	0.025	4.78E-01	8.64E-01
AG↔T2pR	-0.003	9.43E-01	9.77E-01	0.015	6.78E-01	9.11E-01	0.065	6.76E-02	4.23E-01	0.012	7.37E-01	9.39E-01
AG↔STSp	0.036	3.12E-01	7.52E-01	0.065	6.77E-02	4.23E-01	0.026	4.67E-01	8.55E-01	-0.013	7.15E-01	9.31E-01
F2p↔F3orb	0.065	6.92E-02	4.26E-01	0.048	1.80E-01	6.13E-01	0.045	2.05E-01	6.31E-01	0.030	4.05E-01	8.03E-01
Fusa↔T2pR	0.021	5.65E-01	8.83E-01	0.065	6.90E-02	4.26E-01	0.018	6.08E-01	8.95E-01	-0.015	6.72E-01	9.07E-01
T2p↔RolS	0.018	6.10E-01	8.95E-01	0.022	5.37E-01	8.80E-01	0.064	7.07E-02	4.27E-01	0.035	3.23E-01	7.58E-01
F3tv↔T1a.HeschlR	-0.041	2.48E-01	6.87E-01	-0.065	7.02E-02	4.27E-01	0.051	1.49E-01	5.75E-01	0.027	4.57E-01	8.52E-01
RolS↔T1R	0.004	9.10E-01	9.74E-01	0.064	7.33E-02	4.29E-01	0.048	1.75E-01	6.08E-01	0.063	7.74E-02	4.36E-01
T1↔RolS	0.064	7.37E-02	4.29E-01	0.061	8.90E-02	4.60E-01	-0.004	9.17E-01	9.74E-01	-0.019	5.94E-01	8.89E-01
PT↔T2pR	-0.017	6.26E-01	8.95E-01	-0.064	7.24E-02	4.29E-01	-0.002	9.47E-01	9.77E-01	-0.005	8.85E-01	9.71E-01
AG↔F2antR	0.063	7.54E-02	4.35E-01	-0.029	4.11E-01	8.08E-01	0.004	9.15E-01	9.74E-01	-0.024	4.94E-01	8.69E-01
F2p↔T2pR	0.052	1.43E-01	5.71E-01	-0.022	5.30E-01	8.80E-01	0.063	7.74E-02	4.36E-01	0.006	8.64E-01	9.69E-01
Fusa↔AG	-0.018	6.10E-01	8.95E-01	-0.063	7.73E-02	4.36E-01	0.033	3.61E-01	7.76E-01	-0.013	7.16E-01	9.31E-01
T3p↔T1R	0.040	2.58E-01	6.96E-01	0.027	4.48E-01	8.46E-01	-0.014	6.98E-01	9.22E-01	0.062	7.99E-02	4.44E-01
PT↔T2p	-0.047	1.89E-01	6.14E-01	-0.027	4.54E-01	8.48E-01	-0.062	8.19E-02	4.45E-01	-0.041	2.54E-01	6.92E-01
Pole↔T3p	-0.008	8.18E-01	9.51E-01	-0.009	8.02E-01	9.49E-01	-0.062	8.06E-02	4.45E-01	-0.018	6.11E-01	8.95E-01
F3opd↔Fusa	0.062	8.24E-02	4.45E-01	0.037	3.03E-01	7.36E-01	0.041	2.50E-01	6.89E-01	0.005	8.97E-01	9.72E-01
RolS↔PrecR	0.047	1.86E-01	6.13E-01	0.041	2.51E-01	6.91E-01	0.061	8.77E-02	4.56E-01	0.020	5.83E-01	8.83E-01
T2pR↔T1R	0.061	8.77E-02	4.56E-01	-0.035	3.28E-01	7.63E-01	-0.040	2.65E-01	6.96E-01	-0.056	1.14E-01	5.27E-01
PrF3op↔T1a.HeschlR	0.002	9.53E-01	9.79E-01	-0.017	6.41E-01	8.98E-01	-0.022	5.34E-01	8.80E-01	-0.061	8.96E-02	4.61E-01
F3opd↔T1a.HeschlR	0.018	6.19E-01	8.95E-01	0.060	9.06E-02	4.62E-01	-0.012	7.27E-01	9.35E-01	0.012	7.42E-01	9.41E-01
F2p↔SMG	0.003	9.35E-01	9.77E-01	-0.005	8.79E-01	9.71E-01	0.060	9.04E-02	4.62E-01	-0.022	5.38E-01	8.80E-01
PrF3op↔F3td	0.044	2.13E-01	6.40E-01	0.059	9.65E-02	4.78E-01	-0.047	1.89E-01	6.14E-01	-0.060	9.25E-02	4.70E-01
F3td↔T2pR	-0.010	7.89E-01	9.49E-01	0.017	6.42E-01	8.98E-01	-0.037	2.95E-01	7.28E-01	-0.059	9.67E-02	4.78E-01
T1↔Fusa	-0.010	7.74E-01	9.49E-01	0.059	9.57E-02	4.78E-01	-0.027	4.55E-01	8.49E-01	-0.024	4.99E-01	8.69E-01
STSp↔F3td	-0.059	9.57E-02	4.78E-01	-0.029	4.15E-01	8.12E-01	-0.022	5.40E-01	8.80E-01	-0.009	8.02E-01	9.49E-01
Pole↔F3orb	0.039	2.80E-01	7.12E-01	0.018	6.19E-01	8.95E-01	0.059	9.95E-02	4.85E-01	0.046	1.99E-01	6.28E-01
F2p↔T3p	0.022	5.35E-01	8.80E-01	-0.058	1.03E-01	5.00E-01	-0.012	7.30E-01	9.36E-01	-0.003	9.39E-01	9.77E-01
F2p↔PrecR	-0.058	1.05E-01	5.05E-01	-0.023	5.19E-01	8.77E-01	-0.012	7.29E-01	9.36E-01	0.013	7.16E-01	9.31E-01
F2p↔STSp	0.015	6.83E-01	9.13E-01	-0.005	8.88E-01	9.71E-01	0.006	8.72E-01	9.69E-01	0.058	1.07E-01	5.12E-01
F3tv↔T2p	0.011	7.49E-01	9.45E-01	-0.009	7.98E-01	9.49E-01	0.057	1.10E-01	5.24E-01	0.021	5.52E-01	8.82E-01
F2antR↔PrecR	-0.008	8.18E-01	9.51E-01	-0.052	1.46E-01	5.71E-01	-0.054	1.32E-01	5.61E-01	-0.057	1.13E-01	5.27E-01
SMG↔T2pR	-0.041	2.47E-01	6.86E-01	-0.056	1.15E-01	5.27E-01	-0.034	3.36E-01	7.65E-01	0.009	8.12E-01	9.49E-01
STSp↔SMG	0.056	1.14E-01	5.27E-01	0.023	5.11E-01	8.75E-01	0.008	8.23E-01	9.54E-01	0.013	7.18E-01	9.31E-01
T2ml↔F3opdR	-0.049	1.73E-01	6.08E-01	-0.042	2.37E-01	6.78E-01	-0.056	1.18E-01	5.39E-01	-0.015	6.83E-01	9.13E-01
F3opd↔F3td	0.055	1.26E-01	5.49E-01	0.020	5.67E-01	8.83E-01	0.056	1.19E-01	5.40E-01	0.045	2.10E-01	6.34E-01
STSp↔F3orb	0.005	8.92E-01	9.72E-01	-0.019	5.94E-01	8.89E-01	-0.055	1.21E-01	5.45E-01	-0.036	3.11E-01	7.52E-01
PT↔T1R	-0.037	2.96E-01	7.28E-01	0.055	1.23E-01	5.47E-01	0.032	3.72E-01	7.82E-01	0.052	1.45E-01	5.71E-01
STSp↔T1a.HeschlR	-0.011	7.55E-01	9.46E-01	-0.042	2.44E-01	6.83E-01	-0.035	3.32E-01	7.63E-01	-0.055	1.24E-01	5.47E-01
AG↔SMG	-0.011	7.59E-01	9.46E-01	-0.031	3.91E-01	7.95E-01	-0.021	5.65E-01	8.83E-01	-0.055	1.23E-01	5.47E-01
PT↔F3orb	-0.055	1.26E-01	5.49E-01	0.002	9.65E-01	9.82E-01	-0.021	5.58E-01	8.83E-01	0.017	6.39E-01	8.97E-01
PT↔T1	0.032	3.75E-01	7.82E-01	0.019	5.91E-01	8.87E-01	0.054	1.27E-01	5.49E-01	-0.006	8.56E-01	9.66E-01
Fusa↔T1a	0.054	1.27E-01	5.49E-01	0.007	8.43E-01	9.64E-01	0.031	3.91E-01	7.95E-01	0.004	9.01E-01	9.73E-01
F3orb↔PrecR	-0.002	9.60E-01	9.80E-01	-0.055	1.25E-01	5.49E-01	-0.008	8.30E-01	9.57E-01	0.007	8.41E-01	9.63E-01
PT↔F3opdR	-0.032	3.73E-01	7.82E-01	0.020	5.80E-01	8.83E-01	-0.002	9.45E-01	9.77E-01	-0.054	1.28E-01	5.50E-01
T2p↔F2antR	0.021	5.56E-01	8.83E-01	0.048	1.77E-01	6.10E-01	0.021	5.48E-01	8.80E-01	0.054	1.32E-01	5.61E-01
F2p↔Fusa	-0.053	1.37E-01	5.67E-01	-0.038	2.87E-01	7.16E-01	-0.054	1.33E-01	5.63E-01	-0.035	3.24E-01	7.58E-01
Pole↔F3opd	0.007	8.38E-01	9.62E-01	0.053	1.36E-01	5.67E-01	-0.005	8.91E-01	9.72E-01	0.012	7.31E-01	9.36E-01
F3td↔PrecR	0.053	1.37E-01	5.67E-01	0.031	3.84E-01	7.93E-01	-0.019	6.02E-01	8.93E-01	-0.033	3.51E-01	7.74E-01
F3tv↔T1a	0.053	1.36E-01	5.67E-01	0.040	2.67E-01	6.98E-01	0.009	8.11E-01	9.49E-01	0.004	9.18E-01	9.74E-01
T2ml↔Fusa	-0.047	1.87E-01	6.13E-01	-0.002	9.45E-01	9.77E-01	-0.036	3.17E-01	7.56E-01	0.053	1.37E-01	5.67E-01
T1a.HeschlR↔PrecR	0.033	3.61E-01	7.76E-01	0.022	5.45E-01	8.80E-01	0.052	1.46E-01	5.71E-01	0.022	5.47E-01	8.80E-01
F3td↔T1a.HeschlR	-0.026	4.69E-01	8.57E-01	-0.003	9.37E-01	9.77E-01	0.010	7.87E-01	9.49E-01	0.052	1.44E-01	5.71E-01
T2p↔T1R	-0.034	3.45E-01	7.73E-01	0.024	4.95E-01	8.69E-01	-0.052	1.42E-01	5.71E-01	-0.041	2.48E-01	6.87E-01
Pole↔T2p	0.020	5.81E-01	8.83E-01	0.052	1.46E-01	5.71E-01	0.002	9.47E-01	9.77E-01	0.016	6.60E-01	9.00E-01
T2p↔T3p	0.010	7.72E-01	9.49E-01	-0.052	1.45E-01	5.71E-01	-0.029	4.08E-01	8.04E-01	-0.004	9.01E-01	9.73E-01
F3tv↔F3orb	0.052	1.45E-01	5.71E-01	0.018	6.23E-01	8.95E-01	-0.016	6.58E-01	8.99E-01	-0.015	6.79E-01	9.12E-01
T1a↔T2p	-0.024	4.96E-01	8.69E-01	0.052	1.48E-01	5.73E-01	-0.002	9.55E-01	9.79E-01	-0.017	6.35E-01	8.97E-01
Prec↔STSp	0.008	8.27E-01	9.55E-01	-0.051	1.50E-01	5.76E-01	0.013	7.10E-01	9.29E-01	-0.006	8.67E-01	9.69E-01
T1a↔F2antR	0.005	8.83E-01	9.71E-01	0.051	1.54E-01	5.82E-01	0.012	7.39E-01	9.39E-01	0.032	3.64E-01	7.77E-01
PrF3op↔AG	0.051	1.55E-01	5.82E-01	0.035	3.20E-01	7.57E-01	-0.003	9.22E-01	9.76E-01	-0.021	5.51E-01	8.82E-01
T2pR↔T1a.HeschlR	-0.025	4.84E-01	8.66E-01	-0.023	5.17E-01	8.77E-01	0.051	1.54E-01	5.82E-01	0.023	5.11E-01	8.75E-01
SMG↔F2antR	-0.009	8.08E-01	9.49E-01	-0.034	3.40E-01	7.67E-01	-0.051	1.56E-01	5.82E-01	-0.018	6.09E-01	8.95E-01
T2ml↔T1	-0.025	4.79E-01	8.65E-01	-0.029	4.20E-01	8.19E-01	-0.051	1.54E-01	5.82E-01	-0.027	4.48E-01	8.46E-01

C.2. (Endo)phenotypic correlation between both resting-state functional connectivity and language-related scores

Endophenotype	Language task overall accuracy			Working Memory			Oral reading recognition			Vocabulary Comprehension		
	r	p value	FDR p value	r	p value	FDR p value	r	p value	FDR p value	r	p value	FDR p value
F3tv↔PrecR	0.051	1.55E-01	5.82E-01	-0.037	2.98E-01	7.28E-01	0.012	7.37E-01	9.39E-01	0.019	5.87E-01	8.85E-01
Prec↔T1a.HeschlR	0.012	7.38E-01	9.39E-01	0.005	8.95E-01	9.72E-01	-0.050	1.57E-01	5.84E-01	-0.006	8.69E-01	9.69E-01
F3orb↔F2antR	0.034	3.45E-01	7.73E-01	-0.050	1.59E-01	5.86E-01	0.045	2.05E-01	6.31E-01	-0.015	6.65E-01	9.01E-01
T1↔T2p	0.050	1.60E-01	5.87E-01	0.024	5.06E-01	8.73E-01	-0.017	6.36E-01	8.97E-01	-0.026	4.67E-01	8.55E-01
PT↔Fusa	-0.004	9.18E-01	9.74E-01	-0.030	3.97E-01	7.98E-01	-0.050	1.63E-01	5.94E-01	-0.048	1.78E-01	6.12E-01
Pole↔Fusa	-0.024	5.02E-01	8.70E-01	-0.006	8.60E-01	9.69E-01	-0.050	1.64E-01	5.94E-01	-0.034	3.38E-01	7.65E-01
Prec↔T2ml	-0.036	3.13E-01	7.52E-01	-0.049	1.66E-01	5.97E-01	-0.021	5.65E-01	8.83E-01	-0.016	6.54E-01	8.99E-01
PT↔T3p	-0.011	7.48E-01	9.45E-01	0.016	6.61E-01	9.00E-01	0.007	8.46E-01	9.64E-01	-0.049	1.67E-01	6.00E-01
Fusa↔T2p	-0.015	6.64E-01	9.01E-01	0.045	2.06E-01	6.31E-01	-0.044	2.22E-01	6.54E-01	-0.049	1.72E-01	6.08E-01
Prec↔Fusa	-0.009	7.93E-01	9.49E-01	0.025	4.81E-01	8.66E-01	0.018	6.14E-01	8.95E-01	0.049	1.74E-01	6.08E-01
STSp↔T1a	0.048	1.75E-01	6.88E-01	-0.022	5.42E-01	8.80E-01	-0.007	8.35E-01	9.62E-01	0.011	7.55E-01	9.46E-01
F2p↔T2p	-0.048	1.75E-01	6.88E-01	-0.000	9.99E-01	9.99E-01	-0.026	4.61E-01	8.55E-01	0.040	2.62E-01	6.66E-01
T2p↔PrecR	-0.047	1.84E-01	6.13E-01	-0.016	6.44E-01	8.98E-01	-0.016	6.49E-01	8.99E-01	-0.022	5.37E-01	8.80E-01
PrF3op↔F3tv	-0.016	6.52E-01	8.99E-01	0.004	9.17E-01	9.74E-01	0.026	4.74E-01	8.60E-01	0.047	1.86E-01	6.13E-01
Fusa↔PrecR	-0.011	7.67E-01	9.49E-01	-0.047	1.83E-01	6.13E-01	-0.008	8.18E-01	9.51E-01	-0.017	6.25E-01	8.95E-01
T2ml↔SMG	-0.019	5.87E-01	8.85E-01	0.030	3.93E-01	7.95E-01	-0.032	3.74E-01	7.82E-01	-0.047	1.84E-01	6.13E-01
Prec↔F3td	0.004	9.04E-01	9.74E-01	-0.003	9.32E-01	9.77E-01	0.009	8.02E-01	9.49E-01	0.048	1.83E-01	6.13E-01
Pole↔F2antR	0.014	6.92E-01	9.20E-01	0.048	1.81E-01	6.13E-01	0.035	3.21E-01	7.57E-01	0.008	8.23E-01	9.54E-01
F3orb↔T1a.HeschlR	0.004	9.03E-01	9.74E-01	-0.020	5.79E-01	8.83E-01	-0.047	1.83E-01	6.13E-01	-0.030	4.01E-01	7.99E-01
AG↔T1a	0.032	3.75E-01	7.82E-01	0.047	1.89E-01	6.14E-01	0.019	5.90E-01	8.87E-01	0.045	2.07E-01	6.31E-01
F2p↔F3opdR	0.033	3.59E-01	7.76E-01	0.003	9.33E-01	9.77E-01	0.047	1.91E-01	6.15E-01	-0.016	6.58E-01	8.99E-01
STSp↔F2antR	0.009	8.03E-01	9.49E-01	0.046	1.92E-01	6.17E-01	0.046	2.01E-01	6.31E-01	0.009	7.95E-01	9.49E-01
F2p↔PrF3op	0.000	9.95E-01	9.98E-01	0.046	1.94E-01	6.20E-01	-0.014	7.01E-01	9.24E-01	0.010	7.72E-01	9.49E-01
PT↔F3opd	-0.016	6.53E-01	8.99E-01	-0.029	4.09E-01	8.04E-01	-0.046	1.95E-01	6.22E-01	-0.039	2.80E-01	7.12E-01
STSp↔T1R	-0.045	2.03E-01	6.31E-01	0.016	6.52E-01	8.99E-01	-0.046	1.98E-01	6.28E-01	-0.040	2.58E-01	6.66E-01
SMG↔F3td	0.046	1.99E-01	6.28E-01	-0.017	6.39E-01	8.97E-01	0.032	3.70E-01	7.82E-01	-0.016	6.59E-01	8.99E-01
Pole↔F2p	-0.045	2.05E-01	6.31E-01	-0.013	7.20E-01	9.31E-01	-0.010	7.85E-01	9.49E-01	-0.002	9.58E-01	9.79E-01
Pole↔T1R	0.045	2.06E-01	6.31E-01	-0.005	8.85E-01	9.71E-01	0.040	2.65E-01	6.96E-01	0.022	5.40E-01	8.80E-01
PrF3op↔SMG	0.021	5.63E-01	8.83E-01	0.045	2.08E-01	6.33E-01	-0.010	7.76E-01	9.49E-01	-0.023	5.24E-01	8.80E-01
F3opd↔T1R	-0.033	3.55E-01	7.76E-01	0.033	3.61E-01	7.76E-01	0.041	2.50E-01	6.89E-01	0.045	2.09E-01	6.34E-01
Pole↔STSp	-0.015	6.72E-01	9.07E-01	-0.045	2.10E-01	6.34E-01	-0.022	5.30E-01	8.80E-01	0.007	8.51E-01	9.64E-01
Fusa↔PrF3op	0.044	2.15E-01	6.43E-01	0.002	9.66E-01	9.82E-01	0.042	2.37E-01	6.78E-01	0.018	6.23E-01	8.95E-01
PT↔AG	0.003	9.24E-01	9.76E-01	0.006	8.61E-01	9.69E-01	0.033	3.61E-01	7.76E-01	0.044	2.21E-01	6.53E-01
Prec↔F3opd	0.043	2.23E-01	6.54E-01	0.008	8.13E-01	9.49E-01	-0.024	4.98E-01	8.69E-01	-0.021	5.62E-01	8.83E-01
F3td↔T1R	-0.002	9.53E-01	9.79E-01	0.020	5.68E-01	8.83E-01	0.020	5.72E-01	8.83E-01	0.043	2.30E-01	6.67E-01
T2pR↔PrecR	-0.042	2.37E-01	6.78E-01	0.023	5.14E-01	8.77E-01	-0.029	4.13E-01	8.10E-01	-0.005	8.87E-01	9.71E-01
T1a↔T2pR	0.042	2.39E-01	6.78E-01	0.038	2.86E-01	7.15E-01	0.003	9.40E-01	9.77E-01	0.002	9.55E-01	9.79E-01
PrF3op↔PrecR	0.022	5.33E-01	8.80E-01	-0.003	9.42E-01	9.77E-01	0.031	3.86E-01	7.94E-01	-0.042	2.40E-01	6.78E-01
F2antR↔T1a.HeschlR	-0.033	3.57E-01	7.76E-01	-0.003	9.24E-01	9.76E-01	-0.026	4.65E-01	8.55E-01	-0.042	2.40E-01	6.78E-01
PT↔F2p	-0.017	6.34E-01	8.97E-01	0.022	5.44E-01	8.80E-01	0.042	2.43E-01	6.82E-01	0.003	9.34E-01	9.77E-01
T3p↔PrecR	-0.041	2.45E-01	6.85E-01	-0.022	5.45E-01	8.80E-01	-0.021	5.48E-01	8.80E-01	-0.007	8.41E-01	9.63E-01
RoL↔T1a.HeschlR	0.041	2.45E-01	6.85E-01	-0.032	3.69E-01	7.81E-01	0.016	6.49E-01	8.99E-01	0.014	6.66E-01	9.22E-01
Prec↔PT	0.041	2.52E-01	6.91E-01	0.030	3.93E-01	7.95E-01	-0.017	6.30E-01	8.95E-01	-0.010	7.75E-01	9.49E-01
F3tv↔F2antR	-0.007	8.52E-01	9.64E-01	0.040	2.63E-01	6.96E-01	0.032	3.75E-01	7.82E-01	0.027	4.53E-01	8.88E-01
SMG↔T1R	0.040	2.57E-01	6.96E-01	0.004	9.09E-01	9.74E-01	-0.006	8.68E-01	9.69E-01	0.022	5.40E-01	8.80E-01
F3tv↔F3td	-0.040	2.63E-01	6.96E-01	-0.015	6.83E-01	9.13E-01	-0.003	9.41E-01	9.77E-01	-0.021	5.49E-01	8.80E-01
F3orb↔T2pR	-0.004	9.16E-01	9.74E-01	0.039	2.75E-01	7.09E-01	0.033	3.55E-01	7.76E-01	-0.003	9.39E-01	9.77E-01
Pole↔T2pR	0.039	2.74E-01	7.09E-01	0.024	5.03E-01	8.70E-01	0.011	7.57E-01	9.46E-01	0.020	5.82E-01	8.83E-01
Prec↔F3tv	0.039	2.77E-01	7.09E-01	0.032	3.68E-01	7.81E-01	-0.029	4.19E-01	8.17E-01	-0.030	4.00E-01	7.99E-01
Prec↔T1R	-0.028	4.30E-01	8.26E-01	0.039	2.77E-01	7.09E-01	-0.017	6.27E-01	8.95E-01	-0.030	4.05E-01	8.03E-01
F3tv↔RoL	0.031	3.90E-01	7.95E-01	-0.016	6.52E-01	8.99E-01	-0.038	2.83E-01	7.12E-01	-0.018	6.08E-01	8.95E-01
Prec↔T2p	-0.021	5.54E-01	8.83E-01	-0.038	2.83E-01	7.12E-01	-0.005	8.87E-01	9.71E-01	-0.000	9.96E-01	9.98E-01
STSp↔T2pR	0.038	2.83E-01	7.12E-01	0.031	3.85E-01	7.94E-01	0.012	7.47E-01	9.45E-01	-0.002	9.57E-01	9.79E-01
T1↔T2pR	0.025	4.90E-01	8.68E-01	0.010	7.82E-01	9.49E-01	0.005	8.97E-01	9.72E-01	0.038	2.86E-01	7.15E-01
T2ml↔T3p	0.027	4.52E-01	8.48E-01	0.025	4.90E-01	8.68E-01	0.033	3.50E-01	7.74E-01	0.038	2.92E-01	7.25E-01
Prec↔Pole	-0.003	9.35E-01	9.77E-01	-0.009	7.93E-01	9.49E-01	0.018	6.11E-01	8.95E-01	0.037	2.94E-01	7.27E-01
T3p↔RoL	-0.037	2.96E-01	7.28E-01	-0.028	4.24E-01	8.20E-01	-0.035	3.29E-01	7.63E-01	-0.011	7.53E-01	9.46E-01
F3tv↔T3p	-0.024	5.07E-01	8.73E-01	-0.037	2.98E-01	7.28E-01	-0.029	4.23E-01	8.20E-01	0.030	4.08E-01	8.04E-01
PrF3op↔T1a	0.005	8.83E-01	9.71E-01	0.001	9.77E-01	9.90E-01	-0.037	3.00E-01	7.32E-01	0.003	9.35E-01	9.77E-01
AG↔RoL	-0.036	3.10E-01	7.52E-01	0.004	9.08E-01	9.74E-01	-0.009	8.00E-01	9.49E-01	0.016	6.44E-01	8.98E-01
F3opd↔RoL	-0.016	6.51E-01	8.99E-01	0.036	3.15E-01	7.54E-01	0.018	6.21E-01	8.95E-01	-0.019	6.01E-01	8.93E-01
F3opd↔T2p	0.022	5.41E-01	8.80E-01	0.033	3.50E-01	7.74E-01	-0.004	9.01E-01	9.73E-01	0.035	3.22E-01	7.57E-01
T1↔SMG	-0.035	3.31E-01	7.63E-01	-0.026	4.69E-01	8.57E-01	0.034	3.36E-01	7.65E-01	0.025	4.84E-01	8.66E-01
F3orb↔F3opdR	-0.016	6.57E-01	8.99E-01	0.001	9.84E-01	9.93E-01	-0.017	6.33E-01	8.97E-01	-0.034	3.38E-01	7.65E-01
T2ml↔RoL	0.008	8.22E-01	9.54E-01	0.034	3.38E-01	7.65E-01	0.004	9.08E-01	9.74E-01	-0.019	5.88E-01	8.85E-01
F3opd↔T1	-0.031	3.81E-01	7.88E-01	0.032	3.65E-01	7.77E-01	0.018	6.23E-01	8.95E-01	0.025	4.87E-01	8.67E-01
PT↔SMG	-0.028	4.30E-01	8.26E-01	0.032	3.65E-01	7.77E-01	-0.026	4.71E-01	8.58E-01	0.010	7.88E-01	9.49E-01
STSp↔RoL	-0.017	6.29E-01	8.95E-01	0.020	5.79E-01	8.83E-01	-0.005	8.81E-01	9.71E-01	-0.032	3.65E-01	7.77E-01
T1↔F3td	-0.032	3.67E-01	7.80E-01	0.019	5.99E-01	8.92E-01	-0.012	7.26E-01	9.35E-01	0.005	8.90E-01	9.72E-01

Endophenotype	Language task overall accuracy			Working Memory			Oral reading recognition			Vocabulary Comprehension		
	r	p value	FDR p value	r	p value	FDR p value	r	p value	FDR p value	r	p value	FDR p value
F3td↔F2antR	-0.011	7.61E-01	9.46E-01	0.002	9.48E-01	9.77E-01	-0.022	5.33E-01	8.80E-01	-0.032	3.69E-01	7.81E-01
F2p↔T1a.HeschlR	0.020	5.81E-01	8.83E-01	-0.031	3.77E-01	7.85E-01	0.010	7.82E-01	9.49E-01	0.002	9.63E-01	9.81E-01
T3p↔T2pR	-0.019	5.92E-01	8.87E-01	-0.031	3.80E-01	7.88E-01	0.008	8.13E-01	9.49E-01	-0.009	8.06E-01	9.49E-01
Fusa↔T1R	-0.010	7.79E-01	9.49E-01	0.003	9.27E-01	9.77E-01	-0.031	3.86E-01	7.94E-01	-0.030	3.94E-01	7.95E-01
Pole↔RolS	-0.028	4.40E-01	8.38E-01	-0.006	8.59E-01	9.69E-01	-0.030	3.94E-01	7.95E-01	-0.014	6.86E-01	9.14E-01
AG↔T3p	-0.030	3.94E-01	7.95E-01	-0.018	6.08E-01	8.95E-01	-0.009	8.06E-01	9.49E-01	-0.007	8.51E-01	9.64E-01
F2p↔F2antR	-0.000	9.97E-01	9.98E-01	-0.024	5.09E-01	8.74E-01	-0.007	8.46E-01	9.64E-01	-0.030	3.99E-01	7.99E-01
AG↔PrecR	-0.004	9.00E-01	9.73E-01	0.024	4.93E-01	8.69E-01	0.029	4.09E-01	8.04E-01	-0.020	5.74E-01	8.83E-01
T3p↔F3opdR	-0.000	9.91E-01	9.96E-01	0.001	9.85E-01	9.93E-01	-0.029	4.22E-01	8.19E-01	-0.010	7.82E-01	9.49E-01
F2p↔T1R	-0.009	7.93E-01	9.49E-01	0.001	9.77E-01	9.90E-01	-0.005	8.92E-01	9.72E-01	-0.029	4.22E-01	8.19E-01
Pole↔T1a	0.014	7.03E-01	9.25E-01	-0.001	9.87E-01	9.93E-01	0.029	4.23E-01	8.19E-01	0.014	6.91E-01	9.20E-01
F3opdR↔PrecR	-0.009	8.01E-01	9.49E-01	0.019	6.03E-01	8.94E-01	0.028	4.30E-01	8.26E-01	-0.002	9.62E-01	9.81E-01
Prec↔T1	0.009	8.11E-01	9.49E-01	0.027	4.41E-01	8.39E-01	0.024	4.96E-01	8.69E-01	-0.026	4.59E-01	8.54E-01
Prec↔T1a	0.018	6.20E-01	8.95E-01	0.024	5.04E-01	8.72E-01	0.027	4.44E-01	8.42E-01	0.022	5.32E-01	8.80E-01
PT↔T2ml	-0.015	6.66E-01	9.02E-01	-0.027	4.54E-01	8.48E-01	0.006	8.73E-01	9.69E-01	0.022	5.46E-01	8.80E-01
Prec↔F3opdR	0.010	7.79E-01	9.49E-01	-0.019	5.96E-01	8.89E-01	0.026	4.65E-01	8.55E-01	0.006	8.69E-01	9.69E-01
T1a↔T3p	-0.003	9.24E-01	9.76E-01	0.023	5.17E-01	8.77E-01	0.008	8.16E-01	9.51E-01	0.025	4.76E-01	8.62E-01
F3opd↔F3opdR	0.024	4.93E-01	8.69E-01	0.004	9.07E-01	9.74E-01	-0.000	9.97E-01	9.98E-01	-0.012	7.32E-01	9.37E-01
Pole↔PT	-0.005	8.95E-01	9.72E-01	0.013	7.17E-01	9.31E-01	-0.023	5.19E-01	8.77E-01	0.003	9.44E-01	9.77E-01
Prec↔RolS	-0.018	6.14E-01	8.95E-01	-0.023	5.19E-01	8.77E-01	0.016	6.58E-01	8.99E-01	0.016	6.63E-01	9.01E-01
PrF3op↔T2pR	0.016	6.49E-01	8.99E-01	0.023	5.18E-01	8.77E-01	-0.022	5.38E-01	8.80E-01	-0.009	8.09E-01	9.49E-01
SMG↔T2p	0.017	6.37E-01	8.97E-01	-0.004	9.15E-01	9.74E-01	0.003	9.35E-01	9.77E-01	-0.022	5.32E-01	8.80E-01
T1R↔PrecR	-0.019	5.88E-01	8.85E-01	-0.019	5.96E-01	8.89E-01	-0.007	8.46E-01	9.64E-01	-0.022	5.42E-01	8.80E-01
AG↔T2p	-0.004	9.14E-01	9.74E-01	-0.020	5.69E-01	8.83E-01	0.021	5.64E-01	8.83E-01	0.007	8.37E-01	9.62E-01
F2antR↔T2pR	0.006	8.62E-01	9.69E-01	-0.021	5.65E-01	8.83E-01	0.011	7.55E-01	9.46E-01	-0.005	8.98E-01	9.72E-01
T1↔PrF3op	0.017	6.35E-01	8.97E-01	-0.018	6.13E-01	8.95E-01	0.008	8.27E-01	9.55E-01	0.003	9.23E-01	9.76E-01
STSp↔PrecR	0.011	7.67E-01	9.49E-01	-0.012	7.38E-01	9.39E-01	-0.018	6.16E-01	8.95E-01	-0.001	9.79E-01	9.90E-01
Prec↔T3p	0.017	6.32E-01	8.97E-01	0.013	7.20E-01	9.31E-01	0.005	8.80E-01	9.71E-01	-0.009	8.03E-01	9.49E-01
F3td↔T3p	0.015	6.84E-01	9.14E-01	-0.007	8.50E-01	9.64E-01	0.013	7.25E-01	9.34E-01	-0.016	6.52E-01	8.99E-01
T2p↔F3td	-0.016	6.57E-01	8.99E-01	0.011	7.54E-01	9.46E-01	0.008	8.25E-01	9.55E-01	0.013	7.20E-01	9.31E-01
Prec↔AG	-0.012	7.31E-01	9.36E-01	-0.010	7.72E-01	9.49E-01	-0.005	8.83E-01	9.71E-01	0.014	7.01E-01	9.24E-01
Fusa↔T3p	-0.007	8.47E-01	9.64E-01	0.007	8.53E-01	9.64E-01	-0.001	9.75E-01	9.89E-01	0.005	8.98E-01	9.72E-01

C.3 (Endo)phenotypic correlation between both task activations and language-related scores

Table C.3 – Pearson Correlation between both task activations and language-related scores. Significant correlation was observed after false discovery rate correction at $p < 0.05$.

Endophenotype	Language task overall accuracy			Working Memory			Oral reading recognition			Vocabulary Comprehension		
	r	p value	FDR p value	r	p value	FDR p value	r	p value	FDR p value	r	p value	FDR p value
Pole	0.217	7.56E-10	3.78E-08	0.148	3.13E-05	4.15E-04	0.176	6.28E-07	1.57E-05	0.239	1.10E-11	1.10E-09
STSp	0.130	2.47E-04	1.65E-03	0.133	1.87E-04	1.34E-03	0.108	2.38E-03	1.13E-02	0.183	2.31E-07	7.70E-06
T1a	0.175	8.10E-07	1.62E-05	0.142	6.29E-05	5.57E-04	0.101	4.61E-03	2.01E-02	0.138	9.82E-05	7.55E-04
T1R	0.095	7.85E-03	3.14E-02	0.147	3.32E-05	4.15E-04	0.165	3.05E-06	5.08E-05	0.105	3.07E-03	1.39E-02
T2ml	0.142	6.68E-05	5.57E-04	0.145	4.62E-05	5.13E-04	0.142	6.68E-05	5.57E-04	0.126	3.87E-04	2.42E-03
T1a.HeschlR	0.119	8.22E-04	4.83E-03	0.093	9.37E-03	3.47E-02	0.085	1.64E-02	5.42E-02	0.046	1.95E-01	3.82E-01
PT	0.058	1.03E-01	2.58E-01	0.115	1.17E-03	6.48E-03	0.091	1.06E-02	3.79E-02	0.055	1.20E-01	2.83E-01
SMG	-0.114	1.37E-03	7.19E-03	-0.020	5.77E-01	6.76E-01	-0.030	4.00E-01	5.27E-01	-0.048	1.81E-01	3.61E-01
T1	0.093	8.65E-03	3.33E-02	0.113	1.49E-03	7.43E-03	0.077	3.03E-02	8.92E-02	0.040	2.68E-01	4.79E-01
AG	0.030	3.99E-01	5.27E-01	0.082	2.11E-02	6.61E-02	0.055	1.22E-01	2.83E-01	0.100	5.09E-03	2.12E-02
F2antR	-0.087	1.51E-02	5.20E-02	-0.027	4.52E-01	5.58E-01	-0.019	5.89E-01	6.77E-01	-0.073	4.01E-02	1.11E-01
Fusa	0.028	4.25E-01	5.45E-01	0.048	1.74E-01	3.61E-01	0.085	1.68E-02	5.42E-02	0.063	7.92E-02	2.08E-01
F3orb	0.010	7.84E-01	8.43E-01	0.077	3.01E-02	8.92E-02	0.038	2.81E-01	4.94E-01	0.036	3.19E-01	5.02E-01
F3opd	-0.026	4.66E-01	5.69E-01	0.035	3.29E-01	5.06E-01	0.075	3.49E-02	9.98E-02	0.032	3.70E-01	5.27E-01
T2pR	-0.065	6.80E-02	1.84E-01	0.030	4.00E-01	5.27E-01	0.061	8.77E-02	2.25E-01	-0.016	6.58E-01	7.39E-01
F2p	0.045	2.10E-01	4.03E-01	0.056	1.18E-01	2.83E-01	0.031	3.79E-01	5.27E-01	0.053	1.36E-01	3.10E-01
T2p	-0.052	1.46E-01	3.24E-01	0.050	1.57E-01	3.41E-01	-0.020	5.81E-01	6.76E-01	0.004	9.02E-01	9.20E-01
F3td	-0.048	1.78E-01	3.61E-01	0.030	4.08E-01	5.30E-01	0.031	3.89E-01	5.27E-01	0.037	3.03E-01	5.02E-01
T3p	0.007	8.35E-01	8.79E-01	0.049	1.71E-01	3.61E-01	0.035	3.21E-01	5.02E-01	0.016	6.54E-01	7.39E-01
RolS	-0.043	2.25E-01	4.25E-01	0.043	2.31E-01	4.28E-01	0.025	4.85E-01	5.85E-01	-0.011	7.54E-01	8.19E-01
Prec	-0.009	7.97E-01	8.48E-01	0.040	2.66E-01	4.79E-01	0.038	2.91E-01	4.96E-01	0.005	8.98E-01	9.20E-01
F3opdR	-0.036	3.19E-01	5.02E-01	-0.034	3.39E-01	5.13E-01	0.028	4.40E-01	5.58E-01	-0.038	2.92E-01	4.96E-01
PrF3op	-0.002	9.53E-01	9.63E-01	0.036	3.09E-01	5.02E-01	0.031	3.79E-01	5.27E-01	-0.006	8.66E-01	9.03E-01
F3tv	-0.020	5.73E-01	6.76E-01	0.034	3.44E-01	5.13E-01	-0.000	9.90E-01	9.90E-01	-0.014	6.92E-01	7.61E-01
PrecR	-0.032	3.64E-01	5.27E-01	0.027	4.48E-01	5.58E-01	-0.030	3.94E-01	5.27E-01	-0.014	6.88E-01	7.61E-01

APPENDIX **D**

D.1 Genomic loci associated with heritable language anatomical connectivities using MOSTest

Table D.1 – Genomic loci associated with heritable language anatomical connectivities using MOSTest. Lead SNP: ID of the lead SNPs within each locus. Position: position of the SNP in the hg19 human reference genome. mvgwasP discovery -British-: MOSTest association P value obtained using the discovery sample. mvgwasP replication -non British-: MOSTest association P value obtained using the independent replication sample. Functional category: Functional consequence of the SNP on the gene obtained from ANNOVAR. 'Central' phenotypes: the phenotypes that contributed most to the multivariate association considering the genome-wide association threshold ($5e - 8$).

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mvgwasP (discovery -British-)	mvgwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
6	rs6688111	1	47980916	intergenic	G	T	0.360	2.01E-56	4.27E-02	AL365458.1	MO Fronto Insular tracts5 Right	6.73E-16	-8.075324
8	rs61788401	1	59373734	nCRNA intronic	A	T	0.109	3.63E-14	2.67E-01	LINC01135	L3 Superior Longitudinal Fasciculus II Left	3.63E-08	5.507745
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Fronto Insular tract4 Right	2.97E-15	-7.891989
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Superior Longitudinal Fasciculus II Right	7.30E-12	-6.851627
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Arcuate Anterior Segment Right	7.20E-20	-9.124649
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Fronto Insular tract3 Right	2.23E-14	-7.636839
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	MD Arcuate Anterior Segment Right	3.25E-08	5.527584
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Arcuate Long Segment Left	3.78E-20	-9.194018
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Fronto Insular tract3 Right	2.27E-13	-7.331844
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Inferior Fronto Occipital fasciculus Right	4.40E-08	-5.473817
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	FA Arcuate Long Segment Left	1.07E-08	-5.718978
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Corpus callosum	9.13E-09	-5.746152
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Arcuate Posterior Segment Right	1.14E-14	-7.722946
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Fronto Insular tract3 Left	6.26E-12	-6.873663
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	MD Superior Longitudinal Fasciculus III Left	1.56E-08	5.655382
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Inferior Fronto Occipital fasciculus Left	1.02E-10	-6.463870
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Frontal Aslant Tract Right	4.30E-15	-7.845731
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Frontal Inferior longitudinal Left	1.59E-08	-5.651815
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Superior Longitudinal Fasciculus II Left	9.41E-11	-6.67183
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	L3 Arcuate Long Segment Left	2.25E-08	5.591641
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Frontal Aslant Tract Left	2.36E-17	-8.744477
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Superior Longitudinal Fasciculus III Right	4.44E-17	-8.401499
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Arcuate Posterior Segment Left	1.40E-18	-8.797344
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Fronto Insular tract5 Left	3.02E-13	-7.293440
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Arcuate Long Segment Right	1.38E-13	-7.398005
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Inferior Longitudinal Right	8.61E-09	-5.756002
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Fronto Insular tract2 Left	1.25E-09	-6.073372
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	MD Arcuate Posterior Segment Left	4.36E-08	5.475330
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Arcuate Anterior Segment Left	5.75E-14	-7.513694
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Superior Longitudinal Fasciculus I Right	4.45E-09	-5.866297
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	L2 Arcuate Long Segment Left	3.25E-11	6.634737
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Superior Longitudinal Fasciculus I Left	8.36E-11	-6.493977
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ISOVF Fronto Insular tract3 Right	2.53E-08	-5.741105
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	MD Arcuate Long Segment Left	2.06E-10	6.359044
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Uncinate Right	1.09E-12	-7.118951
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Uncinate Left	4.15E-12	-6.931883
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Inferior Longitudinal Left	6.95E-10	-6.167145
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Superior Longitudinal Fasciculus III Left	4.35E-19	-8.927577
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ISOVF Fronto Insular tract4 Right	8.08E-09	-5.766822
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	L3 Arcuate Anterior Segment Right	3.64E-08	5.507350
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	L2 Arcuate Posterior Segment Left	2.30E-08	5.879950
10	rs2780818	1	65307994	intronic	G	T	0.223	3.43E-37	6.04E-04	JAK1	ICVF Frontal Inferior longitudinal Right	1.59E-16	-8.248888
12	rs3767967	1	94002233	intronic	G	A	0.391	2.00E-25	1.71E-02	FNBPAL	OD Superior Longitudinal Fasciculus II Right	1.57E-11	-5.654032
12	rs3767967	1	94002233	intronic	G	A	0.391	2.00E-25	1.71E-02	FNBPAL	MO Superior Longitudinal Fasciculus I Left	2.44E-08	-5.494500
12	rs3767967	1	94002233	intronic	G	A	0.391	2.00E-25	1.71E-02	FNBPAL	OD Superior Longitudinal Fasciculus II Left	9.73E-09	-5.735333
15	rs3841064	1	153919858	upstream:downstream	AC	A	0.491	1.14E-12	4.69E-02	CRTC2	ICVF Arcuate Posterior Segment Right	5.62E-09	5.827759
15	rs3841064	1	153919858	upstream:downstream	AC	A	0.491	1.14E-12	4.69E-02	CRTC2	ICVF Arcuate Posterior Segment Left	3.06E-10	6.295846
15	rs3841064	1	153919858	upstream:downstream	AC	A	0.491	1.14E-12	4.69E-02	CRTC2	ICVF Arcuate Anterior Segment Left	9.69E-09	5.736030
15	rs3841064	1	153919858	upstream:downstream	AC	A	0.491	1.14E-12	4.69E-02	CRTC2	ICVF Superior Longitudinal Fasciculus III Left	3.22E-08	5.529218
16	rs4745	1	155106227	exonic	T	A	0.479	1.20E-20	1.60E-01	EFNA1	OD Fronto Insular tract3 Right	1.82E-08	5.628755
19	rs35306826	1	180956936	intronic	A	T	0.408	6.65E-16	3.00E-03	STX6	ICVF Uncinate Right	2.07E-09	-5.992424
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	FA Inferior Fronto Occipital fasciculus Right	4.81E-08	-5.458275
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	FA Uncinate Right	2.04E-08	-5.608183
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	FA Inferior Fronto Occipital fasciculus Left	3.92E-08	-5.494500
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	FA Uncinate Left	5.57E-09	-5.829170
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	FA Arcuate Anterior Segment Right	1.55E-09	-6.039372
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	FA Fronto Insular tract4 Right	1.31E-10	-6.425766
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	FA Arcuate Long Segment Left	1.32E-10	-6.425243
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	MO Inferior Fronto Occipital fasciculus Right	2.80E-10	-6.309694
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	OD Arcuate Long Segment Left	3.43E-13	7.276474
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	OD Inferior Fronto Occipital fasciculus Left	5.87E-09	5.820496
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	FA Arcuate Long Segment Right	1.93E-10	-6.367021
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	OD Superior Longitudinal Fasciculus III Right	5.92E-09	5.818894
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	FA Fronto Insular tract3 Right	3.68E-11	-6.616568
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	OD Arcuate Long Segment Right	5.15E-13	7.221302
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	OD Arcuate Anterior Segment Right	2.32E-11	6.684432
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	OD Fronto Insular tract4 Right	9.50E-09	5.739337
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	OD Inferior Fronto Occipital fasciculus Right	4.51E-11	6.686259
20	rs10922200	1	197321440	intronic	T	C	0.110	2.17E-22	1.33E-01	CRB1	OD Fronto Insular tract3 Right	3.92E-13	7.258305
24	rs1452628	1	215139887	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Frontal Aslant Tract Left	1.16E-15	8.008297
24	rs1452628	1	215139887	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L1 Arcuate Anterior Segment Right	2.59E-08	5.657129
24	rs1452628	1	215139887	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Inferior Fronto Occipital fasciculus Left	1.31E-09	6.066640
24	rs1452628	1	215139887	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Superior Longitudinal Fasciculus II Right	2.43E-14	7.625774
24	rs1452628	1	215139887	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	MO Frontal Aslant Tract Left	3.53E-08	5.512863
24	rs1452628	1	215139887	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	MD Arcuate Anterior Segment Right	1.44E-09	5.050899
24	rs1452628	1	215139887	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L1 Frontal Superior Longitudinal Right	8.31E-14	7.465263
24	rs1452628	1	215139887	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Frontal Inferior longitudinal Right	1.21E-15	8.003749
24	rs1452628	1	215139887	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	MD Frontal Aslant Tract Right	2.98E-10	6.299814

D.1. Genomic loci associated with heritable language anatomical connectivities using MOSTest

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mgwasP (discovery - British)	mgwasP (replication - non British)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Inferior Longitudinal Right	1.01E-13	7.439071
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Superior Longitudinal Fasciculus III Left	3.00E-17	8.440710
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	MO Frontal Aslant Tract Left	1.45E-10	6.743338
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Arcuate Anterior Segment Left	6.40E-17	8.337530
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L3 Frontal Aslant Tract Left	3.08E-08	5.536681
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	OD Corpus callosum	1.39E-08	5.675166
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Arcuate Posterior Segment Left	1.60E-10	6.910532
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L1 Arcuate Anterior Segment Left	8.00E-10	6.133483
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L3 Superior Longitudinal Fasciculus III Right	7.96E-11	6.501320
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Frontal Superior Longitudinal Right	8.57E-12	6.620799
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Inferior Longitudinal Left	3.94E-14	7.584427
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Frontal Aslant Tract Right	4.80E-15	7.832073
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	OD Superior Longitudinal Fasciculus I Left	2.06E-10	6.356845
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	MD Superior Longitudinal Fasciculus III Left	1.60E-12	7.054531
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	OD Superior Longitudinal Fasciculus I Right	1.67E-13	7.732076
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Arcuate Long Segment Left	4.61E-12	6.916993
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L1 Arcuate Posterior Segment Left	9.64E-10	6.112520
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Superior Longitudinal Fasciculus II Left	4.54E-08	5.468561
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L2 Superior Longitudinal Fasciculus III Right	3.47E-10	6.276107
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L3 Superior Longitudinal Fasciculus III Left	2.01E-10	6.360294
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Superior Longitudinal Fasciculus I Left	1.42E-12	7.082466
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	MO Frontal Aslant Tract Right	4.87E-08	5.466051
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Superior Longitudinal Fasciculus I Right	8.46E-10	6.136307
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L2 Superior Longitudinal Fasciculus III Left	2.75E-08	5.559294
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	MD Frontal Aslant Tract Left	1.15E-11	6.789115
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L1 Superior Longitudinal Fasciculus II Right	4.54E-08	5.468561
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	MD Arcuate Anterior Segment Left	1.26E-11	6.773146
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	MD Frontal Inferior Longitudinal Left	1.04E-08	5.732423
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L2 Superior Longitudinal Fasciculus III Right	1.53E-11	6.740881
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L1 Frontal Aslant Tract Left	5.24E-18	8.646500
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	MD Frontal Inferior Longitudinal Right	4.11E-10	6.249785
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	MD Arcuate Posterior Segment Left	1.06E-08	5.720891
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L1 Frontal inferior longitudinal Right	1.73E-15	7.955953
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L1 Superior Longitudinal Fasciculus I Right	8.39E-16	8.048485
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Corpus callosum	3.11E-17	8.442507
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L3 Arcuate Anterior Segment Left	1.23E-08	5.650193
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Frontal Inferior Longitudinal Right	4.27E-11	6.914875
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Arcuate Anterior Segment Right	1.07E-12	7.121270
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L2 Arcuate Anterior Segment Left	2.34E-09	5.972572
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L1 Superior Longitudinal Fasciculus III Left	9.47E-15	7.750299
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L1 Frontal Aslant Tract Right	1.47E-12	6.813452
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Superior Longitudinal Fasciculus III Right	1.87E-17	8.501368
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L1 Corpus callosum	1.08E-12	7.194466
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Arcuate Long Segment Right	4.37E-12	6.947497
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	ISOVF Arcuate Posterior Segment Right	2.75E-11	6.914875
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	MD Superior Longitudinal Fasciculus III Right	2.29E-12	7.015655
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L3 Frontal Inferior Longitudinal Right	4.93E-08	5.453957
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	MD Superior Longitudinal Fasciculus I Right	1.99E-11	6.795175
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L3 Arcuate Anterior Segment Right	1.33E-08	5.682588
24	rs1452628	1	21513987	intergenic	T	A	0.385	4.97E-54	9.96E-03	RP11-323K10.1	L1 Superior Longitudinal Fasciculus I Left	9.47E-15	7.746150
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	ISOVF Inferior Longitudinal Right	2.15E-08	5.598845
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	ISOVF Superior Longitudinal Fasciculus III Left	4.27E-11	6.914875
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	ISOVF Arcuate Anterior Segment Left	1.09E-12	7.118477
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	ISOVF Arcuate Posterior Segment Left	6.76E-11	6.525994
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	L3 Superior Longitudinal Fasciculus III Right	2.23E-08	5.593048
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	ISOVF Inferior Longitudinal Left	1.07E-12	6.914875
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	MD Superior Longitudinal Fasciculus III Right	3.63E-09	5.990043
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	ISOVF Arcuate Long Segment Left	8.14E-09	5.765637
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	MD Arcuate Anterior Segment Left	3.64E-09	5.990051
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	L2 Superior Longitudinal Fasciculus III Right	6.24E-11	6.914875
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	L1 Frontal Aslant Tract Left	1.56E-10	6.399594
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	L1 Arcuate Anterior Segment Left	4.99E-08	5.456883
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	L2 Superior Longitudinal Fasciculus III Left	1.99E-11	6.795175
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	L1 Frontal Aslant Tract Right	2.78E-09	6.843948
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	ISOVF Superior Longitudinal Fasciculus III Right	7.56E-09	5.777199
24	rs2601645	1	21518552	intronic	T	C	0.084	2.12E-21	6.34E-02	CKNK2	MD Superior Longitudinal Fasciculus III Right	6.24E-09	5.869922
25	rs1138753	2	27148654	intronic	T	C	0.384	1.04E-42	1.99E-03	DPYS15	MD Superior Longitudinal Fasciculus I Left	4.44E-12	6.259521
25	rs1138753	2	27148654	intronic	T	C	0.384	1.04E-42	1.99E-03	DPYS15	MD Superior Longitudinal Fasciculus I Left	6.45E-11	6.533907
25	rs1138753	2	27148654	intronic	T	C	0.384	1.04E-42	1.99E-03	DPYS15	OD Superior Longitudinal Fasciculus II Left	6.83E-10	6.170080
26	rs3505623	2	37063240	intergenic	CT	C	0.462	7.04E-37	2.08E-01	AC007382.1	OD Frontal Superior Longitudinal Right	4.34E-08	5.478232
26	rs3505623	2	37063240	intergenic	CT	C	0.462	7.04E-37	2.08E-01	AC007382.1	ISOVF Superior Longitudinal Fasciculus II Left	2.44E-08	5.716485
26	rs11891807	2	48288816	ncRNA intronic	G	T	0.385	2.93E-16	5.22E-02	AC07807.4	MO Frontal Insular tract Right	2.43E-08	5.578464
29	2.56116193	TTGATCTTGGCCCTCCCTCC	2	56116									

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mgwasp (discovery -British-)	mgwasp (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L1 Inferior Longitudinal Left	1.42E-10	-6.414158
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Superior Longitudinal Fasciculus II Right	3.84E-12	-6.943855
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Arcuate Anterior Segment Right	3.75E-10	-6.264164
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Frontal Insular tracts Right	1.48E-10	-6.407570
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	MD Arcuate Anterior Segment Right	5.12E-09	-5.843240
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L2 Inferior Longitudinal Left	5.94E-14	-7.509462
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	MD Inferior Longitudinal Right	5.96E-12	-6.880593
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L1 Inferior Fronto Occipital fasciculus Left	1.05E-09	-6.101993
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Arcuate Long Segment Left	4.42E-10	-6.238554
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	MD Frontal Aslant Tract Right	5.20E-09	-5.840553
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L3 Superior Longitudinal Fasciculus I Left	5.08E-13	-7.223186
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Fronto Insular tracts Right	5.25E-11	-6.65652
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Fronto Insular tracts Right	1.35E-09	-6.060788
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	FA Superior Longitudinal Fasciculus I Left	2.35E-09	-5.971554
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L3 Frontal Aslant Tract Left	2.40E-09	-5.968352
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Inferior Fronto Occipital fasciculus Right	9.99E-11	-6.468648
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	MD Uncinate Right	6.25E-12	-6.873877
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	MD Inferior Fronto Occipital fasciculus Right	8.72E-13	-7.149285
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Corpus callosum	2.25E-11	-6.680113
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L3 Superior Longitudinal Fasciculus III Right	1.79E-09	-6.016055
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Arcuate Posterior Segment Right	6.25E-09	-5.809952
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	MD Corpus callosum	5.30E-12	-6.897253
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L2 Inferior Fronto Occipital fasciculus Right	6.57E-10	-6.176125
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L1 Arcuate Long Segment Right	3.81E-08	-5.499589
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L3 Arcuate Posterior Segment Right	4.01E-08	-5.490439
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Fronto Insular tracts Left	2.69E-09	-5.946929
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	MD Superior Longitudinal Fasciculus III Left	1.35E-10	-6.421772
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Inferior Fronto Occipital fasciculus Left	3.46E-12	-6.957512
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L1 Superior Longitudinal Fasciculus II Left	2.71E-09	-5.948219
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Frontal Aslant Tract Right	2.19E-11	-7.016661
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L1 Inferior Fronto Occipital fasciculus Right	9.22E-11	-6.479284
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Frontal Superior Longitudinal Left	2.74E-08	-5.559099
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L1 Arcuate Posterior Segment Left	4.00E-13	-7.255522
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	MD Fronto Insular tracts Left	1.15E-10	-6.445209
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	MD Arcuate Posterior Segment Right	3.28E-12	-6.965255
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	FA Superior Longitudinal Fasciculus II Right	2.02E-08	-5.609972
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L2 Uncinate Right	4.23E-08	-5.480942
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L2 Corpus callosum	5.73E-10	-6.197547
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L1 Uncinate Left	1.03E-12	-7.125875
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L2 Superior Longitudinal Fasciculus III Right	7.51E-12	-6.847624
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L3 Superior Longitudinal Fasciculus III Left	6.73E-09	-5.797503
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L1 Uncinate Right	2.78E-11	-6.657625
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Superior Longitudinal Fasciculus II Left	1.89E-11	-6.714385
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L3 Inferior Fronto Occipital fasciculus Right	3.78E-10	-6.262762
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L2 Frontal Aslant Tract Left	2.95E-11	-6.649022
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L2 Superior Longitudinal Fasciculus III Left	1.15E-10	-6.445393
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L3 Inferior Fronto Occipital fasciculus Left	4.56E-09	-5.862525
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L3 Arcuate Long Segment Left	1.50E-08	-5.661758
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Frontal Aslant Tract Left	3.11E-14	-7.593450
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	MD Frontal Aslant Tract Left	4.84E-11	-6.575894
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L1 Superior Longitudinal Fasciculus II Right	1.65E-10	-6.399763
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Superior Longitudinal Fasciculus III Right	2.84E-13	-7.301674
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L2 Superior Longitudinal Fasciculus II Right	1.11E-12	-7.116329
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Frontal Superior Longitudinal Right	1.34E-10	-6.422650
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	MD Arcuate Anterior Segment Left	1.52E-09	-6.041865
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L3 Arcuate Posterior Segment Left	1.77E-11	-6.724194
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L1 Superior Longitudinal Fasciculus III Right	5.46E-11	-6.436177
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	MD Superior Longitudinal Fasciculus II Left	5.73E-13	-7.206804
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L1 Frontal Aslant Tract Left	1.94E-08	-5.616839
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L2 Inferior Fronto Occipital fasciculus Left	1.10E-10	-6.451918
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Arcuate Posterior Segment Left	2.04E-12	-7.031511
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L2 Uncinate Left	4.26E-11	-6.594731
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Arcuate Long Segment Right	3.52E-08	-5.513313
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L3 Inferior Longitudinal Left	1.03E-09	-6.104600
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L3 Corpus callosum	7.20E-09	-5.786267
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L3 Uncinate Left	1.34E-10	-6.422321
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Inferior Longitudinal Right	2.30E-10	-6.339963
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Fronto Insular tracts Left	1.18E-12	-7.108204
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L2 Superior Longitudinal Fasciculus I Right	1.74E-12	-7.053910
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L2 Superior Longitudinal Fasciculus I Left	1.54E-13	-7.403267
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L1 Inferior Longitudinal Right	5.46E-11	-6.527708
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	MD Arcuate Posterior Segment Left	2.44E-16	-8.193103
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Arcuate Anterior Segment Left	1.38E-10	-6.418407
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Superior Longitudinal Fasciculus I Right	1.28E-15	-7.996247
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	MD Superior Longitudinal Fasciculus I Right	1.69E-13	-7.371610
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	MD Arcuate Long Segment Right	2.14E-10	-6.350957
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L3 Uncinate Right	1.27E-10	-6.431300
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Fronto Insular tracts Right	7.18E-13	-7.176040
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L2 Arcuate Long Segment Left	3.16E-08	-5.532600
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	ICVF Superior Longitudinal Fasciculus I Left	4.68E-15	-7.835254
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L2 Frontal Aslant Tract Right	1.73E-09	-6.020910
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	FA Superior Longitudinal Fasciculus I Right	1.27E-08	-5.690551
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A	L2 Inferior Longitudinal Right	7.60E-11	-6.508347
49	rs6442409	3	13835087	intergenic	C	T	0.343	1.20E-25	2.94E-01	WNT7A</			

D.1. Genomic loci associated with heritable language connectivities using MOSTest

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mgvwasP (discovery -British-)	mgvwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
72	r3170987940	ACTCT A	3	170887049	intronic	A	ACTCT	0.455	2.71E-14	7.93E-02	TNIK	9.33E-09	5.742494
72	r3170987940	ACTCT A	3	170987040	intronic	A	ACTCT	0.455	2.71E-14	7.93E-02	TNIK	4.13E-08	5.485005
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	1.48E-09	-6.046966
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	1.46E-08	-6.665789
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	5.12E-10	-6.215248
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	3.91E-08	-5.494767
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	6.12E-10	-6.187278
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	8.37E-11	-6.493747
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	1.31E-09	-5.924591
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	1.07E-08	-5.749064
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	2.16E-08	-5.598444
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	2.64E-09	-5.952408
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	1.18E-09	-6.083558
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	1.05E-09	-6.101431
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	4.87E-09	-5.851703
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	3.96E-10	-6.255507
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	2.24E-08	-5.592626
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	1.29E-09	-6.068114
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	1.51E-09	-6.043451
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	1.31E-08	-5.533811
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	9.09E-09	-5.745821
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	4.85E-08	-5.456874
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	1.67E-10	-6.389048
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	1.00E-08	-5.730239
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	6.54E-10	-6.176943
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	2.81E-08	-5.553056
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	1.59E-09	-6.034471
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	3.74E-09	-5.895467
73	rs9647379		3	171785168	intronic	C	G	0.410	5.56E-12	2.28E-01	FNDC3B	5.35E-09	-5.835794
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	2.19E-08	-5.591566
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	1.40E-10	-6.415792
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	1.99E-10	-6.362051
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	1.02E-08	-5.726704
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	4.43E-08	-5.472603
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	6.82E-10	-6.170280
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	1.66E-11	-6.733381
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	1.02E-08	-5.498978
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	8.00E-10	-6.144987
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	5.56E-09	-5.616819
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	4.13E-08	-5.483329
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	4.51E-11	-6.586364
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	1.42E-09	-6.052866
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	1.68E-10	-6.388323
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	2.78E-08	-5.554889
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	6.04E-09	-5.815685
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	6.93E-09	-5.792699
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	8.88E-10	-6.128355
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	5.53E-10	-6.203224
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	2.46E-09	-5.627717
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	3.32E-09	-5.333359
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	3.27E-09	-5.917310
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	8.02E-10	-6.144571
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	1.70E-08	-5.640440
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	2.77E-09	-5.944915
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	1.28E-08	-5.688848
75	rs3772795		3	183388403	intronic	G	A	0.255	2.84E-13	6.43E-01	KLHL24	2.46E-09	-5.964393
77	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	3.99E-08	-5.491078
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	1.35E-08	-5.679882
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	1.19E-10	-6.440017
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	2.73E-08	-5.576798
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	5.45E-17	-8.375854
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	1.06E-21	-9.570958
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	3.17E-10	-6.290227
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	4.11E-09	-5.879709
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	5.09E-09	-5.841886
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	5.01E-19	-8.912072
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	4.51E-08	-5.469481
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	4.27E-13	-7.320707
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	1.74E-08	-5.679318
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	8.24E-11	-6.327717
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	6.55E-09	-5.802008
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	6.37E-20	-9.137876
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	6.31E-09	-5.808250
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	1.07E-08	-5.719064
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	3.19E-08	-5.530655
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	6.23E-14	-7.503160
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	1.02E-08	-5.469481
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	3.33E-18	-8.699588
85	rs6822050		4	53943333	intronic	A	G	0.273	6.88E-12	2.35E-01	SCFD2	5.70E-10	-6.198491
85	rs6822050		4	53943333									

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mgwasP (discovery -British-)	mgwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Frontal Aslant Tract Left	8.22E-15	-7.764066
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Inferior Fronto Occipital fasciculus Left	5.78E-19	-8.896224
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Superior Longitudinal Fasciculus II Right	9.20E-16	-8.035886
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	L1 Inferior Longitudinal Left	5.33E-10	-6.209028
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Frontal Inferior longitudinal Right	6.02E-11	-6.524259
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Inferior Longitudinal Right	4.66E-16	-8.120021
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Superior Longitudinal Fasciculus III Left	1.24E-16	-8.279063
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Arcuate Anterior Segment Left	7.53E-13	-7.69408
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Arcuate Posterior Segment Left	3.25E-13	-7.283955
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Uncinate Left	4.55E-08	-5.468247
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Frontal Superior Longitudinal Right	2.35E-08	-5.584337
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Inferior Fronto Occipital fasciculus Right	3.24E-17	-8.437857
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Inferior Longitudinal Left	4.13E-18	-8.675029
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Frontal Aslant Tract Right	1.90E-19	-9.018783
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	L1 Inferior Fronto Occipital fasciculus Right	1.61E-08	-5.649119
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Arcuate Long Segment Left	5.35E-19	-8.904707
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	MD Arcuate Posterior Segment Right	2.37E-08	-5.582730
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Superior Longitudinal Fasciculus II Left	1.56E-15	-7.971976
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Superior Longitudinal Fasciculus I Left	1.43E-12	-7.081019
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Superior Longitudinal Fasciculus I Right	7.86E-14	-7.472725
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	L1 Superior Longitudinal Fasciculus I Right	2.75E-10	-6.312001
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Corpus callosum	4.64E-19	-8.020552
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	MD Superior Longitudinal Fasciculus I Right	1.41E-08	-5.671943
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Arcuate Anterior Segment Right	7.61E-20	-9.181536
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	L1 Arcuate Posterior Segment Right	1.25E-09	-6.073356
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Superior Longitudinal Fasciculus III Right	7.53E-22	-9.606612
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	L1 Corpus callosum	1.96E-12	-7.937331
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Arcuate Long Segment Right	6.50E-24	-10.083995
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	ISOVF Arcuate Posterior Segment Right	9.36E-21	-9.343050
96	rs1882325	4	185126789	intronic	T	G	0.350	7.55E-17	1.30E-01	ENPP6	L1 Superior Longitudinal Fasciculus I Left	1.44E-09	-6.053075
100	rs34592828	5	75995909	exonic	A	G	0.045	6.32E-10	4.43E-01	IQGAP2;CTD-2384B11.2	L3 Fronto Insular tracts Left	4.57E-08	5.467468
101	rs145824047	5	81800313	intergenic	A	G	0.396	4.50E-10	8.02E-01	CTD-2015A6.1	L2 Superior Longitudinal Fasciculus I Left	2.38E-09	5.969787
101	rs145824047	5	81800313	intergenic	A	G	0.396	4.50E-10	8.02E-01	CTD-2015A6.1	ICVF Superior Longitudinal Fasciculus I Right	2.65E-08	-5.563281
101	rs752021310	5	81866968	intergenic	C	CTAAT	0.447	3.20E-23	8.87E-03	CTD-2015A6.2	L2 Superior Longitudinal Fasciculus I Left	4.93E-09	-5.495052
101	rs752021310	5	81866968	intergenic	C	CTAAT	0.447	3.20E-23	8.87E-03	CTD-2015A6.2	OD Superior Longitudinal Fasciculus II Left	1.78E-09	-6.066256
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Superior Longitudinal Fasciculus II Left	1.66E-23	-9.975102
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Inferior Longitudinal Left	6.81E-27	-10.737232
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Fronto Insular tracts Right	1.06E-19	9.082184
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Frontal Aslant Tract Right	8.37E-17	-8.325912
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Inferior Fronto Occipital fasciculus Right	2.56E-18	8.738640
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Arcuate Anterior Segment Right	2.42E-13	-7.324442
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Superior Longitudinal Fasciculus III Right	1.28E-17	8.545063
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Superior Longitudinal Fasciculus II Right	4.83E-20	-9.167744
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Uncinate Right	9.19E-12	6.886881
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Frontal Inferior longitudinal Right	5.28E-12	-6.897765
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Inferior Longitudinal Left	6.09E-17	-8.363389
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Superior Longitudinal Fasciculus II Right	3.21E-43	13.783430
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Inferior Longitudinal Left	2.89E-19	8.972950
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Inferior Fronto Occipital fasciculus Left	2.56E-18	8.738640
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Frontal Aslant Tract Left	1.03E-14	7.737323
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Arcuate Anterior Segment Right	4.01E-43	13.767252
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Uncinate Left	4.05E-13	7.253748
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Fronto Insular tracts Right	9.78E-17	8.307416
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Arcuate Anterior Segment Right	1.07E-18	-8.27836
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Inferior Longitudinal Left	1.09E-25	-10.424755
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Inferior Longitudinal Right	1.07E-28	-11.113815
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Inferior Fronto Occipital fasciculus Left	9.57E-20	-9.093740
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Arcuate Long Segment Left	1.71E-44	13.993418
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Arcuate Posterior Segment Right	3.48E-17	8.149308
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Superior Longitudinal Fasciculus I Left	1.44E-23	-10.005001
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Fronto Insular tracts Right	1.73E-16	8.239398
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Fronto Insular tracts Left	3.45E-16	8.149803
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Arcuate Anterior Segment Right	3.08E-12	6.937879
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Superior Longitudinal Fasciculus I Left	5.51E-21	9.308976
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Frontal Aslant Tract Left	2.63E-15	-7.907178
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Inferior Fronto Occipital fasciculus Right	2.74E-45	14.123123
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Arcuate Anterior Segment Left	1.68E-13	-7.371795
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Arcuate Long Segment Left	3.62E-11	6.618973
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Uncinate Right	8.52E-21	-9.353032
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Inferior Fronto Occipital fasciculus Right	8.71E-30	-11.335804
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Corpus callosum	1.80E-48	14.631010
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Superior Longitudinal Fasciculus III Right	2.70E-19	-8.980335
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Arcuate Posterior Segment Right	6.08E-38	12.86784
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Frontal Inferior longitudinal Right	5.11E-10	6.215768
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Corpus callosum	3.20E-25	-10.375750
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Inferior Fronto Occipital fasciculus Right	1.10E-23	-10.032299
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Arcuate Long Segment Right	1.80E-16	-8.230978
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Arcuate Posterior Segment Right	2.24E-18	-8.744212
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Fronto Insular tracts Left	1.98E-17	8.404888
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Superior Longitudinal Fasciculus III Left	2.53E-13	7.317192
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Superior Longitudinal Fasciculus III Left	5.10E-20	-9.161866
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Inferior Fronto Occipital fasciculus Left	4.38E-46	14.251594
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Superior Longitudinal Fasciculus II Left	1.04E-10	-6.365670

D.1. Genomic loci associated with heritable language anatomical connectivities using MOSTest

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mgwasP (discovery -British-)	mgwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Arcuate Anterior Segment Left	7.21E-11	6.516128
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Arcuate Anterior Segment Right	3.49E-10	-6.275274
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Inferior Fronto Occipital Fasciculus Left	9.52E-28	-10.917384
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Arcuate Long Segment Left	4.44E-23	-9.893455
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Arcuate Long Segment Left	7.09E-19	-8.873440
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Frontal Aslant Tract Left	4.04E-40	13.258209
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Frontal Aslant Tract Left	1.27E-16	-8.275975
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Superior Longitudinal Fasciculus II Right	9.00E-11	-6.482926
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Superior Longitudinal Fasciculus III Right	1.77E-42	13.659461
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Superior Longitudinal Fasciculus II Right	1.07E-18	-8.282012
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Frontal Superior Longitudinal Right	2.19E-35	-12.244077
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Arcuate Anterior Segment Left	1.68E-17	-8.514318
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Arcuate Posterior Segment Left	3.13E-14	-7.592898
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Superior Longitudinal Fasciculus III Right	1.14E-14	-7.722874
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Superior Longitudinal Fasciculus II Left	3.25E-22	-9.692444
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Frontal Aslant Tract Right	1.44E-17	8.531560
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Frontal Aslant Tract Left	4.55E-10	-6.233906
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Inferior Fronto Occipital Fasciculus Left	1.72E-23	-9.988109
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Arcuate Posterior Segment Left	9.39E-37	12.603724
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Uncinate Left	1.40E-17	-8.534760
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Arcuate Long Segment Right	2.54E-10	6.324950
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Fronto Insular tracts Left	1.76E-21	9.518091
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Arcuate Long Segment Right	1.12E-41	13.524638
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Inferior Longitudinal Left	6.22E-25	-10.319860
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Corpus callosum	9.70E-24	-10.044036
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Uncinate Left	5.77E-20	-9.148559
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Inferior Longitudinal Right	1.08E-40	14.349357
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Arcuate Long Segment Right	1.52E-21	-9.533713
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Arcuate Posterior Segment Left	1.48E-08	5.664306
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Fronto Insular tracts Left	4.38E-12	6.924223
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Superior Longitudinal Fasciculus I Right	5.05E-24	-10.108881
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Corpus callosum	3.92E-22	9.673302
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Superior Longitudinal Fasciculus I Left	8.83E-23	-8.824556
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	OD Arcuate Long Segment Right	1.70E-09	6.024277
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Frontal Inferior longitudinal Right	1.90E-09	6.006486
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Inferior Longitudinal Right	2.28E-18	-8.742265
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Arcuate Posterior Segment Left	3.22E-21	-9.455477
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Arcuate Anterior Segment Left	5.34E-39	13.063247
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Arcuate Anterior Segment Left	2.64E-13	-7.311471
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Superior Longitudinal Fasciculus I Right	5.11E-10	14.870634
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Arcuate Posterior Segment Right	1.77E-09	6.017361
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Inferior Longitudinal Right	2.97E-21	9.463958
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Superior Longitudinal Fasciculus I Right	1.14E-21	-9.562940
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Arcuate Long Segment Right	2.36E-24	-10.183992
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Uncinate Right	2.11E-22	-9.736150
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Fronto Insular tracts Right	3.87E-22	9.674549
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Arcuate Long Segment Left	8.95E-11	-6.483729
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Superior Longitudinal Fasciculus I Left	1.09E-50	14.936710
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Frontal Aslant Tract Right	6.09E-20	-9.142730
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Superior Longitudinal Fasciculus I Right	2.46E-25	-10.400744
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Arcuate Long Segment Right	1.61E-13	-7.377843
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	FA Superior Longitudinal Fasciculus II Left	1.74E-19	9.028315
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Inferior Longitudinal Right	1.11E-27	-10.903403
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Arcuate Posterior Segment Right	1.70E-18	-8.775296
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Arcuate Anterior Segment Left	1.55E-13	-7.383214
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Superior Longitudinal Fasciculus III Left	6.40E-17	-8.357934
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Superior Longitudinal Fasciculus II Right	3.74E-21	-9.439771
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Frontal Aslant Tract Right	1.23E-08	-5.695006
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Corpus callosum	9.53E-19	-8.840494
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Arcuate Long Segment Left	6.36E-25	-10.390780
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Uncinate Right	5.67E-34	12.151007
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Superior Longitudinal Fasciculus I Right	1.06E-25	-10.480289
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Uncinate Left	5.57E-35	12.339237
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Inferior Longitudinal Left	3.03E-46	14.277367
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Superior Longitudinal Fasciculus III Left	5.59E-40	13.233906
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Inferior Longitudinal Right	5.12E-28	-10.973555
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Uncinate Left	1.97E-18	-8.758834
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Frontal Superior Longitudinal Right	2.20E-09	-5.982425
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Arcuate Posterior Segment Right	9.99E-20	-9.089033
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Superior Longitudinal Fasciculus II Left	6.26E-21	-9.385589
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Superior Longitudinal Fasciculus III Right	9.65E-19	-8.839089
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Frontal Inferior longitudinal Right	3.96E-08	-5.492809
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Inferior Fronto Occipital Fasciculus Left	1.39E-29	-11.294734
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Frontal Superior Longitudinal Right	3.86E-08	-5.497014
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L3 Arcuate Anterior Segment Right	9.60E-20	-9.002389
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L2 Arcuate Posterior Segment Left	1.83E-16	-8.232991
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	ICVF Frontal Inferior longitudinal Right	2.05E-30	11.461794
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	MD Superior Longitudinal Fasciculus I Left	2.18E-22	-9.733031
102	rs11749904	5	82782530	intronic	C	T	0.380	6.09E-29	4.43E-02	VCAN	L1 Superior Longitudinal Fasciculus I Left	2.14E-08	-5.599967
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Superior Longitudinal Fasciculus II Left	5.44E-09	5.831315
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Inferior Longitudinal Left	6.57E-14	7.496263

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mgwasP (discovery -British-)	mgwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Fronto Insular tract4 Right	3.64E-32	-11.805896
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Frontal Aslant Tract Right	3.32E-11	6.631745
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Fronto Insular tract5 Right	3.33E-11	6.962911
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	FA Inferior Fronto Occipital fasciculus Right	1.37E-11	-6.760879
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	FA Superior Longitudinal Fasciculus III Right	5.92E-11	-6.545660
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Superior Longitudinal Fasciculus II Right	2.00E-09	5.998050
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	FA Uncinate Right	4.18E-10	-6.247190
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L1 Inferior Longitudinal Left	3.41E-10	6.278886
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Superior Longitudinal Fasciculus II Right	1.12E-23	-10.030157
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Fronto Insular tract2 Right	4.37E-08	5.475121
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	FA Inferior Longitudinal Left	6.47E-09	-5.804067
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	FA Inferior Fronto Occipital fasciculus Left	6.05E-11	-6.542565
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	FA Frontal Aslant Tract Left	8.54E-11	-6.490825
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Arcuate Anterior Segment Right	5.51E-25	-10.323584
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	FA Uncinate Left	7.25E-16	-8.066150
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Fronto Insular tract3 Right	1.94E-29	-11.265423
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Arcuate Anterior Segment Right	9.75E-12	6.810189
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Inferior Longitudinal Left	3.20E-12	6.968825
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Inferior Longitudinal Right	6.96E-13	7.180310
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L1 Inferior Fronto Occipital fasciculus Left	3.68E-13	7.266737
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Arcuate Long Segment Left	1.73E-24	-10.213161
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Frontal Aslant Tract Right	2.51E-12	7.002939
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ISOVF Inferior Longitudinal Right	6.05E-09	-5.815480
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Superior Longitudinal Fasciculus I Left	3.88E-13	7.259339
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L1 Fronto Insular tract4 Right	1.90E-09	6.005892
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Fronto Insular tract2 Right	9.65E-22	-9.580601
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Fronto Insular tract5 Right	3.38E-24	-10.148181
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	FA Arcuate Anterior Segment Right	2.98E-08	-5.542577
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Fronto Insular tract5 Right	2.11E-10	6.353067
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	FA Superior Longitudinal Fasciculus I Left	5.35E-13	-7.216068
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Fronto Insular Tract Left	1.87E-11	6.715763
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	FA Fronto Insular tract4 Right	2.65E-08	-5.563005
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Inferior Fronto Occipital fasciculus Right	3.77E-25	-10.359923
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L1 Arcuate Anterior Segment Left	1.43E-11	6.754910
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	FA Arcuate Long Segment Left	3.68E-09	-5.898117
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Uncinate Right	2.35E-20	9.245159
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Fronto Insular tract3 Right	2.24E-09	5.979323
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Inferior Fronto Occipital fasciculus Right	9.32E-14	7.450224
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Corpus callosum	4.33E-30	-11.399621
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Superior Longitudinal Fasciculus III Right	1.98E-13	7.305119
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Arcuate Posterior Segment Right	5.77E-23	-9.867319
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Corpus callosum	2.57E-15	7.910342
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Fronto Insular tract3 Right	1.20E-10	6.439869
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Inferior Fronto Occipital fasciculus Right	4.88E-12	6.908957
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Fronto Insular tract4 Right	1.41E-14	7.695473
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Arcuate Posterior Segment Right	3.98E-08	5.494571
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Fronto Insular tract3 Left	1.06E-23	-10.035687
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	FA Superior Longitudinal Fasciculus III Left	4.73E-12	-6.913438
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Superior Longitudinal Fasciculus III Left	3.70E-16	8.148023
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Inferior Fronto Occipital fasciculus Left	9.73E-30	-11.326238
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Fronto Insular tract4 Right	6.36E-28	-10.953925
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Frontal Superior Longitudinal Left	3.24E-17	-8.437509
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Fronto Insular tract5 Left	4.89E-09	5.850868
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Arcuate Posterior Segment Right	2.72E-10	6.313945
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Uncinate Right	1.28E-10	8.275555
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Corpus callosum	2.05E-15	7.960340
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Fronto Insular tract5 Left	7.21E-10	6.164482
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L1 Uncinate Left	4.80E-12	6.911254
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Superior Longitudinal Fasciculus III Right	2.85E-12	6.985002
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Superior Longitudinal Fasciculus III Left	3.41E-15	7.749877
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L1 Uncinate Right	1.57E-14	7.681192
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Frontal inferior longitudinal Left	3.69E-14	-7.571394
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Superior Longitudinal Fasciculus III Left	7.24E-24	-10.073487
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Inferior Fronto Occipital fasciculus Right	3.78E-15	7.861195
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Frontal Aslant Tract Left	1.67E-14	7.674138
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Fronto Insular tract5 Left	4.34E-09	-5.870851
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Superior Longitudinal Fasciculus III Left	1.70E-14	7.671690
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Arcuate Anterior Segment Right	6.70E-09	5.798346
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Inferior Fronto Occipital fasciculus Left	2.66E-17	8.460688
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Arcuate Long Segment Left	2.17E-13	7.337708
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Fronto Insular tract5 Left	1.50E-25	-10.447966
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Frontal Aslant Tract Left	8.09E-13	7.195950
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Fronto Insular tract4 Right	3.34E-14	7.584401
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Superior Longitudinal Fasciculus III Right	2.23E-27	-10.839948
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Superior Longitudinal Fasciculus II Right	1.13E-09	6.089697
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Frontal Superior Longitudinal Right	2.14E-27	-10.843346
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Arcuate Anterior Segment Left	3.94E-14	7.562991
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Arcuate Posterior Segment Left	2.50E-10	6.327445
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Superior Longitudinal Fasciculus III Right	4.70E-12	6.914465
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Superior Longitudinal Fasciculus II Left	2.25E-10	6.343008
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	FA Fronto Insular Tract Right	6.03E-10	-6.189584
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Inferior Fronto Occipital fasciculus Left	1.20E-13	7.417311
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Arcuate Posterior Segment Left	1.31E-27	-10.888193
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Uncinate Left	1.07E-21	9.596993
102	rs160184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Fr		

D.1. Genomic loci associated with heritable language anatomical connectivities using MOSTest

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mygwasP (discovery -British-)	mygwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Fronto Insular tracts Right	1.34E-08	5.68118
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Uncinate Right	1.34E-20	9.305959
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Fronto Insular tracts Right	5.44E-25	-10.32492
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Arcuate Long Segment Left	5.10E-09	5.843830
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Superior Longitudinal Fasciculus I Left	1.04E-30	-11.520304
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Frontal Aslant Tract Right	2.73E-12	6.991090
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	FA Superior Longitudinal Fasciculus I Right	1.80E-13	-7.362729
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Fronto Insular tract4 Right	1.00E-08	5.730159
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Inferior Longitudinal Right	1.08E-11	6.759933
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Arcuate Anterior Segment Left	1.52E-10	6.397387
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L1 Superior Longitudinal Fasciculus III Left	2.43E-11	6.677325
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Superior Longitudinal Fasciculus II Right	1.28E-08	5.689141
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L1 Frontal Aslant Tract Right	3.17E-08	5.53548
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L1 Corpus callosum	1.50E-11	6.794672
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Arcuate Long Segment Left	7.94E-14	7.472771
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Uncinate Right	1.14E-38	-13.007257
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Fronto Insular tract4 Left	1.73E-10	-6.382109
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Superior Longitudinal Fasciculus I Right	4.23E-16	8.131766
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Uncinate Left	1.88E-40	-13.315482
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Fronto Insular tract3 Right	1.90E-11	6.173418
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Inferior Longitudinal Left	6.19E-28	-10.956462
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Superior Longitudinal Fasciculus III Left	5.91E-29	-11.167090
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Inferior Longitudinal Right	1.02E-14	7.736500
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Uncinate Left	1.21E-22	9.792542
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Arcuate Posterior Segment Right	4.83E-09	8.852767
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Superior Longitudinal Fasciculus II Left	3.05E-10	6.296238
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Superior Longitudinal Fasciculus III Right	9.78E-14	7.443781
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Inferior Fronto Occipital fasciculus Left	5.30E-18	8.646753
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L3 Arcuate Anterior Segment Right	3.93E-11	5.333271
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	L2 Arcuate Posterior Segment Left	9.53E-10	6.174767
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	ICVF Frontal Inferior longitudinal	8.71E-21	-9.350600
102	rs1610184	5	82822037	intronic	G	A	0.480	4.03E-25	3.63E-02	VCAN	MD Superior Longitudinal Fasciculus I Left	1.87E-12	7.043585
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L3 Superior Longitudinal Fasciculus II Left	3.80E-08	5.499868
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	MD Inferior Longitudinal Left	5.38E-11	5.659915
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Fronto Insular tract4 Right	5.87E-11	-6.546986
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L1 Inferior Longitudinal Left	1.07E-09	6.098036
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Superior Longitudinal Fasciculus II Right	1.34E-08	-5.680892
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Fronto Insular tract3 Right	5.70E-09	-5.825353
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L2 Inferior Longitudinal Left	1.50E-08	5.661298
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	MD Inferior Longitudinal Right	1.57E-10	6.398742
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L1 Inferior Fronto Occipital fasciculus Left	2.11E-13	7.341875
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Arcuate Long Segment Left	6.60E-09	-5.806961
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Fronto Insular tracts Right	2.60E-09	-5.955293
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Fronto Insular tracts Left	1.83E-10	-6.374661
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Inferior Fronto Occipital fasciculus Right	3.51E-15	-7.871307
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	MD Uncinate Right	2.53E-10	6.325435
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	MD Inferior Fronto Occipital fasciculus Right	7.37E-12	8.680231
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Corpus callosum	2.25E-13	-7.333056
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Arcuate Posterior Segment Right	1.18E-10	-6.442358
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	MD Corpus callosum	9.29E-09	5.743138
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Inferior Fronto Occipital fasciculus Left	2.23E-13	-7.334358
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Frontal Aslant Tract Right	3.48E-09	-5.907264
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L1 Inferior Fronto Occipital fasciculus Right	2.54E-10	6.324795
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L2 Uncinate Right	2.02E-08	5.610607
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L1 Uncinate Left	1.29E-08	5.697958
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Superior Longitudinal Fasciculus II Left	6.02E-09	-5.979266
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L3 Inferior Fronto Occipital fasciculus Right	6.65E-11	6.528414
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L1 Fronto Insular tracts Right	8.49E-11	6.401638
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L3 Inferior Fronto Occipital fasciculus Left	5.57E-11	6.554913
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Frontal Aslant Tract Left	2.50E-08	-5.573475
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Superior Longitudinal Fasciculus III Right	2.45E-09	-5.964439
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Arcuate Posterior Segment Left	5.50E-11	-6.556615
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L2 Uncinate Left	9.57E-12	6.812799
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Fronto Insular tracts Left	9.30E-10	-6.121013
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Arcuate Long Segment Right	6.57E-09	-5.801562
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L3 Inferior Longitudinal Left	1.23E-09	6.607683
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L3 Uncinate Left	1.37E-11	6.761474
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Inferior Longitudinal Right	3.57E-16	-8.152479
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L1 Inferior Longitudinal Right	4.21E-10	6.245884
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L1 Fronto Insular tract2 Right	8.32E-11	6.404315
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Arcuate Anterior Segment Left	1.53E-09	-6.040801
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Superior Longitudinal Fasciculus I Right	5.16E-10	-6.214298
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L3 Uncinate Right	6.49E-11	6.532081
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Superior Longitudinal Fasciculus I Left	5.31E-11	-6.562049
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	FA Superior Longitudinal Fasciculus II Left	6.42E-10	-6.179765
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	MO Fronto Insular tracts Right	4.93E-08	5.453714
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	OD Fronto Insular tract3 Right	1.81E-08	-5.629376
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	MO Fronto Insular tract2 Right	3.31E-08	5.524145
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	L1 Corpus callosum	1.16E-10	6.444237
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Uncinate Right	6.60E-14	-7.488548
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Uncinate Left	2.54E-14	-7.619770
102	rs309588	5	82852578	nRNA intronic	T	A	0.204	5.16E-28	1.55E-01	VCAN.VCAN-AS1	ICVF Inferior Longitudinal Left	3.92E-16	-8.140975
102	rs309588	5	8										

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mgwasP (discovery -British-)	mgwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L3 Fronto Insular tract2 Right	3.61E-14	-7.574321
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	FA Inferior Longitudinal Left	1.36E-37	12.814352
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	FA Inferior Fronto Occipital fasciculus Left	7.60E-29	12.688203
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	FA Frontal Aslant Tract Left	8.62E-22	9.592198
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ICVF Arcuate Anterior Segment Right	8.33E-59	16.160107
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	FA Uncinate Left	2.01E-19	12.603867
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ICVF Fronto Insular tract3 Right	2.99E-63	16.788730
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	MD Arcuate Anterior Segment Right	3.49E-27	-10.798826
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L2 Inferior Longitudinal Left	3.98E-51	-15.040544
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	MD Inferior Longitudinal Right	1.07E-51	-15.127479
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L1 Inferior Fronto Occipital fasciculus Left	5.34E-37	-12.708015
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ICVF Arcuate Long Segment Left	5.62E-59	16.220398
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	MD Frontal Aslant Tract Right	2.65E-25	-10.393648
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ISOVF Inferior Longitudinal Right	4.67E-23	9.888535
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L3 Superior Longitudinal Fasciculus I Left	6.34E-38	-12.873556
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L1 Fronto Insular tract4 Right	5.10E-11	-6.697959
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ICVF Fronto Insular tract2 Right	3.55E-50	14.895094
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ICVF Fronto Insular tract5 Right	1.87E-57	15.976378
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	FA Arcuate Anterior Segment Right	3.91E-19	8.939611
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	MD Fronto Insular tract5 Right	2.40E-13	-7.324107
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	FA Superior Longitudinal Fasciculus I Left	1.66E-29	11.279085
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L3 Frontal Aslant Tract Left	8.83E-26	-10.497878
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	OD Corpus callosum	9.43E-18	8.580713
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	FA Fronto Insular tract4 Right	1.13E-13	7.424574
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ICVF Inferior Fronto Occipital fasciculus Right	6.67E-92	20.332209
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L1 Arcuate Anterior Segment Left	2.49E-26	-10.616861
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	FA Arcuate Long Segment Left	2.09E-20	9.257859
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	MD Uncinate Right	1.10E-52	-15.239933
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	MD Uncinate Left	4.31E-14	-11.431244
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L3 Fronto Insular tract3 Right	5.43E-12	-6.893713
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	MD Inferior Fronto Occipital fasciculus Right	9.29E-54	-15.439575
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ICVF Corpus callosum	1.49E-90	20.179390
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L3 Superior Longitudinal Fasciculus III Right	3.56E-28	-11.008467
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ISOVF Frontal Superior Longitudinal Right	4.98E-09	5.847708
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ICVF Arcuate Posterior Segment Right	1.09E-67	17.386177
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ISOVF Inferior Fronto Occipital fasciculus Right	3.11E-23	9.929254
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	FA Frontal Inferior longitudinal Right	4.13E-14	7.567504
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	MD Corpus callosum	3.88E-47	-14.419935
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	MD Fronto Insular tract3 Right	6.23E-18	-8.628263
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ISOVF Inferior Longitudinal Left	1.99E-14	7.653707
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L2 Inferior Fronto Occipital fasciculus Right	1.73E-41	-13.492504
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L3 Fronto Insular tract4 Right	3.41E-22	-9.687356
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L1 Arcuate Long Segment Right	3.25E-17	-8.436820
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ISOVF Frontal Aslant Tract Right	1.24E-09	6.078517
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	OD Arcuate Posterior Segment Right	2.55E-09	5.958035
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L3 Arcuate Posterior Segment Right	4.84E-28	-10.978763
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ICVF Fronto Insular tract5 Left	3.03E-46	-14.277343
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	FA Superior Longitudinal Fasciculus III Left	3.07E-23	9.903253
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L2 Frontal Superior Longitudinal Left	1.14E-10	-6.447239
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	MD Superior Longitudinal Fasciculus III Left	1.18E-35	-12.463672
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ICVF Inferior Fronto Occipital fasciculus Left	6.32E-92	20.334910
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L1 Superior Longitudinal Fasciculus II Left	9.86E-19	-8.836666
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L1 Fronto Insular tract3 Right	2.34E-14	-7.630382
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ICVF Frontal Aslant Tract Right	1.51E-66	17.232664
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L1 Inferior Fronto Occipital fasciculus Right	3.24E-31	-11.620692
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ICVF Frontal Superior Longitudinal Left	5.69E-47	14.393957
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L3 Fronto Insular tract5 Left	2.85E-09	-5.949824
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L1 Arcuate Posterior Segment Left	4.68E-22	-9.655101
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	MD Fronto Insular tract2 Left	1.12E-14	-7.724840
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	MD Arcuate Posterior Segment Right	4.19E-37	-12.729090
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	FA Superior Longitudinal Fasciculus II Right	1.42E-24	-10.232062
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L2 Uncinate Right	6.87E-49	-14.699584
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L2 Corpus callosum	3.12E-45	-14.113869
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	MD Fronto Insular tract5 Left	3.84E-08	-5.497972
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L1 Uncinate Left	2.28E-24	-10.186294
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L2 Superior Longitudinal Fasciculus III Right	3.07E-31	-11.626057
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L3 Superior Longitudinal Fasciculus III Left	5.75E-31	-11.571516
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	FA Frontal Superior Longitudinal Left	2.92E-10	6.302974
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L1 Uncinate Right	2.91E-29	-11.229709
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ICVF Frontal Inferior longitudinal Left	3.49E-41	13.440660
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ICVF Superior Longitudinal Fasciculus II Left	1.33E-05	17.106213
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L3 Inferior Fronto Occipital fasciculus Right	1.18E-54	-15.569282
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L2 Frontal Aslant Tract Left	4.00E-29	-11.201635
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ISOVF Superior Longitudinal Fasciculus I Right	3.56E-08	5.511349
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L2 Fronto Insular tract5 Right	5.26E-09	-5.838574
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	ICVF Fronto Insular tract5 Left	1.95E-13	7.352316
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L2 Superior Longitudinal Fasciculus III Left	4.09E-32	-11.796055
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	FA Arcuate Anterior Segment Left	1.68E-12	7.065300
102	rs10052710	5	82860025	nCRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN.VCAN-AS1	L2 Arcuate Anterior Segment Right	3.98E-21	-9.433104
102	rs												

D.1. Genomic loci associated with heritable language anatomical connectivities using MOSTest

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mvgwasP (discovery -British-)	mvgwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ICVF Superior Longitudinal Fasciculus III Right	1.62E-66	17.228440
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L2 Superior Longitudinal Fasciculus II Right	1.73E-29	-11.275508
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ICVF Frontal Superior Longitudinal Right	1.78E-53	15.395450
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	OD Inferior Fronto Occipital Fasciculus Left	7.85E-13	7.163734
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	MD Arcuate Anterior Segment Left	1.00E-28	-11.110989
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L3 Arcuate Posterior Segment Left	1.11E-21	-9.565999
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ISOVF Uncinate Right	2.41E-09	5.967349
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	MD Frontal Inferior longitudinal Left	1.15E-14	-7.72154
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L1 Superior Longitudinal Fasciculus III Right	2.08E-23	-9.969208
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	MD Superior Longitudinal Fasciculus II Left	5.31E-33	-11.966718
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	FA Frontal Aslant Tract Right	1.87E-20	9.269461
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L1 Frontal Aslant Tract Left	5.26E-18	-8.847552
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L2 Inferior Fronto Occipital Fasciculus Left	6.45E-47	-14.384758
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ICVF Arcuate Posterior Segment Left	3.55E-54	16.931661
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L2 Uncinate Left	4.87E-50	-14.873906
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L2 Fronto Insular tract3 Left	5.99E-13	-7.200603
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	FA Arcuate Long Segment Right	9.70E-21	9.338259
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ICVF Fronto Insular tract5 Left	2.50E-52	15.222791
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ICVF Arcuate Long Segment Right	2.12E-59	16.253178
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L3 Inferior Longitudinal Left	2.55E-54	-15.519731
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	OD Superior Longitudinal Fasciculus III Right	1.12E-08	5.719349
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	MD Fronto Insular tract5 Left	5.77E-20	-9.348502
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L3 Corpus callosum	3.89E-43	-13.769543
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L3 Uncinate Left	1.35E-52	-15.263124
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ICVF Inferior Longitudinal Right	1.23E-95	20.749887
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L3 Arcuate Long Segment Right	3.67E-31	-11.609333
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	FA Arcuate Posterior Segment Left	7.25E-15	7.78012
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ICVF Fronto Insular tract2 Left	6.99E-37	12.688869
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L2 Superior Longitudinal Fasciculus I Right	2.15E-31	-11.655714
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	FA Fronto Insular tract3 Right	3.71E-09	5.896533
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	FA Corpus callosum	3.20E-38	12.026243
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L2 Superior Longitudinal Fasciculus I Left	4.15E-31	-11.599427
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	MD Frontal Inferior longitudinal Right	4.32E-19	-8.928500
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L1 Inferior Longitudinal Right	1.64E-33	-12.05362
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	MD Arcuate Posterior Segment Left	7.99E-32	-11.739891
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L1 Frontal Inferior Longitudinal Right	1.46E-09	-6.048054
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L1 Fronto Insular tract2 Right	2.28E-11	-6.687862
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ICVF Arcuate Anterior Segment Left	1.30E-68	17.659499
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	OD Arcuate Anterior Segment Left	5.07E-09	5.844979
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L1 Superior Longitudinal Fasciculus I Right	3.98E-16	-8.33924
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ISOVF Corpus callosum	5.18E-09	5.841244
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L3 Arcuate Anterior Segment Left	3.27E-19	-8.059343
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ICVF Superior Longitudinal Fasciculus I Right	8.73E-88	19.861713
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	FA Arcuate Posterior Segment Right	2.46E-15	7.915887
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	FA Inferior Longitudinal Right	3.79E-44	13.434695
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	OD Inferior Longitudinal Right	2.52E-14	7.627040
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	MD Superior Longitudinal Fasciculus I Right	2.55E-38	-12.943630
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	FA Fronto Insular tract5 Right	1.18E-10	6.441556
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	MD Arcuate Long Segment Right	6.07E-35	-12.332317
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	OD Inferior Fronto Occipital Fasciculus Right	1.94E-12	7.125738
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	MD Fronto Insular tract3 Right	7.48E-13	-7.170331
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L3 Uncinate Right	2.50E-55	-15.667974
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ICVF Fronto Insular tract3 Right	2.25E-46	14.298044
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	FA Frontal Inferior longitudinal Left	3.44E-08	5.517370
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L2 Arcuate Long Segment Left	4.73E-23	-9.887253
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ICVF Superior Longitudinal Fasciculus I Left	1.28E-85	19.609669
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ISOVF Arcuate Anterior Segment Right	2.76E-08	5.555774
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L3 Frontal Superior Longitudinal Left	2.85E-08	-5.550447
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L2 Frontal Aslant Tract Right	4.98E-27	-10.765956
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	FA Superior Longitudinal Fasciculus I Right	2.80E-33	12.019693
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L2 Arcuate Long Segment Right	1.15E-25	-10.472907
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L2 Fronto Insular tract4 Right	6.60E-11	-6.529048
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	FA Superior Longitudinal Fasciculus II Left	1.32E-23	10.04218
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L1 Fronto Insular tract3 Left	1.02E-08	-5.727486
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L2 Inferior Longitudinal Right	1.51E-45	-14.164917
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L1 Arcuate Posterior Segment Right	1.54E-27	-10.873801
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L2 Arcuate Anterior Segment Left	2.60E-22	-9.745273
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L1 Superior Longitudinal Fasciculus III Left	1.12E-27	-10.902361
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L3 Superior Longitudinal Fasciculus II Right	5.09E-30	-11.382888
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L1 Frontal Aslant Tract Right	1.11E-15	-8.044571
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ISOVF Superior Longitudinal Fasciculus III Right	2.04E-11	6.703276
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	MD Fronto Insular tract2 Right	6.39E-13	-7.191925
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L1 Corpus callosum	4.50E-39	-13.076252
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	MD Arcuate Long Segment Left	1.61E-36	-12.621229
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ICVF Uncinate Right	5.29E-94	20.568161
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ICVF Fronto Insular tract4 Left	8.55E-21	9.352838
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L3 Superior Longitudinal Fasciculus I Right	1.34E-44	-14.010869
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	ICVF Uncinate Left	2.47E-01	20.267982
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.1E-06	VCAN.VCAN-AS1	L3 Fronto Insular tract3 Right	1.23E-18	-8.811597
102	rs10052710	5	82860025	ncRNA intronic	G	T	0						

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mgwasP (discovery -British-)	mgwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN-VCAN-AS1	L3 Fronto Insular tract3 Left	8.32E-14	-7.465223
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN-VCAN-AS1	MD Inferior Fronto Occipital fasciculus Left	7.15E-08	-16.036083
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN-VCAN-AS1	L3 Arcuate Anterior Segment Right	9.03E-25	-10.276080
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN-VCAN-AS1	L2 Arcuate Posterior Segment Left	6.89E-28	-10.046752
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN-VCAN-AS1	ICVF Frontal Inferior longitudinal Right	2.96E-47	14.438477
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN-VCAN-AS1	MD Superior Longitudinal Fasciculus I Left	7.63E-36	-12.498301
102	rs10052710	5	82860025	ncRNA intronic	G	T	0.195	7.72E-111	6.11E-06	VCAN-VCAN-AS1	L1 Superior Longitudinal Fasciculus I Left	7.89E-17	-8.332809
105	rs115136616	5	93021116	intronic	T	C	0.047	6.85E-16	1.73E-01	FAM172A	MO Inferior Fronto Occipital fasciculus Right	6.09E-10	-6.188109
108	rs7729285	5	122754416	intronic	G	A	0.431	9.63E-11	2.12E-02	CEP120	OD Frontal Aslant Tract Left	4.21E-08	5.481863
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Superior Longitudinal Fasciculus II Right	4.29E-09	-5.872429
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Arcuate Anterior Segment Right	1.31E-12	-7.093482
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Arcuate Long Segment Left	1.90E-18	-8.793229
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	L3 Superior Longitudinal Fasciculus I Left	1.08E-08	-5.772267
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Inferior Fronto Occipital fasciculus Right	1.59E-10	-6.401701
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	FA Arcuate Long Segment Left	8.72E-10	-6.131196
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Corpus callosum	1.38E-19	-9.054096
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Arcuate Posterior Segment Right	1.67E-15	-7.963845
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	MD Corpus callosum	1.92E-09	6.004444
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	L3 Arcuate Posterior Segment Right	2.79E-10	6.350222
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Inferior Fronto Occipital fasciculus Left	4.05E-09	-5.882072
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Frontal Aslant Tract Right	1.67E-08	-5.643236
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	MD Arcuate Posterior Segment Right	1.92E-09	6.004774
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	L2 Corpus callosum	8.33E-09	-7.761588
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	L3 Superior Longitudinal Fasciculus III Left	2.38E-08	5.582024
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	L3 Arcuate Long Segment Left	1.44E-12	7.084597
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Frontal Aslant Tract Left	6.75E-09	-5.739574
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Superior Longitudinal Fasciculus III Right	8.18E-10	-6.141472
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Frontal Superior Longitudinal Right	3.81E-08	-5.499387
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Arcuate Posterior Segment Left	7.51E-15	-7.775261
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Arcuate Long Segment Right	2.91E-10	-8.205626
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	L3 Corpus callosum	5.14E-10	6.246288
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Inferior Longitudinal Right	8.70E-13	-7.149618
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	L3 Arcuate Long Segment Right	8.55E-11	-6.490505
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	FA Corpus callosum	1.05E-08	-5.723085
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	MD Arcuate Posterior Segment Left	5.17E-09	5.841690
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Arcuate Anterior Segment Left	2.94E-11	-6.949799
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Superior Longitudinal Fasciculus I Right	1.23E-10	-6.439765
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	FA Arcuate Posterior Segment Right	2.32E-08	-5.886520
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	FA Inferior Longitudinal Right	4.30E-08	-5.478195
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	MD Arcuate Long Segment Right	1.58E-09	6.035883
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Superior Longitudinal Fasciculus I Left	1.02E-11	-6.803649
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	MD Arcuate Long Segment Left	1.21E-10	6.438217
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Inferior Longitudinal Left	1.23E-13	-7.413313
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Superior Longitudinal Fasciculus III Left	4.55E-13	-7.237987
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	L3 Inferior Longitudinal Right	2.04E-08	5.608463
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	ICVF Frontal Inferior longitudinal Right	4.43E-08	-5.472616
109	rs4150210	5	139721913	intronic	A	T	0.259	1.79E-45	3.91E-03	HBEGF	MD Superior Longitudinal Fasciculus I Left	4.47E-08	5.471889
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Frontal Aslant Tract Left	5.82E-10	-6.195533
112	rs79318212	6	1365244	intergenic	C	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	L3 Fronto Insular tract3 Right	6.77E-09	-5.796631
112	rs79318212	6	1365244	intergenic	C	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	L3 Frontal Inferior longitudinal Left	1.12E-08	-5.712593
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Inferior Fronto Occipital fasciculus Left	9.14E-09	-5.745933
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Superior Longitudinal Fasciculus II Right	5.13E-12	-6.501921
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Frontal Inferior longitudinal Right	3.38E-09	-5.912477
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Inferior Longitudinal Left	1.12E-08	-5.712204
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Superior Longitudinal Fasciculus III Left	1.73E-13	-7.368442
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Arcuate Anterior Segment Left	1.46E-09	-6.048258
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	MD Fronto Insular tract3 Right	1.33E-11	-6.764804
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Arcuate Posterior Segment Left	1.91E-09	-6.005233
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	L3 Superior Longitudinal Fasciculus III Right	9.49E-09	-5.739634
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Inferior Longitudinal Left	2.28E-09	-5.979312
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Frontal Aslant Tract Right	5.99E-09	-5.871123
112	rs79318212	6	1365244	intergenic	C	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Arcuate Long Segment Left	1.05E-08	-5.720895
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Superior Longitudinal Fasciculus II Left	6.80E-12	-6.861857
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	L2 Fronto Insular tract3 Right	1.61E-10	-6.394204
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	MD Frontal Inferior longitudinal Left	3.51E-09	-5.905781
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	L1 Superior Longitudinal Fasciculus III Right	3.13E-09	-5.924546
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Corpus callosum	3.44E-08	-5.517474
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Frontal Inferior longitudinal Left	1.15E-08	-5.705228
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Arcuate Anterior Segment Right	1.38E-08	-5.676228
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	L1 Frontal Aslant Tract Right	1.96E-08	-5.615074
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Superior Longitudinal Fasciculus III Right	4.30E-13	-7.245699
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Arcuate Long Segment Right	5.38E-09	-5.828625
112	rs79318212	6	1365244	intergenic	C	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	ISOVF Arcuate Posterior Segment Right	1.89E-10	-6.377154
112	rs79318212	6	1365244	intergenic	C	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	MD Superior Longitudinal Fasciculus III Right	2.69E-09	-5.575691
112	rs79318212	6	1365244	intergenic	G	A	0.106	1.45E-14	5.92E-01	RP4-668J24.2	L1 Fronto Insular tract3 Right	3.13E-09	-5.924638
113	rs1800562	6	26093141	exonic	A	G	0.078	1.83E-10	5.26E-02	HFE	ICVF Fronto Insular tract3 Right	2.35E-08	-5.584043
113	rs1800562	6	26093141	exonic	A	G	0.078	1.83E-10	5.26E-02	HFE	ICVF Fronto Insular tract3 Right	2.13E-10	-6.351889

D.1. Genomic loci associated with heritable language anatomical connectivities using MOSTest

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mygwasP (discovery -British-)	mygwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate Z	univariate Z
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	ICVF Arcuate Anterior Segment Right	2.60E-08	-5.666167
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	L1 Fronto Insular tract4 Right	3.97E-13	7.250620
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	ICVF Inferior Fronto Occipital fasciculus Right	2.94E-09	-5.934977
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	MD Inferior Fronto Occipital fasciculus Right	3.93E-08	5.494021
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	ICVF Corpus callosum	3.70E-09	-5.894504
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	L1 Superior Longitudinal Fasciculus II Left	2.20E-10	6.345283
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	L1 Fronto Insular tract3 Right	5.96E-10	6.191563
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	ICVF Frontal Aslant Tract Right	3.59E-08	-5.590742
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	L2 Corpus callosum	1.12E-08	5.773060
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	ICVF Superior Longitudinal Fasciculus III Right	2.76E-08	-5.555774
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	ICVF Inferior Longitudinal Right	2.08E-08	-5.542630
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	L2 Superior Longitudinal Fasciculus I Left	2.73E-08	5.557684
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	OD Fronto Insular tract4 Right	2.01E-13	-7.348143
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	MO Fronto Insular tract4 Right	1.01E-11	6.805337
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	OD Fronto Insular tract3 Right	2.07E-08	-5.605954
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	MD Inferior Fronto Occipital fasciculus Left	3.27E-08	5.526334
113	rs146863150	6	26408472	intronic	CATAT	C	0.103	7.93E-22	5.93E-04	BTN3A1	L2 Arcuate Posterior Segment Left	4.68E-09	5.740156
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	ICVF Arcuate Anterior Segment Right	7.21E-09	-5.785976
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	MD Inferior Longitudinal Right	4.20E-08	5.482432
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	MD Frontal Aslant Tract Right	3.30E-08	5.524620
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	L1 Fronto Insular tract4 Right	6.93E-14	7.489196
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	ICVF Fronto Insular tract2 Right	2.96E-08	-5.543756
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	ICVF Inferior Fronto Occipital fasciculus Right	1.26E-10	-6.431894
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	MD Inferior Fronto Occipital fasciculus Right	4.04E-09	5.882638
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	ICVF Corpus callosum	2.66E-10	-6.137386
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	L3 Superior Longitudinal Fasciculus III Right	4.73E-08	5.461322
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	MD Corpus callosum	2.44E-08	5.577441
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	ICVF Inferior Fronto Occipital fasciculus Left	6.29E-09	-5.808930
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	L1 Superior Longitudinal Fasciculus II Left	4.34E-12	6.925611
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	L1 Fronto Insular tract3 Right	1.28E-10	6.429275
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	ICVF Frontal Aslant Tract Right	5.54E-09	-5.830134
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	L2 Corpus callosum	1.50E-09	6.044877
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	L2 Superior Longitudinal Fasciculus III Right	1.28E-08	5.688032
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	L3 Inferior Fronto Occipital fasciculus Right	1.00E-08	5.720438
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	ICVF Superior Longitudinal Fasciculus III Right	7.93E-09	-5.769418
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	MD Superior Longitudinal Fasciculus II Left	1.43E-08	5.470182
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	L3 Corpus callosum	3.18E-08	5.532312
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	ICVF Inferior Longitudinal Right	1.67E-09	-6.027422
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	L2 Superior Longitudinal Fasciculus I Right	4.02E-08	5.489855
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	L2 Superior Longitudinal Fasciculus I Left	3.33E-09	5.914022
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	OD Fronto Insular tract4 Right	1.34E-14	-7.701769
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	MO Fronto Insular tract4 Right	5.17E-13	7.220820
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	OD Fronto Insular tract3 Right	3.72E-08	-5.503509
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	ICVF Inferior Longitudinal Left	3.50E-08	-5.514405
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	MD Superior Longitudinal Fasciculus III Right	2.15E-08	5.599325
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	OD Superior Longitudinal Fasciculus II Left	4.95E-09	-5.848971
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	MD Inferior Fronto Occipital fasciculus Left	4.85E-09	5.852100
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	L3 Arcuate Anterior Segment Right	2.31E-08	5.586711
113	rs341936306	6	27425644	intronic	C	G	0.104	1.17E-27	1.29E-03	ZNF184	L2 Arcuate Posterior Segment Left	3.33E-09	5.914387
113	rs71540548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	4.58E-03	HNRNPA1P1	L1 Fronto Insular tract4 Right	3.47E-12	6.957306
113	rs71540548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	4.58E-03	HNRNPA1P1	L1 Fronto Insular tract4 Right	3.47E-12	6.957306
113	rs71540548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	4.58E-03	HNRNPA1P1	L1 Fronto Insular tract4 Right	3.47E-12	6.957306
113	rs71540548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	4.58E-03	HNRNPA1P1	L1 Superior Longitudinal Fasciculus II Left	4.23E-08	5.480995
113	rs71540548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	4.58E-03	HNRNPA1P1	L1 Superior Longitudinal Fasciculus II Left	4.23E-08	5.480995
113	rs71540548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	4.58E-03	HNRNPA1P1	L1 Superior Longitudinal Fasciculus II Left	4.23E-08	5.480995
113	rs71540548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	4.58E-03	HNRNPA1P1	L1 Fronto Insular tract3 Right	1.98E-08	5.615684
113	rs71540548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	4.58E-03	HNRNPA1P1	L1 Fronto Insular tract3 Right	1.98E-08	5.615684
113	rs71540548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	4.58E-03	HNRNPA1P1	L1 Fronto Insular tract3 Right	1.98E-08	5.615684
113	rs71540548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	4.58E-03	HNRNPA1P1	L1 Fronto Insular tract4 Right	4.03E-12	-6.936293
113	rs71540548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	4.58E-03	HNRNPA1P1	OD Fronto Insular tract4 Right	4.03E-12	-6.936293
113	rs71540548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	4.58E-03	HNRNPA1P1	OD Fronto Insular tract4 Right	4.03E-12	-6.936293
113	rs71540548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	4.58E-03	HNRNPA1P1	MO Fronto Insular tract4 Right	1.21E-08	5.597879
113	rs71540548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	4.58E-03	HNRNPA1P1	MO Fronto Insular tract4 Right	1.21E-08	5.597879
113	rs71540548	6	27493970	intergenic	TTTCTTA	T	0.010	1.41E-13	4.58E-03	HNRNPA1P1	MO Fronto Insular tract4 Right	1.21E-08	5.597879
113	rs34676049	6	28453618	intergenic	G	T	0.108	1.93E-28	4.67E-04	GPX6	MD Superior Longitudinal Fasciculus II Right	2.04E-08	5.608634
113	rs34676049	6	28453618	intergenic	G	T	0.108	1.93E-28	4.67E-04	GPX6	L1 Inferior Longitudinal Left	3.49E-08	5.514751
113	rs34676049	6	28453618	intergenic	G	T	0.108	1.93E-28	4.67E-04	GPX6	ICVF Arcuate Anterior Segment Right	7.32E-09	-5.783392
113	rs34676049	6	28453618	intergenic	G	T	0.108	1.93E-28	4.67E-04	GPX6	MD Arcuate Anterior Segment Right	3.33E-08	5.528453
113	rs34676049	6	28453618	intergenic	G	T	0.108	1.93E-28	4.67E-04	GPX6	MD Inferior Longitudinal Right	1.22E-08	5.696687
113	rs34676049	6	28453618	intergenic	G	T	0.108	1.93E-28	4.67E-04	GPX6	L1 Inferior Fronto Occipital fasciculus Left	1.87E-08	5.624051
113	rs34676049	6	28453618	intergenic	G	T	0.108	1.93E-28	4.67E-04	GPX6	MD Fronto Insular Tract Right	1.34E-08	5.680266
113	rs34676049	6	28453618	intergenic	G	T	0.108	1.93E-28	4.67E-04	GPX6	L1 Fronto Insular tract4 Right	2.47E-13	7.320299
113	rs34676049	6	28453618	intergenic	G	T	0.108	1.93E-28	4.67E-04	GPX6	ICVF Inferior Fronto Occipital fasc		

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mygwasP (discovery -British-)	mygwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	ICVF Arcuate Anterior Segment Right	3.66E-09	-5.899024
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	ICVF Fronto Insular tract3 Right	1.25E-08	-5.692288
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	MD Arcuate Anterior Segment Right	1.27E-08	-5.690448
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	MD Inferior Longitudinal Right	9.78E-09	-5.734440
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L1 Inferior Fronto Occipital Fasciculus Left	9.98E-09	-5.731360
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	MD Frontal Aslant Tract Right	8.89E-10	-6.128223
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L1 Fronto Insular tract4 Right	2.09E-13	-7.342959
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	ICVF Fronto Insular tract2 Right	2.26E-08	-5.590801
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	ICVF Inferior Fronto Occipital fasciculus Right	1.19E-10	-6.440825
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	MD Inferior Fronto Occipital fasciculus Right	1.10E-09	-6.093867
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	ICVF Corpus callosum	1.59E-10	-6.396668
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L3 Superior Longitudinal Fasciculus III Right	4.26E-09	-5.873738
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	MD Corpus callosum	3.01E-09	-5.930985
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	MD Superior Longitudinal Fasciculus III Left	1.26E-09	-6.072826
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	ICVF Inferior Fronto Occipital fasciculus Left	9.24E-10	-6.122008
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L1 Superior Longitudinal Fasciculus II Left	5.25E-13	-7.218046
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L1 Fronto Insular tract3 Right	1.70E-10	-6.385829
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	ICVF Frontal Aslant Tract Right	1.08E-09	-6.097056
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L2 Uncinate Right	2.39E-08	-5.81312
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L2 Corpus callosum	1.99E-10	-6.369065
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L2 Superior Longitudinal Fasciculus III Right	2.57E-09	-5.95911
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L3 Superior Longitudinal Fasciculus III Left	2.76E-08	-5.550499
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L3 Inferior Fronto Occipital Fasciculus Right	3.27E-09	-5.947425
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L2 Superior Longitudinal Fasciculus III Left	1.30E-08	-5.685667
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L3 Inferior Fronto Occipital fasciculus Left	2.65E-08	-5.852922
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	ICVF Frontal Aslant Tract Left	6.04E-09	-5.815518
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	MD Frontal Aslant Tract Left	7.53E-09	-5.778638
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	ICVF Superior Longitudinal Fasciculus III Right	1.91E-09	-6.055011
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	MD Arcuate Anterior Segment Left	3.85E-10	-6.260021
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	MD Superior Longitudinal Fasciculus II Left	7.62E-09	-5.776165
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L1 Frontal Aslant Tract Left	2.27E-10	-6.341587
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L2 Inferior Fronto Occipital fasciculus Left	1.93E-08	-5.617859
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L3 Corpus callosum	7.74E-09	-5.774076
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	ICVF Inferior Longitudinal Right	2.45E-09	-5.964970
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L2 Superior Longitudinal Fasciculus I Right	4.03E-08	-5.489443
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L2 Superior Longitudinal Fasciculus I Left	7.78E-09	-5.773072
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L1 Inferior Longitudinal Right	3.66E-08	-5.506381
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	MD Arcuate Posterior Segment Left	6.51E-09	-5.803170
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	OD Fronto Insular tract2 Right	2.36E-13	-7.326691
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L3 Arcuate Anterior Segment Left	2.24E-08	-5.592437
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	MO Fronto Insular tract4 Right	1.39E-12	-7.087693
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L1 Frontal Aslant Tract Right	2.41E-10	-6.332573
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	MD Arcuate Long Segment Left	3.25E-09	-5.918553
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	ICVF Inferior Longitudinal Left	2.02E-08	-5.609984
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	ICVF Superior Longitudinal Fasciculus III Left	4.44E-08	-5.472580
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L3 Inferior Longitudinal Right	2.75E-08	-5.556880
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	MD Superior Longitudinal Fasciculus III Right	1.37E-09	-6.058673
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	OD Superior Longitudinal Fasciculus II Left	3.88E-11	-6.608602
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	MD Inferior Fronto Occipital fasciculus Left	4.18E-10	-6.247118
113	rs3131856	6	29607101	intergenic	C	T	0.121	4.75E-33	7.66E-04	SUMO2P1	L3 Arcuate Anterior Segment Right	4.67E-09	-5.88363
114	rs9380899	6	39794806	intronic	G	C	0.271	8.76E-12	5.17E-01	DAAM2	L2 Arcuate Posterior Segment Left	1.75E-10	-6.381562
114	rs9380899	6	39794806	intronic	G	C	0.271	8.76E-12	5.17E-01	DAAM2	ISOVF Superior Longitudinal Fasciculus II Right	3.30E-09	-5.915786
114	rs9380899	6	39794806	intronic	G	C	0.271	8.76E-12	5.17E-01	DAAM2	ISOVF Inferior Longitudinal Right	2.06E-10	-6.356746
114	rs9380899	6	39794806	intronic	G	C	0.271	8.76E-12	5.17E-01	DAAM2	L1 Fronto Insular tracts Left	1.88E-08	-5.622600
114	rs9380899	6	39794806	intronic	G	C	0.271	8.76E-12	5.17E-01	DAAM2	MD Fronto Insular tracts Left	1.33E-09	-6.063797
114	rs9380899	6	39794806	intronic	G	C	0.271	8.76E-12	5.17E-01	DAAM2	ISOVF Inferior Fronto Occipital fasciculus Right	8.79E-09	-5.752652
114	rs9380899	6	39794806	intronic	G	C	0.271	8.76E-12	5.17E-01	DAAM2	ISOVF Inferior Longitudinal Left	8.48E-10	-6.135636
114	rs9380899	6	39794806	intronic	G	C	0.271	8.76E-12	5.17E-01	DAAM2	ISOVF Superior Longitudinal Fasciculus II Left	5.77E-11	-6.549600
114	rs9380899	6	39794806	intronic	G	C	0.271	8.76E-12	5.17E-01	DAAM2	ISOVF Superior Longitudinal Fasciculus I Left	1.40E-08	-5.673401
114	rs9380899	6	39794806	intronic	G	C	0.271	8.76E-12	5.17E-01	DAAM2	ISOVF Superior Longitudinal Fasciculus I Right	3.28E-09	-5.917070
114	rs9380899	6	39794806	intronic	G	C	0.271	8.76E-12	5.17E-01	DAAM2	ICVF Fronto Insular tracts Left	2.32E-08	-5.859606
114	rs9380899	6	39794806	intronic	G	C	0.271	8.76E-12	5.17E-01	DAAM2	L2 Fronto Insular tracts Left	4.99E-10	-6.219405
114	rs9380899	6	39794806	intronic	G	C	0.271	8.76E-12	5.17E-01	DAAM2	ISOVF Corpus callosum	4.90E-09	-5.850381
114	rs9380899	6	39794806	intronic	G	C	0.271	8.76E-12	5.17E-01	DAAM2	MD Uncinate Left	2.94E-08	-5.545184
117	rs3799711	6	126075313	intronic	C	T	0.480	2.29E-27	3.15E-01	HEY2	MO Frontal Aslant Tract Right	4.12E-08	-5.485244
118	rs1120786	6	1260808637	intergenic	G	T	0.450	2.50E-22	4.85E-01	MIR588	MD Superior Longitudinal Fasciculus III Left	9.03E-09	-5.748096
118	rs1120786	6	1260808637	intergenic	G	T	0.450	2.50E-22	4.85E-01	MIR588	MO Frontal Aslant Tract Left	2.38E-10	-6.334651
119	rs12192990	6	131021978	intronic	G	C	0.412	5.37E-34	3.98E-01	MIR588	ISOVF Superior Longitudinal Fasciculus II Left	9.93E-10	-6.110826
122	rs798487	7	2802943	intronic	A	G	0.298	4.19E-17	1.46E-01	AMZ1:GNA12	MD Corpus callosum	8.32E-10	-6.138724
122	rs798487	7	2802943	intronic	A	G	0.298	4.19E-17	1.46E-01	AMZ1:GNA12	ISOVF Superior Longitudinal Fasciculus III Left	1.18E-13	-7.419111
122	rs798487	7	2802943	intronic	A	G	0.298	4.19E-17	1.46E-01	AMZ1:GNA12	ISOVF Arcuate Anterior Segment Left	1.94E-09	-6.002895
122	rs798487	7	2802943	intronic	A	G	0.298	4.19E-17	1.46E-01	AMZ1:GNA12	L1 Frontal Superior Longitudinal Left	8.69E-10	-6.131798
122	rs798487	7	2802943	intronic	A	G	0.298	4.19E-17	1.46E-01	AMZ1:GNA12	MO Arcuate Long Segment Left	5.39E-10	-6.207380
122	rs798487	7	2802943	intronic	A	G	0.298	4.19E-17	1.46E-01	AMZ1:GNA12	ISOVF Arcuate Anterior Segment Right	8.09E-10	-6.143071
122	rs798487	7	2802943	intronic	A	G	0.298	4.19E-17	1.46E-01	AMZ1:GNA12	ISOVF Superior Longitudinal Fasciculus III Right	2.16E-10	-6.349821
122	rs798487	7	2802943	intronic	A	G	0.298	4.19E-17	1.46E-01	AMZ1:GNA12	ISOVF Arcuate Posterior Segment Right	3.31E-10	-6.283404
123	rs67248060	7	18899163	intronic	T	C	0.366	1.33E-22	4.78E-01	HDAC9	MD Inferior Longitudinal Left	4.49E-11	-6.869611
123	rs67248060	7	18899163	intronic	T	C	0.366	1.33E-22	4.78E-01	HDAC9	L1 Inferior Longitudinal Left	2.94E-09	-5.934912
123	rs67248060	7	18899163	intronic	T	C	0.366	1.33E-22	4.78E-01	HDAC9	L2 Inferior Longitudinal Left	4.56E-10	-6.233689
123	rs67248060	7	18899163	intronic	T	C	0.366	1.33E-22	4.78E-01	HDAC9	MD Inferior Longitudinal Right	3.21E-10	-6.888206
123	rs67248060	7	18899163	intronic	T	C	0.366	1.33E-22	4.78E-01	HDAC9	ISOVF Superior Longitudinal Fasciculus III Left	4.41E-09	-5.868137
123	rs67248060	7	18899163	intronic	T	C	0.366	1.33E-22	4.78E-01	HDAC9	MD Superior Longitudinal Fasciculus III Left	7.78E-09	-5.773181
123	rs67248060	7	18899163	intronic	T	C	0.366	1.33E-22	4.78E-01	HDAC9	L3 Arcuate Long Segment Left	1.62E-08	

D.1. Genomic loci associated with heritable language anatomical connectivities using MOSTest

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mvgwasP (discovery -British-)	mvgwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	MD Arcuate Anterior Segment Right	9.19E-09	-5.744974
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L1 Inferior Fronto Occipital Fasciculus Left	4.84E-09	-5.852606
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Arcuate Long Segment Left	4.66E-23	9.889944
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L3 Superior Longitudinal Fasciculus I Left	1.37E-10	-6.418751
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Fronto Insular tract5 Right	3.68E-09	5.897804
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Inferior Fronto Occipital Fasciculus Right	2.85E-11	6.659320
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L1 Arcuate Anterior Segment Left	3.56E-09	-5.903230
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Corpus callosum	2.30E-20	9.247327
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Arcuate Posterior Segment Right	3.15E-21	9.457741
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	MD Corpus callosum	2.11E-08	-5.603133
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	OD Superior Longitudinal Fasciculus III Left	9.10E-09	5.746649
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L1 Arcuate Long Segment Right	8.27E-09	-5.764987
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L3 Arcuate Posterior Segment Right	1.85E-09	-6.010395
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Fronto Insular tract3 Left	8.69E-16	8.044037
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	MD Superior Longitudinal Fasciculus III Left	1.94E-08	-5.616888
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Inferior Fronto Occipital Fasciculus Left	3.21E-09	5.920425
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Frontal Aslant Tract Right	1.46E-14	7.690597
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L1 Inferior Fronto Occipital Fasciculus Right	3.93E-10	-6.256966
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Frontal Superior Longitudinal Left	9.51E-13	7.137372
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L1 Arcuate Posterior Segment Left	2.91E-11	-6.651234
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	MD Arcuate Posterior Segment Right	1.53E-12	-7.071857
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Superior Longitudinal Fasciculus II Left	7.31E-17	8.341023
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L3 Arcuate Long Segment Left	3.26E-11	-6.634489
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L1 Arcuate Long Segment Left	1.41E-08	-5.672214
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Frontal Aslant Tract Left	9.18E-14	7.452236
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Superior Longitudinal Fasciculus III Right	3.23E-15	7.881733
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Frontal Superior Longitudinal Right	5.68E-16	8.096022
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	OD Inferior Fronto Occipital Fasciculus Left	4.68E-09	5.858164
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	MD Arcuate Anterior Segment Left	1.98E-08	-5.613436
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L3 Arcuate Posterior Segment Left	4.54E-08	-5.468626
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	MD Superior Longitudinal Fasciculus II Left	2.62E-08	-5.565105
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Arcuate Posterior Segment Left	1.32E-21	9.547923
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Fronto Insular tract5 Left	7.29E-09	5.784136
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Arcuate Long Segment Right	1.51E-21	9.534019
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	OD Superior Longitudinal Fasciculus III Right	9.70E-09	5.735852
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Inferior Longitudinal Right	6.78E-14	7.492089
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L3 Arcuate Long Segment Right	4.39E-12	-6.924090
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L1 Inferior Longitudinal Right	1.43E-08	-5.666635
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	MD Arcuate Posterior Segment Left	2.60E-12	-6.997674
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Arcuate Anterior Segment Left	1.30E-22	9.785637
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Superior Longitudinal Fasciculus I Right	2.48E-19	8.989747
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	OD Inferior Longitudinal Right	3.43E-08	-5.517864
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	MD Superior Longitudinal Fasciculus I Right	3.64E-08	-5.507713
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	MD Arcuate Long Segment Right	5.24E-13	-7.218870
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	OD Inferior Fronto Occipital Fasciculus Right	4.55E-09	5.862860
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Superior Longitudinal Fasciculus I Left	3.54E-19	8.950493
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L1 Arcuate Posterior Segment Right	6.63E-11	-6.528706
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L1 Superior Longitudinal Fasciculus III Left	4.09E-11	-6.600720
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L3 Superior Longitudinal Fasciculus II Right	6.99E-09	-5.791797
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L1 Corpus callosum	1.01E-08	-5.728680
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	MD Arcuate Long Segment Left	2.59E-12	-6.998517
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Uncinate Right	1.50E-10	6.404902
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Fronto Insular tract4 Left	5.70E-09	5.825317
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L3 Superior Longitudinal Fasciculus I Right	1.05E-10	-6.459763
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Uncinate Left	2.48E-08	5.574950
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Inferior Longitudinal Left	6.00E-12	6.879745
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Superior Longitudinal Fasciculus III Left	6.91E-19	8.876381
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L2 Arcuate Posterior Segment Right	6.12E-09	-5.813545
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L3 Arcuate Anterior Segment Right	2.57E-09	-5.957955
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	L2 Arcuate Posterior Segment Left	6.68E-10	-6.173398
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	ICVF Frontal Inferior longitudinal Right	9.06E-11	6.484797
128	rs74504435	7	54949256	intergenic	G	A	0.094	5.65E-41	1.26E-01	SNORA73	MD Superior Longitudinal Fasciculus I Left	2.17E-09	-5.984430
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	ICVF Arcuate Anterior Segment Right	3.13E-10	-6.292061
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	ICVF Arcuate Long Segment Left	2.10E-10	-6.354154
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	ICVF Corpus callosum	3.66E-09	-5.998645
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	ICVF Arcuate Posterior Segment Right	9.84E-09	-5.733533
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	ICVF Fronto Insular tract3 Left	1.36E-09	-6.060690
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	ICVF Frontal Superior Longitudinal Right	1.16E-08	-5.796071
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	ICVF Arcuate Posterior Segment Left	2.76E-09	-5.945157
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	ICVF Arcuate Long Segment Right	3.25E-09	-5.918395
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	ICVF Arcuate Anterior Segment Left	6.14E-11	-6.540215
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	ICVF Superior Longitudinal Fasciculus I Right	8.41E-09	-5.760040
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	ICVF Superior Longitudinal Fasciculus I Left	1.26E-09	-6.071950
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	L1 Superior Longitudinal Fasciculus III Left	2.33E-08	-5.585584
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	ICVF Uncinate Right	8.73E-10	-6.131089
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	ICVF Uncinate Left	4.85E-08	-5.456876
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	ICVF Inferior Longitudinal Left	1.45E-08	-5.667219
128	rs10244020	7	55108994	intronic	C	T	0.472	1.02E-10	1.83E-01	EGFR	ICVF Superior Longitudinal Fasciculus III Left	2.06E-09	-5.993106
129	rs10252734	7	81418566	intergenic	G	T	0.400	5.23E-12	2.30E-01	HGF	MO Uncinate Left	8.14E-09	-5.765552

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mygaswP (discovery -British-)	mygaswP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
131	rs697285	7	96200086	NaN	C	T	0.308	6.42E-37	3.87E-02	SHEM1	L1 Frontal Inferior longitudinal Left	5.40E-09	5.841953
131	rs697285	7	96200086	NaN	C	T	0.308	6.42E-37	3.87E-02	SHEM1	MO Frontal Inferior longitudinal Left	9.64E-16	8.931332
137	rs6973339	7	134388544	intergenic	T	G	0.358	2.90E-10	3.26E-01	ACO09276.4	L3 Fronto Insular tracts Right	2.85E-09	-5.934433
137	rs6973339	7	134388544	intergenic	T	G	0.358	2.90E-10	3.26E-01	ACO09276.4	L1 Fronto Insular tracts Left	2.83E-09	-5.941052
137	rs6973339	7	134388544	intergenic	T	G	0.358	2.90E-10	3.26E-01	ACO09276.4	MD Fronto Insular tracts Left	6.62E-11	-6.528975
137	rs6973339	7	134388544	intergenic	T	G	0.358	2.90E-10	3.26E-01	ACO09276.4	L2 Fronto Insular tracts Left	6.43E-10	-6.170634
137	rs6973339	7	134388544	intergenic	T	G	0.358	2.90E-10	3.26E-01	ACO09276.4	L3 Fronto Insular tracts Left	3.31E-11	-6.632168
137	rs6973339	7	134388544	intergenic	T	G	0.358	2.90E-10	3.26E-01	ACO09276.4	ISOVF Superior Longitudinal Fasciculus III Right	1.60E-10	-6.395215
137	rs6973339	7	134388544	intergenic	T	G	0.358	2.90E-10	3.26E-01	ACO09276.4	MD Fronto Insular tracts Right	9.34E-10	-6.120214
139	rs475995	7	150819169	intronic	T	C	0.305	3.86E-19	3.41E-01	AGAP3	OD Superior Longitudinal Fasciculus II Right	3.48E-09	5.907303
139	rs475995	7	150819169	intronic	T	C	0.305	3.86E-19	3.41E-01	AGAP3	FA Arcuate Anterior Segment Right	1.25E-11	-6.773807
139	rs475995	7	150819169	intronic	T	C	0.305	3.86E-19	3.41E-01	AGAP3	OD Corpus callosum	1.52E-08	5.659817
139	rs475995	7	150819169	intronic	T	C	0.305	3.86E-19	3.41E-01	AGAP3	FA Arcuate Long Segment Left	8.33E-12	-6.832737
139	rs475995	7	150819169	intronic	T	C	0.305	3.86E-19	3.41E-01	AGAP3	OD Arcuate Long Segment Left	2.58E-13	7.314647
139	rs475995	7	150819169	intronic	T	C	0.305	3.86E-19	3.41E-01	AGAP3	FA Arcuate Long Segment Right	3.73E-08	-5.502952
139	rs475995	7	150819169	intronic	T	C	0.305	3.86E-19	3.41E-01	AGAP3	OD Superior Longitudinal Fasciculus III Right	1.38E-12	7.088284
139	rs475995	7	150819169	intronic	T	C	0.305	3.86E-19	3.41E-01	AGAP3	OD Arcuate Long Segment Right	1.32E-09	6.064667
139	rs475995	7	150819169	intronic	T	C	0.305	3.86E-19	3.41E-01	AGAP3	OD Arcuate Anterior Segment Right	1.31E-12	7.093535
139	rs475995	7	150819169	intronic	T	C	0.305	3.86E-19	3.41E-01	AGAP3	OD Superior Longitudinal Fasciculus II Left	3.90E-08	5.541545
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	MD Inferior Longitudinal Left	2.04E-08	5.608267
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	ISOVF Superior Longitudinal Fasciculus II Right	9.86E-09	5.733136
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	L1 Frontal Superior Longitudinal Right	1.18E-12	7.180939
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	ISOVF Frontal Superior Longitudinal Left	1.93E-08	5.618468
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	L2 Inferior Longitudinal Left	9.54E-09	5.738766
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	ISOVF Arcuate Posterior Segment Left	2.35E-10	6.336557
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	MD Corpus callosum	3.67E-08	5.506024
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	L1 Frontal Superior Longitudinal Left	3.38E-09	5.912039
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	L1 Superior Longitudinal Fasciculus II Left	1.00E-10	6.466457
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	L2 Uncinate Right	1.24E-08	5.635566
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	ISOVF Superior Longitudinal Fasciculus II Left	5.65E-09	5.826932
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	L1 Superior Longitudinal Fasciculus II Right	3.67E-09	5.898425
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	MD Superior Longitudinal Fasciculus II Left	2.21E-08	5.597143
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	L2 Uncinate Left	4.41E-08	5.473045
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	L3 Uncinate Left	3.23E-08	5.528486
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	L1 Superior Longitudinal Fasciculus I Right	7.74E-13	7.165669
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	ISOVF Corpus callosum	9.95E-09	5.731514
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	L3 Uncinate Right	1.33E-09	6.063460
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	L2 Inferior Longitudinal Right	2.71E-08	5.558971
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	L2 Arcuate Anterior Segment Left	1.40E-08	5.673169
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	L1 Corpus callosum	4.68E-12	6.914902
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	MD Uncinate Left	4.22E-08	5.481263
140	rs4128399	7	151498679	intronic	C	T	0.160	8.29E-19	2.19E-03	PRKAG2	L1 Superior Longitudinal Fasciculus I Left	7.29E-14	7.482568
141	rs4075359	8	9487813	intronic	C	T	0.376	1.73E-13	7.90E-02	TNKS	L2 Fronto Insular tracts Right	5.10E-09	-5.843722
141	rs4075359	8	9487813	intronic	C	T	0.376	1.73E-13	7.90E-02	TNKS	MO Fronto Insular tracts Right	8.56E-09	5.757000
142	rs71548883	8	23101514	intronic	CT	C	0.384	8.31E-21	3.02E-01	CHMP7	OD Superior Longitudinal Fasciculus II Left	1.81E-10	6.376720
143	rs403185	8	26476721	intronic	G	A	0.408	1.84E-17	5.03E-02	DPYSL2	MO Fronto Insular tract3 Right	1.85E-10	6.373438
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ISOVF Inferior Fronto Occipital fasciculus Left	8.40E-10	6.137168
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Superior Longitudinal Fasciculus II Right	6.81E-13	7.183274
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Arcuate Anterior Segment Right	5.13E-12	6.902002
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	L2 Inferior Longitudinal Left	3.45E-08	-5.516780
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	MD Inferior Longitudinal Right	3.60E-08	-5.590247
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Arcuate Long Segment Left	1.97E-11	6.741319
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ISOVF Inferior Longitudinal Right	2.35E-11	6.682158
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Inferior Fronto Occipital fasciculus Right	1.65E-15	7.965027
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Corpus callosum	1.39E-13	7.397416
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Arcuate Posterior Segment Right	3.98E-15	7.855561
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ISOVF Inferior Fronto Occipital fasciculus Right	4.89E-14	7.534830
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Inferior Fronto Occipital fasciculus Left	3.88E-14	7.564957
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Frontal Aslant Tract Right	3.41E-09	5.910363
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Frontal Inferior longitudinal Left	1.04E-08	5.724078
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Superior Longitudinal Fasciculus II Left	2.53E-12	7.001664
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ISOVF Superior Longitudinal Fasciculus I Right	1.09E-09	6.095389
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Frontal Aslant Tract Left	1.24E-08	5.693650
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Superior Longitudinal Fasciculus III Right	2.04E-11	6.702926
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Frontal Superior Longitudinal Right	4.41E-08	5.473718
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Arcuate Posterior Segment Left	8.50E-14	7.462383
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Arcuate Long Segment Right	3.95E-16	8.140216
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Inferior Longitudinal Right	6.86E-20	9.129391
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Arcuate Anterior Segment Left	5.84E-09	5.811410
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ISOVF Corpus callosum	2.27E-14	7.634487
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Superior Longitudinal Fasciculus I Right	2.04E-13	7.346278
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	L3 Uncinate Right	6.41E-09	-5.805656
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ICVF Superior Longitudinal Fasciculus I Left	1.29E-14	7.706855
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20	2.26E-02	SNX31	ISOVF Arcuate Anterior Segment Right	6.45E-09	5.804698
146	rs2978101	8	101679224	intergenic	G	T	0.461	7.07E-20					

D.1. Genomic loci associated with heritable language anatomical connectivities using MOSTest

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mgwasP (discovery -British-)	mgwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	FA Inferior Frontal Occipital Fasciculus Right	8.73E-09	5.753744
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	FA Superior Longitudinal Fasciculus III Right	1.54E-12	7.071268
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ISOVF Inferior Frontal Occipital Fasciculus Left	1.19E-10	6.440385
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Superior Longitudinal Fasciculus II Right	2.62E-17	8.462566
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	FA Inferior Longitudinal	4.46E-10	6.237167
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	FA Inferior Frontal Occipital Fasciculus Left	2.04E-13	7.346312
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Arcuate Anterior Segment Right	8.43E-25	10.282754
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	MD Arcuate Long Segment Left	2.58E-10	-6.322388
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	MD Inferior Longitudinal Right	6.99E-14	-7.480123
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Arcuate Long Segment Left	4.74E-25	10.338034
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L3 Superior Longitudinal Fasciculus I Left	1.95E-10	-6.549447
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Frontal Insular tracts Right	1.56E-08	5.654623
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Frontal Insular tracts Right	1.86E-08	5.624266
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	FA Arcuate Anterior Segment Right	3.00E-10	6.289794
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	FA Superior Longitudinal Fasciculus I Left	8.94E-11	6.483948
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Inferior Frontal Occipital Fasciculus Right	1.20E-28	11.103688
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	FA Arcuate Long Segment Left	2.30E-11	6.697777
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	MD Inferior Frontal Occipital Fasciculus Right	5.61E-16	-8.091337
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Corpus callosum	1.81E-22	9.755122
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L3 Superior Longitudinal Fasciculus III Right	2.91E-11	-6.651113
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Arcuate Posterior Segment Right	6.60E-19	8.881521
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ISOVF Inferior Frontal Occipital Fasciculus Right	2.59E-08	5.670160
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	FA Frontal Inferior Longitudinal Right	2.59E-09	5.958300
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ISOVF Inferior Longitudinal Left	4.01E-08	5.490478
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L2 Inferior Frontal Occipital Fasciculus Right	5.49E-09	-5.834499
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L3 Arcuate Posterior Segment Right	1.32E-09	-6.064775
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Frontal Insular tracts Left	7.33E-11	6.513045
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	MD Superior Longitudinal Fasciculus III Left	4.72E-09	-5.856874
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Inferior Frontal Occipital Fasciculus Left	6.76E-26	10.523144
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Frontal Aslant Tract Right	2.46E-17	8.469847
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L1 Inferior Frontal Occipital Fasciculus Right	1.83E-12	-7.049957
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	MD Arcuate Posterior Segment Right	4.03E-10	-6.252764
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L3 Superior Longitudinal Fasciculus III Left	1.88E-10	-6.370729
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Frontal Inferior Longitudinal Left	2.70E-11	6.662809
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Superior Longitudinal Fasciculus II Left	5.85E-19	8.894826
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L3 Inferior Frontal Occipital Fasciculus Right	1.10E-16	-8.160504
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L3 Inferior Frontal Occipital Fasciculus Left	4.40E-14	-7.545910
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L3 Arcuate Long Segment Left	2.04E-15	-5.931885
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Frontal Aslant Tract Left	8.66E-18	8.599551
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Superior Longitudinal Fasciculus III Right	6.39E-21	9.383404
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Frontal Superior Longitudinal Right	5.64E-10	6.200257
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	MD Superior Longitudinal Fasciculus II Left	1.03E-08	-5.725007
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L2 Inferior Frontal Occipital Fasciculus Left	2.79E-11	-6.652768
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Arcuate Posterior Segment Left	1.13E-16	8.290203
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	FA Arcuate Long Segment Right	4.90E-12	6.908508
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Frontal Insular tracts Left	4.43E-11	6.888838
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Arcuate Long Segment Right	4.21E-25	10.349527
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L3 Inferior Longitudinal Left	8.86E-14	-7.456818
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L3 Corpus callosum	9.91E-09	-5.745014
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Inferior Longitudinal Right	2.01E-26	10.636774
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L3 Arcuate Long Segment Right	3.36E-17	-8.433295
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	FA Corpus callosum	5.78E-10	6.196290
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L1 Inferior Longitudinal Right	3.81E-11	-6.614288
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	MD Arcuate Posterior Segment Left	1.10E-08	-5.714354
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Arcuate Anterior Segment Left	4.76E-22	9.653247
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L3 Arcuate Anterior Segment Left	1.32E-08	-5.682805
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Superior Longitudinal Fasciculus I Right	2.64E-18	8.726104
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	FA Inferior Longitudinal Right	1.20E-09	6.080603
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	MD Arcuate Long Segment Right	1.36E-13	-7.400986
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Frontal Insular tracts Right	1.92E-08	5.618060
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	MD Inferior Longitudinal Right	1.79E-08	-5.630991
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Superior Longitudinal Fasciculus I Left	5.00E-18	8.653444
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L2 Arcuate Long Segment Right	2.33E-08	-5.581818
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L2 Inferior Longitudinal Right	7.59E-09	-5.777286
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L1 Superior Longitudinal Fasciculus III Left	4.48E-08	-5.470777
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L3 Superior Longitudinal Fasciculus II Right	3.08E-09	-5.927317
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	MD Arcuate Long Segment Left	3.98E-13	-7.256381
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Uncinate Right	4.19E-14	7.554971
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	L3 Superior Longitudinal Fasciculus I Right	4.87E-09	-5.951593
151	r55705857	8	130645692	ncRNA intronic	G	A	0.058	1.60E-24	8.28E-01	CDC26	ICVF Uncinate Left	9.97E-14	7.441269
154	r679038	9	22029080	ncRNA intronic	A	G	0.421	1.68E-08	1.62E-01	RP11-145E5.5;CDKN2B-AS1	ICVF Superior Longitudinal Fasciculus II Right	3.03E-08	5.539814
154	r679038	9	22029080	ncRNA intronic	A	G	0.421	1.68E-08	1.62E-01	RP11-145E5.5;CDKN2B-AS1	ICVF Inferior Frontal Occipital Fasciculus Right	7.19E-09	5.786518
154	r679038	9	22029080	ncRNA intronic	A	G	0.421	1.68E-08	1.62E-01	RP11-145E5.5;CDKN2B-AS1	ICVF Corpus callosum	2.63E-13	7.312065
154	r679038	9	22029080	ncRNA intronic	A	G	0.421	1.68E-08	1.62E-01	RP11-145E5.5;CDKN2B-AS1	ICVF Inferior Frontal Occipital Fasciculus Left	1.15E-08	5.707463
154	r679038	9	22029080	ncRNA intronic	A	G	0.421	1.68E-08	1.62E-01	RP11-145E5.5;CDKN2B-AS1	ICVF Frontal Aslant Tract Right	2.02E-08	5.610549
154	r679038	9	22029080	ncRNA intronic	A	G	0.421	1.68E-08	1.62E-01	RP11-145E5.5;CDKN2B-AS1	ICVF Superior Longitudinal Fasciculus II Left	1.42E-08	5.671400
154	r679038	9	22029080	ncRNA intronic	A	G	0.421	1.68E-08	1.62E-01	RP11-145E5.5;CDKN2B-AS1	ICVF Superior Longitudinal Fasciculus III Right	3.36E-09	5.912738
154	r679038	9	22029080	ncRNA intronic	A	G	0.421	1.68E-08	1.62E-01	RP11-145E5.5;CDKN2B-AS1	ICVF Inferior Longitudinal Right	2.16E-08	5.598799
154	r679038	9	22029080	ncRNA intronic	A	G	0.421	1.68E-08	1.62E-01	RP11-145E5.5;CDKN2B-AS1	ICVF Arcuate Anterior Segment Left	1.94E-08	5.616961
154	r679038	9	22029080	ncRNA intronic	A	G	0.421	1.68E-08	1.62E-01	RP11-145E5.5;CDKN2B-AS1	ICVF Superior Longitudinal Fasciculus I Right	6.17E-09	5.812016
154	r679038	9	22029080	ncRNA intronic	A								

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mgwasP (discovery - British-)	mgwasP (replication - non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
154	rs679038	9	22029080	ncRNA intronic	A	G	0.421	1.68E-08	1.62E-01	RP11-145E5.5;CDKN2B-AS1	ICVF Superior Longitudinal Fasciculus III Left	2.37E-08	5.582359
154	rs679038	9	22029080	ncRNA intronic	A	G	0.421	1.68E-08	1.62E-01	RP11-145E5.5;CDKN2B-AS1	ICVF Frontal Inferior longitudinal Right	1.70E-08	5.640448
158	rs4744186	9	95932546	intergenic	T	C	0.431	2.93E-11	5.89E-02	WNK2	MO Inferior Fronto Occipital fasciculus Right	4.07E-08	5.487279
160	rs919334	9	101637477	intergenic	G	T	0.313	1.75E-23	6.15E-01	RP11-92C4.3	MO Superior Longitudinal Fasciculus II Left	2.34E-09	5.972391
160	rs919334	9	101637477	intergenic	G	T	0.313	1.75E-23	6.15E-01	RP11-92C4.3	OD Superior Longitudinal Fasciculus III Left	4.47E-09	5.866534
160	rs919334	9	101637477	intergenic	G	T	0.313	1.75E-23	6.15E-01	RP11-92C4.3	FA Fronto Insular tract Right	2.12E-08	5.600782
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Frontal Aslant Tract Left	2.22E-11	-6.690899
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Inferior Fronto Occipital fasciculus Left	2.29E-13	-7.330749
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Superior Longitudinal Fasciculus II Right	2.19E-11	-6.692491
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Fronto Insular tract Left	2.54E-08	-5.570679
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Inferior Longitudinal Right	3.38E-13	-7.278310
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Superior Longitudinal Fasciculus III Left	1.15E-14	-7.721283
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	L1 Fronto Insular tracts Left	4.50E-08	-5.460908
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Arcuate Anterior Segment Left	4.20E-10	-6.246323
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	MD Fronto Insular tract Left	3.72E-09	-5.895977
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Arcuate Posterior Segment Left	1.08E-08	-5.718293
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Uncinate Left	6.12E-11	-6.540843
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Inferior Fronto Occipital fasciculus Right	5.58E-12	-6.889935
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Inferior Longitudinal Left	4.75E-08	-5.460336
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Frontal Aslant Tract Right	1.91E-12	-7.400552
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Arcuate Long Segment Left	1.42E-10	-6.441402
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Superior Longitudinal Fasciculus II Left	3.91E-11	-6.607297
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Superior Longitudinal Fasciculus I Left	1.77E-13	-7.365130
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Superior Longitudinal Fasciculus I Right	8.72E-14	-7.458979
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	L3 Fronto Insular tracts Left	9.30E-09	-5.743139
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Corpus callosum	5.37E-11	-6.580209
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Arcuate Anterior Segment Right	9.73E-12	-6.810417
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Superior Longitudinal Fasciculus III Right	1.04E-15	-8.022191
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Arcuate Long Segment Right	2.71E-10	-6.314615
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Arcuate Posterior Segment Right	1.23E-10	-6.438074
161	rs10980607	9	113638375	intronic	T	C	0.216	8.06E-12	5.46E-01	LPAR1	ISOVF Fronto Insular tract Right	4.86E-10	-6.225543
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Frontal Aslant Tract Left	1.34E-16	-8.268834
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L3 Fronto Insular tracts Right	1.82E-08	-5.628285
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Inferior Fronto Occipital fasciculus Left	1.27E-18	-8.808066
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Superior Longitudinal Fasciculus II Right	3.27E-17	-8.436584
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L1 Inferior Longitudinal Left	2.12E-08	-5.602080
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L3 Fronto Insular tract Right	3.15E-13	-7.287974
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	MO Corpus callosum	1.37E-09	-6.059327
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Frontal Inferior longitudinal Right	3.00E-11	-6.646326
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Fronto Insular tracts Left	8.13E-11	-6.499133
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Inferior Longitudinal Right	1.11E-19	-9.077440
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Superior Longitudinal Fasciculus III Left	3.72E-22	-9.678659
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L2 Fronto Insular tract Right	1.13E-08	-5.710529
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L1 Fronto Insular tracts Left	3.39E-20	-9.205697
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Arcuate Anterior Segment Left	3.27E-12	-6.965734
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	OD Corpus callosum	3.02E-08	5.540275
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	MD Fronto Insular tracts Left	2.68E-25	-10.392461
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Arcuate Posterior Segment Left	4.87E-17	-8.389845
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Fronto Insular tract Left	2.02E-09	-5.995977
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	MO Frontal Superior Longitudinal Left	3.02E-09	-5.930693
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	MD Uncinate Right	1.06E-10	-6.458466
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Uncinate Left	1.23E-10	-6.060903
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L3 Fronto Insular tracts Right	2.37E-09	-5.970439
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Frontal Superior Longitudinal Right	6.66E-10	-6.174029
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Inferior Fronto Occipital fasciculus Right	1.95E-16	-8.224850
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Inferior Longitudinal Left	5.03E-12	-6.904810
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Frontal Aslant Tract Right	9.01E-19	-8.846739
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L1 Frontal Superior Longitudinal Left	2.30E-09	-5.975233
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Arcuate Long Segment Left	2.85E-12	-6.984703
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L2 Uncinate Right	2.92E-08	-5.545925
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Superior Longitudinal Fasciculus II Left	4.35E-20	-9.179147
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L1 Uncinate Left	1.26E-13	-7.410279
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Superior Longitudinal Fasciculus I Left	4.63E-22	-9.656220
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Superior Longitudinal Fasciculus I Right	3.41E-18	-8.696898
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ICVF Fronto Insular tracts Left	5.92E-16	8.090975
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L2 Fronto Insular tracts Left	2.19E-23	-9.903924
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L3 Fronto Insular tracts Left	2.98E-24	-10.160383
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L2 Uncinate Left	1.53E-13	-7.384727
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L2 Fronto Insular tracts Right	6.39E-09	-5.806178
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L3 Uncinate Left	1.52E-14	-7.686016
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L1 Inferior Longitudinal Right	1.80E-08	-5.716389
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L1 Fronto Insular tract Right	1.03E-10	-6.139060
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	L1 Superior Longitudinal Fasciculus I Right	1.98E-08	-5.613986
161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Corpus callosum	2.81E-17	-8.454127
161	rs72748148	9											

D.1. Genomic loci associated with heritable language anatomical connectivities using MOSTest

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mgwasP (discovery -British-)	mgwasP (replication -non British-)	Nearst gene	Central endophenotypes	univariate P value	univariate Z
rs161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Fronto Insular tracts Right	1.25E-10	-6.288212
rs161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Fronto Insular tracts Right	1.95E-09	-6.001919
rs161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Superior Longitudinal Fasciculus III Right	1.11E-23	-10.030878
rs161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	MD Fronto Insular tracts Right	1.06E-12	-7.119105
rs161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	MD Superior Longitudinal Fasciculus I Left	6.60E-09	-5.860890
rs161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Fronto Insular tract4 Right	9.59E-12	-8.812479
rs161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Arcuate Long Segment Right	2.13E-17	-8.486320
rs161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	MD Uncinate Left	7.73E-17	-8.335274
rs161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Arcuate Posterior Segment Right	4.16E-22	-9.667040
rs161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	ISOVF Fronto Insular tracts Right	2.65E-17	-8.460913
rs161	rs72748148	9	113677241	intronic	A	G	0.214	1.93E-33	2.41E-02	LPAR1	MD Superior Longitudinal Fasciculus I Left	1.77E-09	-6.017669
rs162	g118957629 AAAG A	9	118957629	intronic	A	AAAG	0.424	1.55E-08	3.09E-02	PAPPA	OD Fronto Insular tracts Right	8.32E-09	-5.761927
rs162	g118957629 AAAG A	9	118957629	intronic	A	AAAG	0.424	1.55E-08	3.09E-02	PAPPA	OD Fronto Insular tracts Right	1.28E-08	-5.688354
rs162	g118957629 AAAG A	9	118957629	intronic	A	AAAG	0.424	1.55E-08	3.09E-02	PAPPA	FA Fronto Insular tract Right	1.99E-09	-5.998593
rs163	rs7030607	9	125454937	intronic	A	G	0.254	7.40E-01	7.35E-02	ASTN2	MD Arcuate Anterior Segment Right	4.44E-09	-5.865524
rs167	rs774455421	10	25041543	intergenic	G	GAGACAGTCAAGTC	0.352	2.01E-09	2.43E-01	ARHGAP21	L3 Fronto Insular tracts Right	2.93E-09	-5.935474
rs167	rs774455421	10	25041543	intergenic	G	GAGACAGTCAAGTC	0.352	2.01E-09	2.43E-01	ARHGAP21	ISOVF Superior Longitudinal Fasciculus III Left	5.63E-09	-5.827325
rs167	rs774455421	10	25041543	intergenic	G	GAGACAGTCAAGTC	0.352	2.01E-09	2.43E-01	ARHGAP21	L2 Fronto Insular tracts Right	2.36E-09	-5.971008
rs167	rs774455421	10	25041543	intergenic	G	GAGACAGTCAAGTC	0.352	2.01E-09	2.43E-01	ARHGAP21	MD Fronto Insular tracts Left	1.37E-11	-6.760876
rs167	rs774455421	10	25041543	intergenic	G	GAGACAGTCAAGTC	0.352	2.01E-09	2.43E-01	ARHGAP21	ISOVF Uncinate Left	2.04E-11	-6.702994
rs167	rs774455421	10	25041543	intergenic	G	GAGACAGTCAAGTC	0.352	2.01E-09	2.43E-01	ARHGAP21	MD Arcuate Anterior Segment Right	1.73E-09	-5.774205
rs167	rs774455421	10	25041543	intergenic	G	GAGACAGTCAAGTC	0.352	2.01E-09	2.43E-01	ARHGAP21	L3 Fronto Insular tracts Left	6.00E-12	-6.879663
rs167	rs774455421	10	25041543	intergenic	G	GAGACAGTCAAGTC	0.352	2.01E-09	2.43E-01	ARHGAP21	ISOVF Corpus callosum	4.83E-08	-5.457413
rs167	rs774455421	10	25041543	intergenic	G	GAGACAGTCAAGTC	0.352	2.01E-09	2.43E-01	ARHGAP21	ISOVF Superior Longitudinal Fasciculus III Right	2.30E-08	-5.879282
rs167	rs774455421	10	25041543	intergenic	G	GAGACAGTCAAGTC	0.352	2.01E-09	2.43E-01	ARHGAP21	ISOVF Fronto Insular tracts Right	2.98E-09	-5.932812
rs169	rs17537179	10	48442894	intronic	G	A	0.267	3.68E-01	3.84E-01	CDP10	MD Arcuate Anterior Segment Left	2.42E-11	-5.460610
rs170	rs122405234	10	105039048	intronic	T	C	0.249	3.28E-10	3.61E-01	VSTM4	L1 Frontal Aslant Tract Left	7.73E-09	-5.774205
rs172	rs1163084	10	105039048	intronic	C	T	0.456	2.15E-15	1.09E-01	INA	L2 Superior Longitudinal Fasciculus II Left	4.97E-10	-6.145560
rs172	rs1163084	10	105039048	intronic	C	T	0.456	2.15E-15	1.09E-01	INA	L2 Superior Longitudinal Fasciculus III Left	1.71E-08	-5.639309
rs172	rs1163084	10	105039048	intronic	C	T	0.456	2.15E-15	1.09E-01	INA	L1 Superior Longitudinal Fasciculus III Right	7.34E-09	-5.783010
rs172	rs1163084	10	105039048	intronic	C	T	0.456	2.15E-15	1.09E-01	INA	MD Superior Longitudinal Fasciculus II Left	4.71E-08	-5.450939
rs172	rs1163084	10	105039048	intronic	C	T	0.456	2.15E-15	1.09E-01	INA	L1 Frontal Inferior longitudinal Right	5.65E-12	-6.384618
rs172	rs1163084	10	105039048	intronic	C	T	0.456	2.15E-15	1.09E-01	INA	L2 Arcuate Long Segment Left	1.05E-08	-5.721264
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L3 Frontal Aslant Tract Right	2.52E-09	-5.960046
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	MD Superior Longitudinal Fasciculus II Right	1.82E-08	-5.628171
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L2 Frontal Inferior longitudinal Right	2.48E-08	-5.574953
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	MD Arcuate Anterior Segment Right	1.20E-10	-6.702925
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	ISOVF Frontal Inferior longitudinal Right	1.27E-08	-5.692023
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	MD Frontal Aslant Tract Right	3.58E-10	-6.274455
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L1 Fronto Insular tract4 Right	3.58E-10	-6.272755
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	ICVF Fronto Insular tracts Right	9.24E-09	-5.744131
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	MD Fronto Insular tracts Right	4.43E-09	-5.871715
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L3 Frontal Aslant Tract Left	1.17E-10	-6.395618
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	MD Inferior Fronto Occipital Fasciculus Right	3.64E-08	-5.977324
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L3 Superior Longitudinal Fasciculus III Right	1.78E-10	-6.379106
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	MD Fronto Insular tracts Right	4.45E-12	-6.922516
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	MD Superior Longitudinal Fasciculus III Left	2.99E-11	-6.651702
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L1 Fronto Insular tracts Right	5.15E-13	-7.287107
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	ISOVF Arcuate Long Segment Left	2.68E-08	-5.560588
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	MD Arcuate Posterior Segment Right	4.35E-08	-5.476952
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	MD Fronto Insular tracts Left	9.18E-10	-6.123049
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L2 Superior Longitudinal Fasciculus III Right	7.95E-13	-7.161967
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L3 Superior Longitudinal Fasciculus III Left	1.07E-09	-6.098268
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L3 Inferior Fronto Occipital Fasciculus Right	6.19E-08	-5.842444
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L2 Frontal Aslant Tract Left	3.90E-10	-6.257958
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L2 Superior Longitudinal Fasciculus III Left	2.87E-12	-6.983901
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L2 Arcuate Anterior Segment Right	3.01E-10	-6.298478
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L3 Inferior Fronto Occipital fasciculus Left	5.11E-10	-6.215735
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	MD Frontal Aslant Tract Right	7.84E-09	-5.871952
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L2 Fronto Insular tracts Right	1.11E-08	-5.485484
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L2 Superior Longitudinal Fasciculus II Right	2.15E-09	-5.985867
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L1 Superior Longitudinal Fasciculus III Right	2.07E-11	-6.700780
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L1 Fronto Insular tracts Right	1.09E-08	-5.716483
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L2 Fronto Insular tracts Right	2.24E-09	-5.979554
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L3 Arcuate Long Segment Right	2.36E-08	-5.819751
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	MD Fronto Insular tracts Right	4.67E-09	-5.915294
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L1 Frontal Inferior longitudinal Right	4.02E-10	-6.253362
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	MD Arcuate Long Segment Right	2.07E-09	-5.992299
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L2 Arcuate Long Segment Left	3.84E-09	-5.890733
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L2 Frontal Aslant Tract Right	5.56E-09	-5.779208
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L1 Fronto Insular tracts Right	1.41E-08	-5.485484
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	MD Arcuate Long Segment Left	5.65E-10	-6.202544
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	MD Superior Longitudinal Fasciculus III Right	1.26E-12	-7.098201
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L3 Frontal Inferior longitudinal Right	2.48E-09	-5.962479
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L2 Fronto Insular tracts Right	7.13E-09	-5.787877
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	MD Inferior Fronto Occipital Fasciculus Left	6.19E-10	-6.185444
rs173	rs11394128	10	105454037	intronic	TG	T	0.485	3.49E-09	3.10E-01	SH3PXD2A	L3 Arcuate Anterior Segment Right	3.00E-09	-5.748548
rs173	rs11191829	10	105614452	intronic	T	G	0.353	3.25E-13	1.84E-01	SH3PXD2A	L3 Frontal Aslant Tract Right	4.51E-08	-5.465700
rs173	rs11191829	10	105614452	intronic	T	G	0.353	3.25E-13	1.84E-01	SH3PXD2A	MD Fronto Insular tracts Right	4.35E-08	-5.475984
rs173	rs11191829	10	105614452										

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mvqwasP (discovery - British-)	mvqwasP (replication - non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
179	rs2071020	11	10478228	intronic	T	C	0.452	4.99E-19	7.47E-01	AMPD3	L1 Arcuate Long Segment Left	3.68E-13	-7.266913
179	rs2071020	11	10478228	intronic	T	C	0.452	4.99E-19	7.47E-01	AMPD3	L1 Superior Longitudinal Fasciculus II Right	1.94E-14	-7.654833
179	rs2071020	11	10478228	intronic	T	C	0.452	4.99E-19	7.47E-01	AMPD3	OD Inferior Fronto Occipital fasciculus Left	5.01E-10	6.218775
179	rs2071020	11	10478228	intronic	T	C	0.452	4.99E-19	7.47E-01	AMPD3	L1 Superior Longitudinal Fasciculus III Right	1.31E-10	-6.425599
179	rs2071020	11	10478228	intronic	T	C	0.452	4.99E-19	7.47E-01	AMPD3	MD Superior Longitudinal Fasciculus II Left	4.38E-08	-5.474896
179	rs2071020	11	10478228	intronic	T	C	0.452	4.99E-19	7.47E-01	AMPD3	L1 Frontal Aslant Tract Left	4.42E-09	-5.867518
179	rs2071020	11	10478228	intronic	T	C	0.452	4.99E-19	7.47E-01	AMPD3	OD Superior Longitudinal Fasciculus III Right	1.07E-21	9.569702
179	rs2071020	11	10478228	intronic	T	C	0.452	4.99E-19	7.47E-01	AMPD3	L1 Inferior Longitudinal Right	8.10E-09	-5.766260
179	rs2071020	11	10478228	intronic	T	C	0.452	4.99E-19	7.47E-01	AMPD3	MD Arcuate Long Segment Right	1.81E-08	-5.629699
179	rs2071020	11	10478228	intronic	T	C	0.452	4.99E-19	7.47E-01	AMPD3	L1 Arcuate Posterior Segment Right	1.88E-09	-6.008128
179	rs2071020	11	10478228	intronic	T	C	0.452	4.99E-19	7.47E-01	AMPD3	L1 Superior Longitudinal Fasciculus III Left	1.54E-08	-5.657168
179	rs2071020	11	10478228	intronic	T	C	0.452	4.99E-19	7.47E-01	AMPD3	L1 Frontal Aslant Tract Right	8.66E-10	-6.132242
179	rs2071020	11	10478228	intronic	T	C	0.452	4.99E-19	7.47E-01	AMPD3	L1 Corpus callosum	2.23E-13	-7.334509
179	rs2071020	11	10478228	intronic	T	C	0.452	4.99E-19	7.47E-01	AMPD3	OD Superior Longitudinal Fasciculus II Left	8.97E-11	6.483333
182	rs2585753	11	19895756	intronic	G	C	0.226	1.52E-11	4.18E-03	NAV2	OD Superior Longitudinal Fasciculus II Right	1.47E-08	5.66564
182	rs2585753	11	19895756	intronic	G	C	0.226	1.52E-11	4.18E-03	NAV2	OD Arcuate Long Segment Left	1.15E-09	6.087171
182	rs2585753	11	19895756	intronic	G	C	0.226	1.52E-11	4.18E-03	NAV2	L1 Arcuate Long Segment Left	2.80E-10	-6.309202
182	rs2585753	11	19895756	intronic	G	C	0.226	1.52E-11	4.18E-03	NAV2	L1 Superior Longitudinal Fasciculus II Right	2.86E-09	-5.939723
182	rs2585753	11	19895756	intronic	G	C	0.226	1.52E-11	4.18E-03	NAV2	OD Fronto Insular tract4 Right	2.65E-08	5.563294
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Inferior Longitudinal Left	3.00E-11	-6.646595
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Fronto Insular tract4 Right	1.29E-08	8.806952
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L3 Frontal Aslant Tract Right	8.73E-09	-5.753635
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Superior Longitudinal Fasciculus II Right	3.32E-09	-5.915153
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	OD Uncinate Left	1.66E-08	5.644237
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Fronto Insular tract3 Right	5.55E-16	8.098833
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Arcuate Anterior Segment Right	3.37E-09	-5.912384
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Frontal Superior Longitudinal Right	8.26E-09	-5.826968
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L2 Inferior Longitudinal Left	3.91E-11	-6.607318
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Inferior Longitudinal Right	1.00E-09	-6.108635
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Inferior Fronto Occipital fasciculus Left	2.56E-10	-6.32429
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Frontal Aslant Tract Right	8.18E-12	-6.835427
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L3 Superior Longitudinal Fasciculus I Left	4.54E-09	-5.862207
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Fronto Insular tract4 Right	2.27E-10	-6.341546
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Fronto Insular tract2 Right	5.67E-09	5.826119
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Fronto Insular tract3 Right	8.18E-19	8.875888
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Fronto Insular tract5 Right	3.72E-08	-5.503679
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L3 Frontal Aslant Tract Left	5.65E-09	-6.828885
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	OD Corpus callosum	1.79E-08	5.638054
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Inferior Fronto Occipital fasciculus Right	4.36E-08	5.475505
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Arcuate Anterior Segment Left	1.43E-09	-6.051881
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Uncinate Right	2.24E-20	-9.250391
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Inferior Fronto Occipital fasciculus Right	4.86E-09	-5.81793
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L3 Superior Longitudinal Fasciculus III Right	3.09E-08	-5.532638
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Arcuate Posterior Segment Right	9.76E-09	5.734837
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Corpus callosum	9.04E-12	-6.821042
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Fronto Insular tract3 Right	1.22E-09	-6.077554
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Superior Longitudinal Fasciculus III Left	5.43E-10	-6.206237
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Frontal Superior Longitudinal Left	1.56E-08	-5.65115
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Inferior Fronto Occipital fasciculus Left	2.19E-10	6.347123
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Superior Longitudinal Fasciculus II Left	5.04E-10	-6.217701
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Fronto Insular tract3 Right	4.62E-12	-6.916677
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Frontal Aslant Tract Right	3.16E-09	5.922854
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Arcuate Posterior Segment Left	5.21E-13	-7.219803
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Fronto Insular tract2 Left	1.44E-11	-6.754264
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Arcuate Posterior Segment Right	8.39E-09	-5.760334
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L2 Uncinate Right	1.64E-14	-7.675854
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L2 Corpus callosum	7.20E-10	-6.161658
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Uncinate Left	2.74E-21	-9.472294
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L2 Superior Longitudinal Fasciculus III Right	1.88E-11	-6.715219
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Uncinate Right	1.83E-21	-9.514327
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L2 Frontal Aslant Tract Left	5.24E-12	-6.899039
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L2 Superior Longitudinal Fasciculus III Left	3.61E-08	-5.508933
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L3 Inferior Fronto Occipital fasciculus Left	7.83E-12	-6.841568
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Arcuate Long Segment Left	1.80E-08	-5.630673
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Frontal Aslant Tract Left	1.84E-08	5.626633
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Frontal Aslant Tract Left	2.74E-13	-7.306380
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Superior Longitudinal Fasciculus II Right	3.19E-10	-6.289041
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Fronto Insular tract4 Right	1.60E-11	-6.738610
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Frontal Superior Longitudinal Right	3.69E-09	5.897623
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Arcuate Anterior Segment Left	4.63E-10	-6.231558
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Superior Longitudinal Fasciculus III Right	2.47E-13	-7.320760
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Superior Longitudinal Fasciculus II Left	1.76E-08	-5.634188
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Frontal Aslant Tract Left	1.49E-12	-7.075807
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L2 Inferior Fronto Occipital fasciculus Left	4.50E-10	-6.235694
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Arcuate Posterior Segment Left	2.88E-08	5.548351
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L2 Uncinate Left	1.33E-17	-8.541140
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Fronto Insular tract5 Left	4.61E-15	7.837180
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L3 Inferior Longitudinal Left	2.40E-09	-5.968362
183	rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L3 Uncinate Left	4.61E-16	-8.121341

D.1. Genomic loci associated with heritable language anatomical connectivities using MOSTest

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mvgsasP (discovery -British-)	mvgsasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
rs37935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Inferior Longitudinal Right	1.69E-09	6.025463	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Fronto Insular tracts Left	3.08E-11	6.642601	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L2 Superior Longitudinal Fasciculus I Left	2.18E-08	-5.597317	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Arcuate Posterior Segment Left	8.22E-11	-6.495578	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Arcuate Anterior Segment Left	6.55E-09	5.809554	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Superior Longitudinal Fasciculus I Right	1.25E-12	-7.100267	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Superior Longitudinal Fasciculus I Right	3.05E-12	-6.975488	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	OD Uncinate Right	1.41E-11	6.756473	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L3 Uncinate Right	9.80E-15	-7.74835	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Fronto Insular tracts Right	1.27E-08	5.690448	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L2 Frontal Aslant Tract Right	6.15E-10	-6.186447	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L2 Arcuate Long Segment Right	2.40E-08	-5.580369	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L2 Fronto Insular tract4 Right	4.54E-09	-5.863325	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L2 Inferior Longitudinal Right	3.28E-09	-5.916980	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Superior Longitudinal Fasciculus III Left	2.11E-11	-6.698349	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Frontal Aslant Tract Right	1.98E-11	-6.740627	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Corpus callosum	6.35E-16	-8.082476	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Arcuate Long Segment Left	4.19E-08	-5.482443	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Uncinate Right	2.31E-19	8.997533	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L3 Superior Longitudinal Fasciculus I Right	4.77E-08	-5.459506	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Uncinate Left	2.40E-20	9.242938	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Inferior Longitudinal Left	1.54E-10	6.193334	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Uncinate Left	1.06E-22	-9.80970	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L2 Arcuate Posterior Segment Right	1.65E-08	-5.645555	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Superior Longitudinal Fasciculus III Right	5.68E-12	-6.887526	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Inferior Fronto Occipital fasciculus Left	1.24E-13	-7.412484	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L2 Arcuate Posterior Segment Left	2.59E-08	-5.566814	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	MD Superior Longitudinal Fasciculus I Left	1.98E-12	-7.035870	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	L1 Superior Longitudinal Fasciculus I Left	4.95E-11	-6.572323	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Superior Longitudinal Fasciculus II Right	4.56E-09	-5.862409	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Arcuate Anterior Segment Right	8.88E-09	-5.750899	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Inferior Fronto Occipital fasciculus Right	3.87E-11	-6.608627	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Corpus callosum	5.59E-10	-6.201643	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Arcuate Posterior Segment Right	4.63E-08	-5.464835	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Inferior Fronto Occipital fasciculus Left	7.78E-09	-5.773175	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Superior Longitudinal Fasciculus II Left	2.95E-09	-5.934521	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Arcuate Long Segment Right	1.29E-09	-6.068043	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Inferior Longitudinal Right	3.38E-11	-6.628824	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Arcuate Anterior Segment Left	1.16E-10	-6.444383	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Superior Longitudinal Fasciculus I Right	3.80E-10	-6.261879	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Superior Longitudinal Fasciculus I Left	1.73E-11	-6.727402	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Uncinate Right	8.05E-09	-5.797369	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Inferior Longitudinal Left	8.25E-13	-7.155860	
rs7935166	11	44636547	intronic	A	G	0.133	4.43E-41	3.00E-01	CD82	ICVF Superior Longitudinal Fasciculus III Left	9.75E-09	-5.734947	
rs11020080	11	92578985	intronic	A	G	0.387	9.29E-12	7.11E-02	FAT3	MO Frontal Aslant Tract Left	1.19E-10	6.440155	
rs11020080	11	92578985	intronic	A	G	0.387	9.29E-12	7.11E-02	FAT3	MO Frontal Aslant Tract Right	3.32E-11	6.631689	
rs636406	11	114031092	intronic	G	C	0.358	2.80E-09	6.12E-01	ZBTB16	ICVF Fronto Insular tract4 Right	5.40E-09	-5.834258	
rs636406	11	114031092	intronic	G	C	0.358	2.80E-09	6.12E-01	ZBTB16	OD Corpus callosum	4.21E-10	-6.246028	
rs10841784	12	21441577	intronic	C	T	0.283	9.57E-09	5.77E-01	SLCO1A2	MO Frontal Aslant Tract Left	3.11E-08	5.534957	
rs7305286	12	28282047	intergen	G	A	0.307	3.24E-17	4.99E-04	CCDC91	OD Inferior Fronto Occipital fasciculus Right	1.90E-08	-5.620457	
rs10878070	12	56478002	intronic	C	A	0.425	4.56E-09	2.26E-01	ERBB3	ICVF Arcuate Posterior Segment Left	1.20E-08	-5.700107	
rs4479956	12	79493956	ncRNA intronic	A	G	0.307	3.57E-34	3.65E-01	RP11-359M6.1	FA Frontal Superior Longitudinal Right	1.97E-09	6.000399	
rs4479956	12	79493956	ncRNA intronic	A	G	0.307	3.57E-34	3.65E-01	RP11-359M6.1	MD Arcuate Anterior Segment Left	4.71E-08	-5.461094	
rs4479956	12	79493956	ncRNA intronic	A	G	0.307	3.57E-34	3.65E-01	RP11-359M6.1	L2 Arcuate Anterior Segment Left	2.59E-08	-5.566930	
rs4479956	12	79493956	ncRNA intronic	A	G	0.307	3.57E-34	3.65E-01	RP11-359M6.1	L3 Frontal Superior Longitudinal Right	1.09E-09	-6.095207	
rs77436242	12	99640159	exonic	C	G	0.048	3.03E-10	5.36E-01	ANKS1B	OD Arcuate Long Segment Right	1.03E-08	-5.726373	
rs12146713	12	106476805	intronic	C	T	0.096	7.00E-35	1.99E-01	NUAK1	MO Frontal Aslant Tract Left	2.42E-19	-8.992431	
rs12146713	12	106476805	intronic	C	T	0.096	7.00E-35	1.99E-01	NUAK1	FA Frontal Aslant Tract Left	4.92E-13	-7.227460	
rs12146713	12	106476805	intronic	C	T	0.096	7.00E-35	1.99E-01	NUAK1	OD Frontal Aslant Tract Right	1.66E-15	7.964608	
rs12146713	12	106476805	intronic	C	T	0.096	7.00E-35	1.99E-01	NUAK1	OD Frontal Aslant Tract Left	5.44E-18	8.643663	
rs12146713	12	106476805	intronic	C	T	0.096	7.00E-35	1.99E-01	NUAK1	MO Frontal Aslant Tract Right	2.09E-17	-8.488629	
rs12146713	12	106476805	intronic	C	T	0.096	7.00E-35	1.99E-01	NUAK1	FA Frontal Aslant Tract Right	6.60E-13	-7.187463	
rs12146713	12	106476805	intronic	C	T	0.096	7.00E-35	1.99E-01	NUAK1	FA Fronto Insular tracts Right	1.01E-08	-5.728880	
rs12146713	12	106476805	intronic	C	T	0.096	7.00E-35	1.99E-01	NUAK1	OD Fronto Insular tract2 Left	3.24E-09	5.910936	
rs536700128	12	107653485	intronic	GA	G	0.243	1.13E-13	3.57E-01	RP11-144F15.1:RFX4	ICVF Inferior Longitudinal Right	1.44E-08	-5.668403	
rs536700128	12	107653485	intronic	GA	G	0.243	1.13E-13	3.57E-01	RP11-144F15.1:RFX4	ICVF Inferior Longitudinal Left	6.48E-09	-5.803986	
rs1114035	12	109102983	intronic	A	T	0.438	1.01E-08	8.17E-01	CORX1C	MO Superior Longitudinal Fasciculus I Left	5.72E-11	-6.550892	
rs7137828	12	119328200	intronic	T	C	0.484	2.99E-29	6.03E-01	ATXN2	ISOVF Arcuate Anterior Segment Left	1.00E-08	-5.729885	
rs7137828	12	119328200	intronic	T	C	0.484	2.99E-29	6.03E-01	ATXN2	ISOVF Arcuate Anterior Segment Right	1.59E-08	-5.653234	
rs10219716	12	112503474	ncRNA exonic	A	T	0.085	1.60E-09	3.98E-01	NAA25:RP11-267L14.3	OD Inferior Fronto Occipital fasciculus Left	2.86E-09	5.939208	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ISOVF Inferior Fronto Occipital fasciculus Left	5.08E-13	7.223065	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ICVF Fronto Longitudinal Fasciculus II Right	4.43E-17	8.400980	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ICVF Arcuate Anterior Segment Right	4.09E-08	8.676232	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ICVF Arcuate Long Segment Left	1.02E-14	7.738411	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ISOVF Inferior Longitudinal Right	6.62E-15	7.791568	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	L3 Superior Longitudinal Fasciculus I Left	4.48E-10	-6.236474	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	FA Superior Longitudinal Fasciculus I Left	1.09E-09	6.095569	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ICVF Inferior Fronto Occipital fasciculus Right	8.70E-18	8.589655	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ICVF Corpus callosum	2.07E-17	8.489798	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ICVF Arcuate Posterior Segment Right	4.28E-13	7.246513	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ISOVF Inferior Fronto Occipital fasciculus Right	2.54E-15	7.91853	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ISOVF Inferior Longitudinal Left	1.03E-12	7.126878	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ICVF Inferior Fronto Occipital fasciculus Left	7.01E-16	8.070367	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ICVF Frontal Aslant Tract Right	6.38E-12	6.534641	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ICVF Superior Longitudinal Fasciculus II Left	1.31E-15	7.993908	
rs9739560	12	118461574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	L3 Inferior Fronto Occipital fasciculus Right	1.80E-08	-5.630211	
rs9739560	12	118461574	intronic	A	T	0.418							

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mvgwasP (discovery -British-)	mvgwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value Z	univariate Z
208	rs9739560	12	118261574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ICVF Superior Longitudinal Fasciculus I Right	7.26E-21	9.369839
208	rs9739560	12	118261574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ICVF Superior Longitudinal Fasciculus I Left	1.43E-21	9.540087
208	rs9739560	12	118261574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	FA Superior Longitudinal Fasciculus I Right	1.56E-08	5.644961
208	rs9739560	12	118261574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	L3 Superior Longitudinal Fasciculus I Right	8.81E-10	-6.129682
208	rs9739560	12	118261574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ICVF Inferior Longitudinal Left	3.4E-18	8.696055
208	rs9739560	12	118261574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ICVF Superior Longitudinal Fasciculus III Left	8.55E-12	6.828063
208	rs9739560	12	118261574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ISOVF Arcuate Long Segment Right	1.34E-08	5.681119
208	rs9739560	12	118261574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	L3 Inferior Longitudinal Right	3.32E-08	-5.235253
208	rs9739560	12	118261574	intronic	A	T	0.418	3.29E-17	1.33E-01	RFC5	ICVF Frontal Inferior longitudinal Right	7.77E-09	5.773470
215	rs11830776	13	111040681	intronic	A	G	0.277	3.42E-10	2.91E-03	COL4A2	L1 Fronto Insular tracts Right	9.24E-09	-5.744197
218	rs160459	14	59074136	intergenic	C	A	0.459	4.19E-31	5.26E-01	DACT1	OD Arcuate Posterior Segment Right	1.04E-10	6.466070
218	rs160459	14	59074136	intergenic	C	A	0.459	4.19E-31	5.26E-01	DACT1	MO Inferior Fronto Occipital fasciculus Right	5.96E-12	-6.880632
218	rs160459	14	59074136	intergenic	C	A	0.459	4.19E-31	5.26E-01	DACT1	OD Inferior Fronto Occipital fasciculus Left	4.44E-11	6.588698
218	rs160459	14	59074136	intergenic	C	A	0.459	4.19E-31	5.26E-01	DACT1	MO Inferior Fronto Occipital fasciculus Left	6.60E-13	-7.187519
218	rs160459	14	59074136	intergenic	C	A	0.459	4.19E-31	5.26E-01	DACT1	FA Arcuate Posterior Segment Left	3.32E-12	-6.904461
218	rs160459	14	59074136	intergenic	C	A	0.459	4.19E-31	5.26E-01	DACT1	OD Arcuate Posterior Segment Left	1.17E-13	7.419671
218	rs160459	14	59074136	intergenic	C	A	0.459	4.19E-31	5.26E-01	DACT1	FA Arcuate Posterior Segment Right	6.66E-10	-6.166933
222	rs714221	14	91910995	intergenic	A	G	0.443	4.74E-08	4.17E-01	SMEK1	MO Corpus callosum	1.18E-08	5.703206
224	rs35523244	14	100283856	intronic	T	C	0.382	2.60E-23	3.21E-01	EML1	FA Fronto Insular tracts Right	4.66E-08	5.463867
227	rs5812099	15	39633877	ncRNA intronic	AG	A	0.090	2.15E-171	2.63E-06	RP11-624L4.1	L2 Frontal Inferior longitudinal Right	5.86E-09	-5.825733
227	rs5812099	15	39633877	ncRNA intronic	AG	A	0.090	2.15E-171	2.63E-06	RP11-624L4.1	MO Arcuate Anterior Segment Right	1.25E-10	6.433193
227	rs5812099	15	39633877	ncRNA intronic	AG	A	0.090	2.15E-171	2.63E-06	RP11-624L4.1	L1 Arcuate Anterior Segment Left	6.52E-15	7.793374
227	rs5812099	15	39633877	ncRNA intronic	AG	A	0.090	2.15E-171	2.63E-06	RP11-624L4.1	L1 Superior Longitudinal Fasciculus II Left	4.70E-09	-5.875297
227	rs5812099	15	39633877	ncRNA intronic	AG	A	0.090	2.15E-171	2.63E-06	RP11-624L4.1	FA Arcuate Anterior Segment Left	2.83E-28	11.026795
227	rs5812099	15	39633877	ncRNA intronic	AG	A	0.090	2.15E-171	2.63E-06	RP11-624L4.1	OD Arcuate Anterior Segment Left	1.67E-67	-17.359497
227	rs5812099	15	39633877	ncRNA intronic	AG	A	0.090	2.15E-171	2.63E-06	RP11-624L4.1	OD Arcuate Anterior Segment Right	4.86E-14	-7.534763
227	rs5812099	15	39633877	ncRNA intronic	AG	A	0.090	2.15E-171	2.63E-06	RP11-624L4.1	L2 Arcuate Anterior Segment Left	2.05E-27	-10.847632
227	rs5812099	15	39633877	ncRNA intronic	AG	A	0.090	2.15E-171	2.63E-06	RP11-624L4.1	MO Arcuate Anterior Segment Left	5.17E-53	15.254999
229	rs12438742	15	61947280	ncRNA intronic	G	C	0.429	1.36E-08	2.57E-01	RP11-507B12.1	MD Fronto Insular tracts Left	1.29E-10	-6.870499
229	rs12438742	15	61947280	ncRNA intronic	G	C	0.429	1.36E-08	2.57E-01	RP11-507B12.1	ISOVF Superior Longitudinal Fasciculus II Left	8.12E-09	-5.766015
229	rs12438742	15	61947280	ncRNA intronic	G	C	0.429	1.36E-08	2.57E-01	RP11-507B12.1	ICVF Fronto Insular tracts Left	5.14E-09	5.843706
229	rs12438742	15	61947280	ncRNA intronic	G	C	0.429	1.36E-08	2.57E-01	RP11-507B12.1	L2 Fronto Insular tracts Left	3.05E-11	-6.644206
229	rs12438742	15	61947280	ncRNA intronic	G	C	0.429	1.36E-08	2.57E-01	RP11-507B12.1	L3 Fronto Insular tracts Left	3.29E-12	-6.964842
229	rs12438742	15	61947280	ncRNA intronic	G	C	0.429	1.36E-08	2.57E-01	RP11-507B12.1	L3 Uncinate Right	8.38E-08	-5.49774
233	rs2324218	15	85693016	intergenic	A	G	0.325	7.28E-30	1.84E-03	PDE8A	L1 Fronto Insular tracts Left	8.81E-11	6.486072
233	rs2324218	15	85693016	intergenic	A	G	0.325	7.28E-30	1.84E-03	PDE8A	ISOVF Arcuate Anterior Segment Left	9.15E-10	-6.123517
233	rs2324218	15	85693016	intergenic	A	G	0.325	7.28E-30	1.84E-03	PDE8A	MD Fronto Insular tracts Left	4.24E-11	6.595551
233	rs2324218	15	85693016	intergenic	A	G	0.325	7.28E-30	1.84E-03	PDE8A	ICVF Fronto Insular tracts Left	5.34E-12	-6.896238
233	rs2324218	15	85693016	intergenic	A	G	0.325	7.28E-30	1.84E-03	PDE8A	L2 Fronto Insular tracts Left	3.56E-10	6.272156
233	rs2324218	15	85693016	intergenic	A	G	0.325	7.28E-30	1.84E-03	PDE8A	L3 Fronto Insular tracts Left	1.51E-08	5.660958
235	rs6224	15	91423543	intronic	T	G	0.475	9.79E-16	6.63E-01	FURIN	OD Superior Longitudinal Fasciculus II Left	1.75E-09	6.019806
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Inferior Longitudinal Left	1.57E-12	-7.079952
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L3 Frontal Aslant Tract Right	1.29E-08	-5.688487
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Frontal Aslant Tract Left	4.19E-31	-11.598536
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Inferior Fronto Occipital fasciculus Left	1.89E-34	-12.240321
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Superior Longitudinal Fasciculus II Right	3.66E-19	-8.948962
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Inferior Longitudinal Left	1.19E-13	-7.417851
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L3 Fronto Insular tracts Right	6.92E-14	-7.489465
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Frontal Superior Longitudinal Right	4.77E-13	-7.231995
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Frontal Superior Longitudinal Left	3.75E-12	-6.946328
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L2 Inferior Longitudinal Left	2.18E-09	-5.983779
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Frontal Inferior longitudinal Right	6.34E-14	-7.500899
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Inferior Longitudinal Right	2.40E-09	-5.967835
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Inferior Fronto Occipital fasciculus Left	2.41E-08	-5.979727
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Fronto Insular tracts Left	8.52E-09	-5.757797
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Frontal Aslant Tract Right	1.32E-12	-7.321298
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Fronto Insular tracts Left	8.32E-09	-5.761852
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Inferior Longitudinal Right	9.00E-33	-11.922783
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Superior Longitudinal Fasciculus III Left	3.03E-36	-12.571383
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L2 Fronto Insular tracts Right	1.10E-09	-6.093798
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Fronto Insular tracts Left	4.81E-15	-7.813842
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Arcuate Anterior Segment Left	1.18E-27	-10.897984
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	FA Fronto Insular tracts Right	1.88E-11	6.714804
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Fronto Insular tracts Left	2.84E-20	-9.224957
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Arcuate Posterior Segment Left	7.14E-27	-10.73527
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Fronto Insular tracts Left	4.10E-23	-9.901495
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Arcuate Anterior Segment Left	2.78E-08	-5.554983
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Uncinate Right	4.20E-11	-6.599882
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Inferior Fronto Occipital fasciculus Left	2.01E-31	-11.661393
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Inferior Fronto Occipital fasciculus Left	3.99E-15	-7.855368
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Inferior Fronto Occipital fasciculus Right	1.38E-32	-11.868683
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Corpus callosum	2.31E-09	-5.974549
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Fronto Insular tracts Right	7.79E-10	-6.149131
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Inferior Longitudinal Left	1.09E-34	-12.284741
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L3 Fronto Insular tracts Right	2.21E-13	-7.335303
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Frontal Aslant Tract Right	5.80E-31	-11.570776
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Superior Longitudinal Fasciculus III Left	7.77E-09	-5.773475
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Frontal Superior Longitudinal Left	2.48E-17	-8.468582
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Arcuate Long Segment Left	1.07E-20	-9.328694
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.			

D.1. Genomic loci associated with heritable language anatomical connectivities using MOSTest

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mgvwasP (discovery -British-)	mgvwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L2 Fronto Insular tracts Left	3.45E-20	-9.203854
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Superior Longitudinal Fasciculus II Right	2.60E-08	-5.562406
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Fronto Insular tract4 Right	4.47E-13	-7.240599
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Frontal Superior Longitudinal Left	2.20E-10	-6.346476
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L2 Fronto Insular tract3 Right	4.65E-10	-6.230620
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Arcuate Anterior Segment Left	2.52E-08	-5.571724
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Uncinate Right	5.95E-13	-7.201692
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Superior Longitudinal Fasciculus III Right	1.52E-08	-5.658797
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Frontal Aslant Tract Left	1.09E-08	-5.716690
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L3 Fronto Insular tracts Left	4.67E-20	-9.174662
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L2 Uncinate Left	2.32E-15	-7.922768
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L2 Fronto Insular tract3 Left	2.10E-11	-6.692442
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L3 Inferior Longitudinal Left	1.51E-08	-5.660341
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Fronto Insular tract3 Left	1.82E-17	-9.504939
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L3 Uncinate Left	1.29E-05	-7.407374
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	FA Fronto Insular tract3 Right	3.13E-09	-5.924379
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Inferior Longitudinal Right	3.28E-10	-6.284919
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Fronto Insular tract2 Right	1.32E-13	-7.402395
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Superior Longitudinal Fasciculus I Right	4.04E-14	-7.559687
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Corpus callosum	8.44E-29	-11.135353
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Fronto Insular tract5 Right	4.21E-14	-7.554257
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Superior Longitudinal Fasciculus I Right	3.19E-11	-6.637685
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Fronto Insular tract1 Right	5.92E-10	-6.192632
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L3 Uncinate Right	1.17E-11	-6.783972
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Arcuate Anterior Segment Right	1.02E-26	-10.700133
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L2 Frontal Aslant Tract Right	1.17E-11	-6.783446
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Fronto Insular tract3 Right	9.48E-18	-8.344223
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L2 Fronto Insular tract4 Right	1.64E-14	-7.676052
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Fronto Insular tract3 Left	1.38E-08	-5.676036
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Fronto Insular tract3 Right	8.82E-18	-8.915556
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L2 Arcuate Anterior Segment Left	6.97E-09	-5.791741
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Superior Longitudinal Fasciculus III Left	3.50E-09	-5.906023
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Frontal Aslant Tract Right	5.33E-10	-6.209126
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Superior Longitudinal Fasciculus III Right	2.29E-35	-12.410405
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Fronto Insular tract2 Right	1.22E-15	-8.002872
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Corpus callosum	1.36E-15	-7.989204
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L3 Fronto Insular tract3 Right	5.44E-11	-6.558371
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Fronto Insular tract4 Right	8.80E-20	-9.102043
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Arcuate Long Segment Right	7.21E-19	-8.871556
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Uncinate Left	8.28E-17	-8.227213
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Arcuate Posterior Segment Right	9.54E-27	-10.705966
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L3 Fronto Insular tract3 Left	4.99E-11	-6.571171
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Frontal Superior Longitudinal Right	1.01E-10	-6.465592
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	ISOVF Fronto Insular tract2 Right	1.22E-19	-9.067351
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	MD Superior Longitudinal Fasciculus I Left	3.70E-11	-6.615631
240	rs1948948	16	51442679	intergenic	T	C	0.438	3.49E-53	5.03E-04	RP11-437L7.1	L1 Superior Longitudinal Fasciculus I Left	2.96E-16	-8.174842
240	16:51495932 CAA C		51495932	intergenic	C	CAA	0.362	1.21E-08	5.29E-02	RP11-437L7.1	ISOVF Inferior Longitudinal Right	3.34E-08	-5.522509
240	16:51495932 CAA C		51495932	intergenic	C	CAA	0.362	1.21E-08	5.29E-02	RP11-437L7.1	ISOVF Superior Longitudinal Fasciculus III Left	7.33E-08	-5.523200
240	16:51495932 CAA C		51495932	intergenic	C	CAA	0.362	1.21E-08	5.29E-02	RP11-437L7.1	ISOVF Arcuate Anterior Segment Left	3.09E-09	-5.788814
240	16:51495932 CAA C		51495932	intergenic	C	CAA	0.362	1.21E-08	5.29E-02	RP11-437L7.1	ISOVF Arcuate Posterior Segment Left	1.95E-08	-5.616352
242	rs55945372	16	86606321	intergenic	C	T	0.188	1.25E-09	4.79E-03	FOXO1	MD Uncinate Left	3.18E-08	-5.531323
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	MD Inferior Longitudinal Left	1.21E-09	-6.078831
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Frontal Aslant Tract Left	1.62E-13	-7.377171
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Inferior Fronto Occipital fasciculus Left	9.21E-15	-7.949760
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	OD Uncinate Left	2.81E-08	-5.552777
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Superior Longitudinal Fasciculus II Right	4.90E-12	-6.908374
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	L1 Inferior Longitudinal Left	2.71E-13	-7.308094
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	MO Frontal Aslant Tract Left	1.29E-08	-5.687231
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	MO Corpus callosum	8.75E-18	-8.589240
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	L2 Inferior Longitudinal Left	2.86E-09	-5.939563
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Frontal Inferior longitudinal Right	7.76E-13	-7.165233
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Inferior Longitudinal Right	6.23E-24	-10.088202
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Superior Longitudinal Fasciculus III Left	2.78E-24	-10.167005
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Arcuate Anterior Segment Left	5.91E-14	-7.599994
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	OD Corpus callosum	7.18E-17	-8.344223
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Arcuate Posterior Segment Left	1.10E-25	-10.477164
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Fronto Insular tract3 Left	7.82E-09	-5.772310
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Uncinate Left	2.87E-09	-5.938600
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Inferior Fronto Occipital fasciculus Right	2.46E-18	-8.733873
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Inferior Longitudinal Left	1.90E-25	-10.425342
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Frontal Aslant Tract Right	3.77E-12	-6.945392
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	OD Superior Longitudinal Fasciculus I Left	5.92E-14	-7.509811
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	L1 Inferior Fronto Occipital fasciculus Right	2.12E-10	-6.351509
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	MO Uncinate Left	7.01E-09	-5.790744
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Arcuate Long Segment Left	4.58E-21	-9.418393
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Superior Longitudinal Fasciculus II Left	4.79E-14	-7.537532
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	L1 Uncinate Left	1.25E-12	-7.099977
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Superior Longitudinal Fasciculus I Left	5.68E-12	-6.887392
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Superior Longitudinal Fasciculus I Right	7.15E-11	-6.517524
243	rs4843552	16	87233516	intergenic	A								

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mgwasP (discovery -British-)	mgwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Arcuate Anterior Segment Right	2.05E-19	9.010571
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	MO Fronto Insular tract5 Right	5.41E-11	6.559071
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Fronto Insular tract5 Right	1.01E-10	6.368767
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Superior Longitudinal Fasciculus III Right	6.60E-25	10.305070
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Arcuate Long Segment Right	3.35E-22	6.689214
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	ISOVF Arcuate Posterior Segment Right	1.32E-31	11.697289
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	MO Superior Longitudinal Fasciculus I Right	1.89E-09	6.007268
243	rs4843552	16	87233516	intergenic	A	G	0.422	6.18E-105	1.00E-05	C16orf95	L1 Superior Longitudinal Fasciculus III Left	2.04E-10	6.358565
244	rs111308992	16	89992613	intronic	T	G	0.131	1.17E-31	9.85E-04	TUBB3;TUBB3	ICVF Superior Longitudinal Fasciculus II Left	2.19E-08	-5.598418
244	rs111308992	16	89992613	intronic	T	G	0.131	1.17E-31	9.85E-04	TUBB3;TUBB3	FA Arcuate Long Segment Right	9.41E-09	-5.741011
244	rs111308992	16	89992613	intronic	T	G	0.131	1.17E-31	9.85E-04	TUBB3;TUBB3	FA Arcuate Posterior Segment Right	2.81E-09	-5.942300
244	rs111308992	16	89992613	intronic	T	G	0.131	1.17E-31	9.85E-04	TUBB3;TUBB3	L2 Arcuate Long Segment Right	3.00E-08	-5.541514
244	rs111308992	16	89992613	intronic	T	G	0.131	1.17E-31	9.85E-04	TUBB3;TUBB3	FA Superior Longitudinal Fasciculus II Left	1.85E-10	-6.373403
244	rs111308992	16	89992613	intronic	T	G	0.131	1.17E-31	9.85E-04	TUBB3;TUBB3	L2 Superior Longitudinal Fasciculus II Left	2.20E-08	-5.595579
245	rs2131704	17	1970898	intronic	T	C	0.425	8.01E-19	6.48E-03	SMG6	ISOVF Frontal Aslant Tract Left	4.82E-09	5.853128
245	rs2131704	17	1970898	intronic	T	C	0.425	8.01E-19	6.48E-03	SMG6	L2 Frontal Inferior longitudinal Left	9.88E-11	6.468839
245	rs2131704	17	1970898	intronic	T	C	0.425	8.01E-19	6.48E-03	SMG6	ISOVF Superior Longitudinal Fasciculus II Right	5.38E-11	6.561031
245	rs2131704	17	1970898	intronic	T	C	0.425	8.01E-19	6.48E-03	SMG6	ISOVF Superior Longitudinal Fasciculus III Left	6.57E-10	6.176070
245	rs2131704	17	1970898	intronic	T	C	0.425	8.01E-19	6.48E-03	SMG6	ISOVF Superior Longitudinal Fasciculus II Left	1.25E-12	7.100065
245	rs2131704	17	1970898	intronic	T	C	0.425	8.01E-19	6.48E-03	SMG6	ICVF Fronto Insular tract1 Left	1.52E-08	-5.959543
245	rs2131704	17	1970898	intronic	T	C	0.425	8.01E-19	6.48E-03	SMG6	L3 Fronto Insular tract1 Left	4.11E-09	-5.79842
245	rs2131704	17	1970898	intronic	T	C	0.425	8.01E-19	6.48E-03	SMG6	ISOVF Arcuate Anterior Segment Right	4.75E-10	6.226599
245	rs2131704	17	1970898	intronic	T	C	0.425	8.01E-19	6.48E-03	SMG6	L2 Arcuate Anterior Segment Left	8.83E-09	-5.751782
245	rs2131704	17	1970898	intronic	T	C	0.425	8.01E-19	6.48E-03	SMG6	ISOVF Superior Longitudinal Fasciculus III Right	1.82E-10	6.375862
245	rs11078805	17	2109109	intronic	A	T	0.371	1.11E-12	1.74E-01	SMG6	ISOVF Superior Longitudinal Fasciculus II Right	5.27E-08	-5.689668
245	rs11078805	17	2109109	intronic	A	T	0.371	1.11E-12	1.74E-01	SMG6	ISOVF Frontal Inferior longitudinal Right	2.02E-09	-5.996412
245	rs11078805	17	2109109	intronic	A	T	0.371	1.11E-12	1.74E-01	SMG6	ISOVF Superior Longitudinal Fasciculus II Left	7.39E-09	-5.781866
248	rs7210118	17	13047713	intergenic	T	C	0.414	1.54E-24	6.24E-02	RP11-488L1.1	MO Arcuate Anterior Segment Right	1.34E-08	-5.681293
248	rs7210118	17	13047713	intergenic	T	C	0.414	1.54E-24	6.24E-02	RP11-488L1.1	OD Fronto Insular tract5 Right	4.02E-08	-5.489397
248	rs7210118	17	13047713	intergenic	T	C	0.414	1.54E-24	6.24E-02	RP11-488L1.1	FA Fronto Insular tract5 Right	1.37E-08	-5.67719
250	rs4994353	17	28386280	intronic	A	T	0.443	9.63E-36	9.98E-02	EFCAB5	L2 Fronto Insular tract5 Right	1.86E-08	-5.624326
250	rs4994353	17	28386280	intronic	A	T	0.443	9.63E-36	9.98E-02	EFCAB5	OD Fronto Insular tract5 Right	8.64E-09	-5.755564
250	rs4994353	17	28386280	intronic	A	T	0.443	9.63E-36	9.98E-02	EFCAB5	MO Fronto Insular tract5 Right	4.12E-13	-7.251425
250	rs4994353	17	28386280	intronic	A	T	0.443	9.63E-36	9.98E-02	EFCAB5	MO Fronto Insular tract3 Right	1.14E-09	-6.088497
250	rs4994353	17	28386280	intronic	A	T	0.443	9.63E-36	9.98E-02	EFCAB5	L1 Fronto Insular tract3 Left	1.32E-09	-6.064151
250	rs4994353	17	28386280	intronic	A	T	0.443	9.63E-36	9.98E-02	EFCAB5	OD Fronto Insular tract3 Right	5.92E-11	-6.545823
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	MD Inferior Longitudinal Left	7.41E-09	-5.678224
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Fronto Insular tract4 Right	3.38E-16	-8.159079
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	FA Inferior Fronto Occipital fasciculus Right	2.93E-12	-7.032553
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ISOVF Inferior Fronto Occipital fasciculus Left	6.66E-12	-6.861728
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Superior Longitudinal Fasciculus II Right	8.44E-15	-7.76720
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	FA Inferior Longitudinal Left	1.88E-10	-6.370832
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	FA Inferior Fronto Occipital fasciculus Left	2.51E-11	-6.672679
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Arcuate Anterior Segment Right	5.95E-17	-8.382328
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Fronto Insular tract3 Right	1.02E-14	-7.736565
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	L2 Inferior Longitudinal Left	5.72E-10	-6.198027
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	MD Inferior Longitudinal Right	1.77E-10	-6.380803
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	L1 Inferior Fronto Occipital fasciculus Left	2.61E-11	-6.667175
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Arcuate Long Segment Left	1.20E-16	-8.283003
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ISOVF Inferior Longitudinal Right	1.45E-14	-7.691506
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Fronto Insular tract2 Right	4.98E-13	-7.225827
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Fronto Insular tract2 Right	1.57E-14	-7.681764
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	FA Superior Longitudinal Fasciculus I Left	3.18E-09	-5.921886
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Inferior Fronto Occipital fasciculus Right	1.65E-29	-11.279678
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	MD Uncinate Right	3.57E-10	-6.27107
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	MD Inferior Fronto Occipital fasciculus Right	1.44E-16	-8.261505
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Corpus callosum	8.45E-21	-9.353854
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Arcuate Posterior Segment Right	3.30E-14	-7.95894
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ISOVF Inferior Fronto Occipital fasciculus Right	6.95E-12	-6.958622
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	MD Corpus callosum	6.07E-09	-5.814762
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ISOVF Inferior Longitudinal Left	3.67E-13	-7.267374
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	L2 Inferior Fronto Occipital fasciculus Right	2.64E-13	-7.311410
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Fronto Insular tract3 Left	6.94E-12	-6.858874
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	MD Superior Longitudinal Fasciculus III Left	1.18E-09	-6.083137
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Inferior Fronto Occipital fasciculus Left	1.27E-29	-11.302757
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Frontal Aslant Tract Right	1.08E-17	-5.684941
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	L1 Inferior Fronto Occipital fasciculus Right	4.27E-09	-5.873306
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Frontal Superior Longitudinal Left	2.55E-08	-5.65652
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	L2 Uncinate Right	1.34E-08	-5.681353
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	L2 Corpus callosum	3.49E-09	-5.906570
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	L3 Superior Longitudinal Fasciculus III Left	8.75E-09	-5.753264
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Frontal Inferior longitudinal Left	3.46E-08	-5.51628
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Superior Longitudinal Fasciculus II Left	1.23E-14	-7.712840
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	L3 Inferior Fronto Occipital fasciculus Right	6.35E-18	-8.626054
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	L3 Inferior Fronto Occipital fasciculus Left	8.88E-17	-8.318906
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Frontal Aslant Tract Left	2.98E-18	-8.712070
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	MD Frontal Aslant Tract Left	2.93E-08	-5.545475
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Superior Longitudinal Fasciculus III Right	1.78E-18	-8.770265
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Frontal Superior Longitudinal Right	1.46E-09	-6.048875
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	L1 Superior Longitudinal Fasciculus III Right	4.70E-08	-5.462389
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	L2 Inferior Fronto Occipital fasciculus Left	1.02E-13	-7.438096
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Arcuate Posterior Segment Left	8.79E-18	-8.588773
252	rs9898671												

D.1. Genomic loci associated with heritable language anatomical connectivities using MOSTest

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mgwasP (discovery -British-)	mgwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	FA Corpus callosum	3.75E-09	-5.894754
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	MD Arcuate Posterior Segment Left	1.33E-08	5.682804
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Arcuate Anterior Segment Left	1.23E-13	-7.413912
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Superior Longitudinal Fasciculus I Right	1.10E-16	-8.293817
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	FA Inferior Longitudinal	1.13E-13	-7.424624
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	MD Arcuate Long Segment Right	1.82E-08	5.628614
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	L3 Uncinate Right	2.50E-10	6.326886
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Fronto Insular tracts Right	5.99E-11	-6.543898
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Superior Longitudinal Fasciculus I Left	1.41E-17	-8.533936
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	FA Superior Longitudinal Fasciculus I Right	2.64E-08	-5.563913
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	L2 Inferior Longitudinal Right	4.25E-11	6.594964
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	L1 Superior Longitudinal Fasciculus III Left	5.66E-09	5.826410
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	MD Arcuate Long Segment Left	1.54E-09	6.040081
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Uncinate Right	4.59E-19	-8.921781
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Uncinate Left	1.66E-18	-8.778083
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Inferior Longitudinal Left	1.95E-24	-10.201391
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Superior Longitudinal Fasciculus III Left	2.73E-17	-8.457517
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	L3 Inferior Longitudinal Right	3.50E-13	7.273592
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	MD Uncinate Left	1.26E-09	6.072453
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	MD Superior Longitudinal Fasciculus III Right	1.04E-08	5.383865
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	MD Inferior Fronto Occipital fasciculus Left	5.12E-17	-8.723900
252	rs9898671	17	42935321	intronic	A	G	0.161	1.30E-21	1.05E-01	EFTUD2	ICVF Frontal Inferior longitudinal Right	3.10E-13	-7.286234
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Superior Longitudinal Fasciculus II Left	4.93E-08	-5.453659
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	MD Inferior Longitudinal Left	1.14E-10	-6.446471
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Frontal Aslant Tract Right	4.27E-11	-6.594373
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L1 Arcuate Anterior Segment Right	8.22E-09	-5.763863
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	FA Superior Longitudinal Fasciculus III Right	1.58E-08	5.652183
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L2 Frontal Inferior longitudinal Right	5.05E-10	-6.217573
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L1 Inferior Longitudinal Left	4.65E-10	-6.230509
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Superior Longitudinal Fasciculus II Right	2.15E-16	8.213282
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Arcuate Anterior Segment Right	3.50E-19	8.951795
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	MD Arcuate Anterior Segment Right	1.83E-12	-7.046851
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	MD Inferior Longitudinal Right	1.04E-10	-6.614662
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L1 Inferior Fronto Occipital fasciculus Left	1.17E-12	-7.108515
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Arcuate Long Segment Left	1.77E-16	8.236456
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	MD Frontal Aslant Tract Right	3.44E-12	-6.953353
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Superior Longitudinal Fasciculus I Left	4.54E-12	-6.919325
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Fronto Insular tract2 Right	1.16E-08	5.705945
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	FA Superior Longitudinal Fasciculus I Left	4.11E-09	5.879756
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Frontal Aslant Tract Left	1.07E-08	-5.719615
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Inferior Fronto Occipital fasciculus Right	1.40E-18	8.797359
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	MD Inferior Fronto Occipital fasciculus Right	8.76E-14	-7.458318
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Corpus callosum	5.96E-19	8.892698
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Superior Longitudinal Fasciculus III Right	6.49E-12	-6.868485
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Arcuate Posterior Segment Right	6.51E-13	7.189411
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	MD Corpus callosum	3.51E-11	-6.623479
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L2 Inferior Fronto Occipital fasciculus Right	4.00E-10	-6.254042
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	MD Superior Longitudinal Fasciculus III Left	5.72E-12	-6.880349
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Inferior Fronto Occipital fasciculus Left	4.64E-20	9.172001
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Frontal Aslant Tract Right	1.31E-18	8.804800
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L1 Inferior Fronto Occipital fasciculus Right	9.66E-10	-6.114912
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Frontal Superior Longitudinal Left	4.94E-08	5.453383
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	MD Arcuate Posterior Segment Right	4.29E-08	-5.478418
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L2 Corpus callosum	6.41E-11	-6.533843
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L2 Superior Longitudinal Fasciculus III Right	7.65E-14	-7.476193
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Superior Longitudinal Fasciculus III Left	2.67E-09	-5.950795
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Frontal Inferior longitudinal Left	2.77E-12	6.988805
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Superior Longitudinal Fasciculus II Left	1.03E-15	8.022761
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Inferior Fronto Occipital fasciculus Right	8.39E-13	-7.154669
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L2 Frontal Aslant Tract Left	1.16E-11	-6.784984
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L2 Superior Longitudinal Fasciculus III Left	9.52E-11	-6.474453
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L2 Arcuate Anterior Segment Right	2.82E-10	-6.308267
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Inferior Fronto Occipital fasciculus Left	2.21E-13	-7.335474
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Arcuate Long Segment Left	2.35E-11	-6.682587
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Frontal Aslant Tract Left	1.84E-16	8.232224
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	MD Frontal Aslant Tract Left	2.82E-10	-6.308636
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Superior Longitudinal Fasciculus III Right	9.63E-20	9.093977
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Frontal Superior Longitudinal Right	8.32E-14	7.465089
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L1 Superior Longitudinal Fasciculus III Right	4.89E-14	-7.543736
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	MD Superior Longitudinal Fasciculus II Left	2.07E-08	-5.605913
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	FA Frontal Aslant Tract Right	3.48E-09	5.997332
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L2 Inferior Fronto Occipital fasciculus Left	1.91E-11	-6.712912
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Arcuate Posterior Segment Left	2.44E-12	7.006712
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	FA Arcuate Long Segment Right	1.01E-08	5.728553
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Arcuate Long Segment Right	5.98E-16	8.089460
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Inferior Longitudinal Left	3.01E-10	-6.298156
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Corpus callosum	5.65E-10	-6.199831
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Inferior Longitudinal Right	1.90E-18	8.762899
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Arcuate Long Segment Right	5.87E-09	-5.820370
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L2 Superior Longitudinal Fasciculus I Left	7.01E-10	-6.152815
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	MD Frontal Inferior longitudinal Right	2.14E-13	-7.339872
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L1 Inferior Longitudinal Right	1.12E-08	-6.088805
252	rs12936234	17	43777951	intronic	C	T	0.350	8.59E-26	4.2				

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mvgnwSP (discovery -British-)	mvgnwSP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
252	rs12936234	17	43177951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L1 Corpus callosum	8.29E-10	-6.139214
252	rs12936234	17	43177951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	MD Arcuate Long Segment Left	2.19E-11	-6.692511
252	rs12936234	17	43177951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Superior Longitudinal Fasciculus I Right	7.68E-10	-6.151304
252	rs12936234	17	43177951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Uncinate Left	1.32E-10	6.424898
252	rs12936234	17	43177951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Inferior Longitudinal Left	1.89E-19	9.019032
252	rs12936234	17	43177951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Superior Longitudinal Fasciculus III Left	7.04E-17	8.346312
252	rs12936234	17	43177951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Inferior Longitudinal Right	5.97E-11	-6.544370
252	rs12936234	17	43177951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L2 Arcuate Posterior Segment Right	9.53E-12	-6.813432
252	rs12936234	17	43177951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	MD Superior Longitudinal Fasciculus III Right	2.10E-14	-7.644257
252	rs12936234	17	43177951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Frontal inferior longitudinal Right	2.44E-10	-6.330988
252	rs12936234	17	43177951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	MD Inferior Fronto Occipital fasciculus Left	2.58E-15	-7.909738
252	rs12936234	17	43177951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L3 Arcuate Anterior Segment Right	5.57E-10	-6.202140
252	rs12936234	17	43177951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	L2 Arcuate Posterior Segment Left	9.41E-09	-5.741056
252	rs12936234	17	43177951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	ICVF Frontal Inferior longitudinal Right	4.68E-15	7.835126
252	rs12936234	17	43177951	intronic	C	T	0.350	8.59E-26	4.26E-01	NMT1	MD Superior Longitudinal Fasciculus I Left	5.42E-11	-6.558962
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Fronto Insular tract4 Right	4.77E-21	9.44109
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L3 Frontal Aslant Tract Right	6.77E-14	-7.492325
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L3 Fronto Insular tract5 Right	2.55E-10	-6.325852
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L1 Arcuate Anterior Segment Right	4.21E-12	-6.029800
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	FA Superior Longitudinal Fasciculus III Right	1.53E-12	7.071917
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	MD Superior Longitudinal Fasciculus II Right	6.00E-10	-6.190473
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L2 Frontal Inferior longitudinal Right	1.67E-16	-8.243386
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Superior Longitudinal Fasciculus II Right	6.59E-16	8.077825
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	FA Frontal Aslant Tract Left	2.51E-08	5.572442
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Arcuate Anterior Segment Right	1.92E-20	9.266726
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Fronto Insular tract3 Right	2.31E-15	7.932565
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	MD Arcuate Anterior Segment Right	9.24E-14	-7.451329
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L2 Fronto Insular tract5 Left	3.94E-08	-5.493468
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Arcuate Long Segment Left	2.87E-15	7.896446
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	MD Frontal Aslant Tract Right	6.58E-12	-6.866481
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L1 Fronto Insular tract4 Right	5.20E-09	-5.840796
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Fronto Insular tract2 Right	3.20E-12	6.988833
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Fronto Insular tract5 Right	1.33E-14	7.702506
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	MD Fronto Insular tract5 Right	3.72E-16	-8.147222
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L3 Frontal Aslant Tract Left	4.06E-10	-6.235884
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Inferior Fronto Occipital fasciculus Right	7.88E-09	5.717488
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L1 Arcuate Anterior Segment Left	4.71E-08	-5.461763
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Corpus callosum	3.47E-18	8.694963
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L3 Superior Longitudinal Fasciculus III Right	1.01E-15	-8.025881
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Arcuate Posterior Segment Right	2.97E-15	7.892327
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	FA Frontal Inferior longitudinal Right	1.29E-14	7.706437
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	MD Corpus callosum	4.72E-10	-6.228287
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	MD Fronto Insular tract3 Right	9.63E-12	-6.811997
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	OD Superior Longitudinal Fasciculus III Left	4.36E-08	5.475244
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L3 Fronto Insular tract4 Right	7.94E-11	-6.501814
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Fronto Insular tract3 Left	6.81E-09	5.795490
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	MD Superior Longitudinal Fasciculus III Left	7.03E-11	-6.520116
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Inferior Fronto Occipital fasciculus Left	1.29E-13	7.406704
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L1 Fronto Insular tract3 Right	4.30E-10	-6.242550
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Frontal Aslant Tract Right	1.47E-23	10.003880
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Frontal Superior Longitudinal Left	1.03E-13	7.439933
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L3 Fronto Insular tract5 Left	5.39E-09	-5.934753
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	FA Fronto Insular tract5 Left	3.70E-08	5.504426
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	OD Arcuate Long Segment Left	6.27E-11	6.537106
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L2 Corpus callosum	2.63E-10	-6.139340
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L2 Superior Longitudinal Fasciculus III Right	1.73E-14	-7.661077
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L3 Superior Longitudinal Fasciculus III Left	1.47E-11	-6.751075
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Frontal Inferior longitudinal Left	1.36E-10	6.420523
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Superior Longitudinal Fasciculus II Left	8.57E-14	7.461195
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L2 Frontal Aslant Tract Left	4.92E-09	-5.840972
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L2 Fronto Insular tract3 Right	5.41E-16	-8.101933
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	FA Arcuate Anterior Segment Left	5.30E-10	6.209936
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L2 Arcuate Anterior Segment Right	6.25E-09	-5.810029
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L1 Arcuate Long Segment Left	6.66E-12	-6.864748
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Frontal Aslant Tract Left	6.76E-20	9.131427
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	MD Frontal Aslant Tract Left	1.25E-09	-6.073605
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	MD Fronto Insular tract4 Right	9.29E-15	-7.748572
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L2 Fronto Insular tract3 Right	6.38E-11	-6.534633
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Superior Longitudinal Fasciculus III Right	2.06E-24	10.196471
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L2 Superior Longitudinal Fasciculus II Right	1.60E-10	-6.395434
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Frontal Superior Longitudinal Right	3.13E-16	8.168255
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	MD Arcuate Anterior Segment Left	2.36E-12	-7.011337
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L3 Arcuate Posterior Segment Left	2.24E-08	-5.592107
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L1 Superior Longitudinal Fasciculus III Right	4.59E-12	-6.917627
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	FA Frontal Aslant Tract Right	2.63E-09	5.953340
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Arcuate Posterior Segment Left	5.64E-18	8.639585
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Fronto Insular tract5 Left	1.55E-16	8.252289
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Arcuate Long Segment Right	5.31E-16	8.104108
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L3 Corpus callosum	8.84E-10	-6.129007
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-				

D.1. Genomic loci associated with heritable language anatomical connectivities using MOSTest

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mygwasP (discovery -British-)	mygwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Superior Longitudinal Fasciculus I Left	1.89E-16	8.228704
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L2 Frontal Aslant Tract Right	9.00E-09	-5.748508
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L2 Fronto Insular tract4 Right	7.75E-16	-8.058115
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L2 Arcuate Anterior Segment Left	3.60E-09	-5.901051
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L1 Superior Longitudinal Fasciculus III Left	3.98E-12	-6.938008
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L3 Superior Longitudinal Fasciculus II Right	6.49E-09	-5.803720
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L1 Frontal Aslant Tract Right	3.32E-08	-5.523667
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	MD Arcuate Long Segment Left	3.01E-09	-5.901251
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L3 Superior Longitudinal Fasciculus I Right	1.63E-09	-6.030634
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Uncinate Left	9.39E-11	6.470456
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Inferior Longitudinal Left	4.52E-08	5.469307
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Superior Longitudinal Fasciculus III Left	3.00E-20	9.218880
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	MD Superior Longitudinal Fasciculus III Right	1.14E-15	-8.010875
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L3 Frontal Inferior longitudinal Right	7.77E-16	-8.057838
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L1 Fronto Insular tract5 Right	6.62E-13	-7.186988
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L3 Arcuate Anterior Segment Right	3.45E-10	-6.277266
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	L2 Arcuate Posterior Segment Left	1.48E-09	-6.046922
252	rs113448888	17	44108482	UTR3	AGCCCTCT	A	0.224	2.09E-100	9.42E-05	KANSL1	ICVF Frontal Inferior longitudinal Right	3.48E-24	10.145278
255	rs62086046	17	62984671	intergenic	C	G	0.081	3.35E-10	1.56E-01	AMZ2p1	MO Corpus callosum	3.13E-08	5.533045
255	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L3 Superior Longitudinal Fasciculus II Left	1.85E-08	-5.625138
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	MD Inferior Longitudinal Left	9.40E-09	-5.741251
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Fronto Insular tract4 Right	1.18E-10	6.442344
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	FA Superior Longitudinal Fasciculus III Right	6.31E-13	7.193666
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	MD Superior Longitudinal Fasciculus II Right	2.29E-08	-5.588580
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Superior Longitudinal Fasciculus II Right	2.34E-09	5.972586
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Arcuate Anterior Segment Right	1.06E-11	6.798595
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	MD Arcuate Anterior Segment Right	2.03E-11	-6.703756
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L2 Inferior Longitudinal Left	9.09E-09	-5.746980
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Arcuate Long Segment Left	4.25E-10	6.244595
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L3 Superior Longitudinal Fasciculus I Left	3.19E-08	-5.539726
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	FA Arcuate Anterior Segment Right	1.95E-10	6.365264
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	FA Superior Longitudinal Fasciculus I Left	2.01E-08	5.611010
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Inferior Fronto Occipital fasciculus Right	2.99E-08	5.542209
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Corpus callosum	2.99E-09	5.932721
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L3 Superior Longitudinal Fasciculus III Right	2.43E-14	-7.625566
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Arcuate Posterior Segment Right	3.41E-11	6.627475
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	FA Frontal Inferior longitudinal Right	5.36E-09	5.835643
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L3 Arcuate Posterior Segment Right	6.48E-10	-6.178688
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Fronto Insular tract3 Left	2.70E-10	6.314979
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	FA Superior Longitudinal Fasciculus III Left	8.29E-10	6.139187
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	MD Superior Longitudinal Fasciculus III Left	4.85E-12	-6.909955
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Inferior Fronto Occipital Fasciculus Left	1.88E-09	6.009659
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Frontal Aslant Tract Right	1.14E-08	5.708940
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L1 Arcuate Posterior Segment Left	8.74E-09	-5.753013
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	MD Arcuate Posterior Segment Right	2.43E-10	-6.331325
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L1 Uncinate Left	4.13E-09	-5.878865
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L2 Superior Longitudinal Fasciculus III Right	6.92E-13	-7.181209
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L3 Superior Longitudinal Fasciculus III Left	1.00E-11	-6.805963
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L1 Uncinate Right	5.78E-10	-6.196222
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Superior Longitudinal Fasciculus II Left	4.58E-09	5.861682
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L2 Frontal Aslant Tract Left	2.18E-08	-5.596940
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L2 Superior Longitudinal Fasciculus III Left	1.49E-10	-6.406449
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L2 Arcuate Anterior Segment Right	5.11E-10	-6.215729
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L3 Arcuate Long Segment Left	2.60E-12	-6.997996
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Frontal Aslant Tract Left	8.70E-10	6.131959
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	MD Frontal Aslant Tract Left	1.24E-08	-5.693994
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Superior Longitudinal Fasciculus III Right	6.66E-14	7.494323
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L3 Arcuate Posterior Segment Left	9.13E-12	-6.816616
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L1 Superior Longitudinal Fasciculus III Right	3.81E-10	-6.261736
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	MD Superior Longitudinal Fasciculus II Left	1.40E-09	-6.055503
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Arcuate Posterior Segment Left	1.39E-11	6.759462
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Arcuate Long Segment Right	4.14E-12	6.932214
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L3 Inferior Longitudinal Left	6.14E-09	-5.813009
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L3 Uncinate Left	3.02E-09	-5.930460
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Inferior Longitudinal Right	4.75E-08	5.460561
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L3 Arcuate Long Segment Right	6.54E-12	-6.867298
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L2 Superior Longitudinal Fasciculus I Left	2.59E-10	-6.321685
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	MD Frontal Inferior longitudinal Right	1.50E-09	-6.044406
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	MD Arcuate Posterior Segment Left	3.49E-12	-6.956464
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Superior Longitudinal Fasciculus I Right	4.66E-08	5.464090
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	MD Arcuate Long Segment Right	7.99E-11	-6.500809
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Superior Longitudinal Fasciculus I Left	3.75E-08	5.502016
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L2 Arcuate Anterior Segment Left	1.38E-09	-6.057307
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L1 Superior Longitudinal Fasciculus III Left	1.20E-08	-5.699958
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L3 Superior Longitudinal Fasciculus II Right	3.90E-08	-5.495300
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	MD Arcuate Long Segment Left	1.25E-10	-6.433398
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Uncinate Right	1.73E-08	5.037310
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Uncinate Left	1.33E-10	6.424104
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Inferior Longitudinal Left	2.20E-08	5.590897
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Superior Longitudinal Fasciculus III Left	8.32E-11	6.494637
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	MD Uncinate Left	4.39E-09	-5.872214
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	MD Superior Longitudinal Fasc		

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mgvwasP (discovery -British-)	mgvwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	MD Inferior Fronto Occipital fasciculus Left	4.15E-08	-5.484500
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	L3 Arcuate Anterior Segment Right	2.85E-12	-6.985743
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	ICVF Frontal inferior longitudinal Right	2.66E-10	6.327254
256	rs62086046	17	65982943	intergenic	T	C	0.204	8.58E-19	3.20E-02	BPTF	MD Superior Longitudinal Fasciculus Left	2.67E-08	-5.561061
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	ICVF Arcuate Anterior Segment Right	8.96E-11	6.483498
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	ICVF Arcuate Long Segment Left	2.48E-09	5.962580
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	L3 Superior Longitudinal Fasciculus III Right	5.97E-09	-5.817535
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	ICVF Inferior Fronto Occipital fasciculus Left	2.41E-09	5.967644
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	ICVF Frontal Aslant Tract Right	3.76E-09	5.894606
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	L2 Superior Longitudinal Fasciculus III Right	3.31E-08	-5.524297
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	ICVF Frontal inferior longitudinal Left	1.49E-08	5.662640
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	ICVF Frontal Aslant Tract Left	1.07E-09	6.098347
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	ICVF Superior Longitudinal Fasciculus III Right	1.58E-13	7.380429
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	ICVF Arcuate Posterior Segment Left	3.61E-08	-5.508947
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	MD Frontal Inferior longitudinal Right	3.69E-08	-5.505205
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	ICVF Arcuate Anterior Segment Left	3.49E-10	6.275271
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	ICVF Fronto Insular tracts Right	1.21E-08	5.698591
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	ICVF Uncinate Left	4.66E-08	5.403951
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	ICVF Superior Longitudinal Fasciculus III Left	2.46E-12	7.005540
259	rs770068798	17	73866798	intergenic	G	GT	0.184	1.36E-13	4.55E-02	TRIM47	ICVF Frontal inferior longitudinal Right	3.80E-12	6.044276
260	rs9066170	18	35025056	intronic	G	A	0.360	1.24E-28	5.40E-01	CELF4	ISOVF Inferior Fronto Occipital fasciculus Right	3.01E-09	-5.932166
264	rs62127993	19	3298646	ncRNA intronic	A	G	0.135	2.58E-10	3.58E-01	AC010649.1	L1 Arcuate Anterior Segment Right	6.42E-09	5.805445
264	rs62127993	19	3298646	ncRNA intronic	A	G	0.135	2.58E-10	3.58E-01	AC010649.1	OD Arcuate Anterior Segment Right	2.49E-09	-5.961935
265	rs7256295	19	4071814	intergenic	A	G	0.223	6.17E-09	9.02E-01	ZBTB7A	L2 Uncinate Left	2.39E-09	5.969056
265	rs7256295	19	4071814	intergenic	A	G	0.223	6.17E-09	9.02E-01	ZBTB7A	L3 Uncinate Left	4.90E-08	-5.454833
265	rs7256295	19	4071814	intergenic	A	G	0.223	6.17E-09	9.02E-01	ZBTB7A	L3 Uncinate Right	1.49E-08	5.662840
267	rs10403210	19	13103638	intergenic	C	T	0.466	4.69E-08	4.17E-02	NFIX	MO Superior Longitudinal Fasciculus I Left	4.92E-08	-5.454230
271	rs429358	19	4541194	exonic	C	T	0.156	1.20E-16	5.37E-02	APOE	MD Inferior Longitudinal Left	1.25E-10	-6.433351
271	rs429358	19	4541194	exonic	C	T	0.156	1.20E-16	5.37E-02	APOE	L1 Inferior Longitudinal Left	6.70E-09	-5.798157
271	rs429358	19	4541194	exonic	C	T	0.156	1.20E-16	5.37E-02	APOE	L2 Inferior Longitudinal Left	8.48E-10	-6.135756
271	rs429358	19	4541194	exonic	C	T	0.156	1.20E-16	5.37E-02	APOE	MD Inferior Longitudinal Right	1.15E-09	-6.087099
271	rs429358	19	4541194	exonic	C	T	0.156	1.20E-16	5.37E-02	APOE	ISOVF Inferior Longitudinal Right	4.27E-08	-5.479396
271	rs429358	19	4541194	exonic	C	T	0.156	1.20E-16	5.37E-02	APOE	ISOVF Inferior Longitudinal Left	1.04E-08	-5.723498
271	rs429358	19	4541194	exonic	C	T	0.156	1.20E-16	5.37E-02	APOE	MD Arcuate Posterior Segment Right	3.10E-08	-5.535967
271	rs429358	19	4541194	exonic	C	T	0.156	1.20E-16	5.37E-02	APOE	L3 Inferior Longitudinal Left	2.37E-09	-5.970221
271	rs429358	19	4541194	exonic	C	T	0.156	1.20E-16	5.37E-02	APOE	L2 Inferior Longitudinal Right	5.25E-10	-6.208393
271	rs429358	19	4541194	exonic	C	T	0.156	1.20E-16	5.37E-02	APOE	L1 Arcuate Posterior Segment Right	4.11E-08	-5.486218
271	rs429358	19	4541194	exonic	C	T	0.156	1.20E-16	5.37E-02	APOE	L3 Inferior Longitudinal Right	2.23E-09	-5.970769
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	L3 Superior Longitudinal Fasciculus II Left	1.84E-09	-6.011205
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	ICVF Superior Longitudinal Fasciculus II Right	1.55E-10	6.400411
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	ICVF Arcuate Anterior Segment Right	3.18E-10	6.289534
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	ICVF Arcuate Long Segment Left	3.34E-11	6.630756
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	L3 Superior Longitudinal Fasciculus I Left	4.35E-08	-5.476129
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	FA Superior Longitudinal Fasciculus II Left	1.43E-10	6.412748
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	FA Arcuate Long Segment Left	2.05E-08	5.607661
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	FA Superior Longitudinal Fasciculus II Right	2.86E-08	5.549502
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	ICVF Superior Longitudinal Fasciculus II Left	2.60E-11	6.667147
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	L3 Arcuate Long Segment Left	1.54E-09	-6.040641
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	L2 Superior Longitudinal Fasciculus II Right	4.50E-08	-5.469894
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	MD Superior Longitudinal Fasciculus II Right	2.06E-09	-5.993312
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	FA Arcuate Long Segment Right	2.48E-08	5.596899
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	ICVF Arcuate Long Segment Right	5.22E-09	5.339391
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	L3 Arcuate Long Segment Right	2.79E-08	-5.554399
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	ICVF Superior Longitudinal Fasciculus I Right	2.94E-09	5.935183
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	L2 Arcuate Long Segment Left	4.46E-08	-5.471526
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	ICVF Superior Longitudinal Fasciculus I Left	3.43E-10	6.277932
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	FA Superior Longitudinal Fasciculus I Right	2.16E-10	6.349478
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	L3 Superior Longitudinal Fasciculus II Right	1.09E-08	-5.715910
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	MD Arcuate Long Segment Left	1.92E-09	-6.004597
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	L3 Superior Longitudinal Fasciculus I Right	1.51E-08	-5.660032
277	rs202393	20	45195738	intronic	C	T	0.160	1.77E-12	4.82E-01	SLC13A3	L2 Superior Longitudinal Fasciculus II Left	2.07E-09	-5.992548
277	rs2694896	20	4523743	intronic	G	A	0.385	4.95E-08	3.64E-01	SLC13A3	ISOVF Frontal Aslant Tract Left	6.36E-09	5.807953
277	rs2694896	20	4523743	intronic	G	A	0.385	4.95E-08	3.64E-01	SLC13A3	ISOVF Inferior Fronto Occipital fasciculus Left	2.56E-08	5.691209
277	rs2694896	20	4523743	intronic	G	A	0.385	4.95E-08	3.64E-01	SLC13A3	ISOVF Superior Longitudinal Fasciculus II Right	5.20E-13	7.219841
277	rs2694896	20	4523743	intronic	G	A	0.385	4.95E-08	3.64E-01	SLC13A3	ISOVF Superior Longitudinal Fasciculus III Left	5.22E-15	7.821460
277	rs2694896	20	4523743	intronic	G	A	0.385	4.95E-08	3.64E-01	SLC13A3	ISOVF Arcuate Anterior Segment Left	5.74E-15	7.809432
277	rs2694896	20	4523743	intronic	G	A	0.385	4.95E-08	3.64E-01	SLC13A3	ISOVF Arcuate Posterior Segment Left	3.74E-14	7.599572
277	rs2694896	20	4523743	intronic	G	A	0.385	4.95E-08	3.64E-01	SLC13A3	ISOVF Inferior Longitudinal Left	6.05E-13	7.199261
277	rs2694896	20	4523743	intronic	G	A	0.385	4.95E-08	3.64E-01	SLC13A3	ISOVF Frontal Aslant Tract Right	4.62E-11	6.582644
277	rs2694896	20	4523743	intronic	G	A	0.385	4.95E-08	3.64E-01	SLC13A3	ISOVF Arcuate Long Segment Left	2.50E-10	6.326993
277	rs2694896	20	4523743	intronic	G	A	0.385	4.95E-08	3.64E-01	SLC13A3	ISOVF Superior Longitudinal Fasciculus II Left	2.44E-16	8.198424
277	rs2694896	20	4523743	intronic	G	A	0.385	4.95E-08	3.64E-01	SLC13A3	ISOVF Arcuate Anterior Segment Right	8.85E-11	6.485454
277	rs2694896	20	4523743	intronic	G	A	0.385	4.95E-08	3.64E-01	SLC13A3	ISOVF Superior Longitudinal Fasciculus III Right	6.40E-13	7.191623
277	rs2694896	20	4523743	intronic	G	A	0.385	4.95E-08	3.64E-01	SLC13A3	ISOVF Arcuate Posterior Segment Right	1.63E-08	5.647213
277	rs6011998	20	45269867	intronic	T	C	0.042	9.31E-10	3.66E-01	SLC13A3	ISOVF Superior Longitudinal Fasciculus III Right	3.12E-08	5.534480
277	rs6011998	20	45269867	intronic	T	C	0.042	9.31E-10	3.66E-01	SLC13A3	ISOVF Superior Longitudinal Fasciculus III Left	1.35E-08	5.680163
277	rs6011998	20	45269867	intronic	T	C	0.042	9.31E-10	3.66E-01	SLC13A3	ISOVF Arcuate Anterior Segment Left	4.07E-10	6.251590
277	rs6011998	20	45269867	intronic	T	C	0.042	9.31E-10	3.66E-01	SLC13A3	ISOVF Arcuate Posterior Segment Left	1.93E-15	7.945830
277	rs6011998	20	45269867	intronic	T	C	0.042	9.31E-10	3.66E-01	SLC13A3	ISOVF Inferior Longitudinal Left	3.16E-09	5.923089
277	rs6011998	20	45269867	intronic	T	C	0.042	9.31E-10	3.66E-01	SLC13A3	ISOVF Arcuate Long Segment Left	3.20E-08	5.529942
277	rs6011998	20	45269867	intronic	T	C	0.042	9.31E-10	3.66E-01	SLC13A3	ISOVF Superior Longitudinal Fasciculus II Left	5.36E-10	6.208230
277	rs6011998	20	45269867	intronic	T	C	0.042	9.31E-10	3.66E-01	SLC13A3	ISOVF Superior Longitudinal Fasciculus III Right	4.98E-09	5.847971
277</													

D.1. Genomic loci associated with heritable language anatomical connectivities using MOSTest

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mvqwasP (discovery -British-)	mvqwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
280	rs6062264	20	61154871	intronic	T	C	0.279	2.12E-08	4.59E-01	Czoorf66	FA Superior Longitudinal Fasciculus II Right	3.05E-08	5.383255
280	rs6062264	20	61154871	intronic	T	C	0.279	2.12E-08	4.59E-01	Czoorf66	ICVF Superior Longitudinal Fasciculus II Left	6.88E-11	6.523167
280	rs6062264	20	61154871	intronic	T	C	0.279	2.12E-08	4.59E-01	Czoorf66	L3 Corpus callosum	1.59E-08	-5.652002
280	rs6062264	20	61154871	intronic	T	C	0.279	2.12E-08	4.59E-01	Czoorf66	FA Corpus callosum	1.46E-10	6.409307
280	rs6062264	20	61154871	intronic	T	C	0.279	2.12E-08	4.59E-01	Czoorf66	FA Superior Longitudinal Fasciculus II Left	1.13E-11	6.788361
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Superior Longitudinal Fasciculus II Right	1.31E-11	-6.767972
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	FA Inferior Longitudinal Left	9.36E-09	-5.742006
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Arcuate Anterior Segment Right	9.40E-12	-6.815330
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Arcuate Long Segment Left	8.38E-12	-6.831812
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Inferior Frontal Occipital fasciculus Right	5.53E-15	-7.814322
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	MD Inferior Frontal Occipital fasciculus Right	4.53E-08	5.468767
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Corpus callosum	8.28E-14	-7.465787
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Arcuate Posterior Segment Right	1.32E-09	-6.064682
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Inferior Frontal Occipital fasciculus Left	7.32E-15	-7.778804
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Frontal Aslant Tract Right	6.98E-12	-6.858133
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Superior Longitudinal Fasciculus II Left	6.33E-10	-6.181802
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	L3 Inferior Frontal Occipital fasciculus Right	1.85E-09	6.010664
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	L3 Inferior Frontal Occipital fasciculus Left	3.18E-10	6.289503
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	L3 Arcuate Long Segment Left	1.1E-08	5.733269
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Frontal Aslant Tract Left	1.24E-12	-7.013399
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Superior Longitudinal Fasciculus III Right	2.96E-09	-5.933574
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Frontal Superior Longitudinal Right	3.56E-09	-5.903279
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	FA Arcuate Long Segment Right	4.24E-08	-5.480486
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Arcuate Long Segment Right	3.95E-12	-6.938846
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Inferior Longitudinal Right	3.75E-13	-7.264328
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	L3 Arcuate Long Segment Right	2.48E-08	-5.574378
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Superior Longitudinal Fasciculus I Right	2.85E-13	-7.301488
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Superior Longitudinal Fasciculus I Left	1.67E-12	-7.059527
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Inferior Longitudinal Left	1.03E-13	-7.437239
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Superior Longitudinal Fasciculus III Left	2.18E-10	-6.347764
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	MD Inferior Frontal Occipital fasciculus Left	6.07E-09	5.814822
282	rs9608859	22	30667277	intergenic	C	T	0.437	2.31E-12	5.74E-01	OSM	ICVF Frontal inferior longitudinal Right	3.70E-08	-5.504648
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	MD Inferior Longitudinal Left	7.22E-09	5.785687
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Superior Longitudinal Fasciculus II Right	1.1E-14	-7.720407
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Arcuate Anterior Segment Right	2.57E-18	-8.720445
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	MD Inferior Longitudinal Right	3.70E-09	5.897937
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	L1 Inferior Frontal Occipital fasciculus Left	5.69E-11	6.651687
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Arcuate Long Segment Left	2.48E-10	-9.08658
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	FA Arcuate Anterior Segment Right	2.66E-09	-5.951242
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	FA Superior Longitudinal Fasciculus I Left	4.67E-10	-6.229846
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Inferior Frontal Occipital fasciculus Right	2.04E-16	-8.210492
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	FA Arcuate Long Segment Left	1.52E-08	-5.659434
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	MD Inferior Frontal Occipital fasciculus Right	1.40E-10	6.416151
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Corpus callosum	9.09E-18	-8.848882
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Arcuate Posterior Segment Right	2.03E-13	-7.346629
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Superior Longitudinal Fasciculus III Left	2.15E-08	-5.593788
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	MD Superior Longitudinal Fasciculus III Left	4.25E-08	5.701210
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Inferior Frontal Occipital fasciculus Left	6.79E-15	-7.788240
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Frontal Aslant Tract Right	2.18E-13	-7.337182
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	L1 Inferior Frontal Occipital fasciculus Right	7.73E-11	6.059777
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	L1 Arcuate Posterior Segment Left	4.27E-08	5.479777
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	L3 Superior Longitudinal Fasciculus III Left	1.21E-08	-6.079902
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Superior Longitudinal Fasciculus II Left	5.29E-15	-7.819283
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	L3 Inferior Frontal Occipital fasciculus Right	9.20E-09	5.447784
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	L3 Inferior Frontal Occipital fasciculus Left	7.38E-09	5.782038
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	L3 Arcuate Long Segment Left	3.85E-12	6.953980
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Frontal Aslant Tract Left	1.08E-13	-7.431097
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Superior Longitudinal Fasciculus III Right	2.33E-14	-7.630974
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Arcuate Posterior Segment Left	1.43E-13	-7.393747
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Arcuate Long Segment Right	3.82E-16	-8.144250
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Inferior Longitudinal Right	1.93E-16	-8.226645
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	L3 Arcuate Long Segment Right	7.28E-09	5.784271
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	FA Corpus callosum	4.41E-08	-5.473407
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	L1 Inferior Longitudinal Right	1.30E-08	5.685912
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	MD Arcuate Posterior Segment Left	1.20E-09	6.080398
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Arcuate Anterior Segment Left	7.30E-14	-7.482401
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Superior Longitudinal Fasciculus I Right	1.02E-17	-8.57281
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	OD Inferior Frontal Occipital fasciculus Right	4.21E-08	-5.486167
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Superior Longitudinal Fasciculus I Left	6.75E-17	-8.351324
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	FA Superior Longitudinal Fasciculus I Right	1.59E-08	-5.66112
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	MD Arcuate Long Segment Left	5.95E-12	6.880805
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Uncinate Right	5.44E-10	-6.205921
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	L3 Superior Longitudinal Fasciculus I Right	4.25E-08	5.479936
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Inferior Longitudinal Left	2.46E-15	-7.915699
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Superior Longitudinal Fasciculus III Left	3.62E-16	-8.150557
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	L3 Inferior Longitudinal Right	3.17E-08	5.531535
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	MD Inferior Frontal Occipital fasciculus Left	1.75E-10	6.381978
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	L3 Arcuate Anterior Segment Right	1.89E-09	6.009789
282	rs2267161	22	30953295	exonic	T	C	0.306	2.46E-12	2.57E-02	GAL3ST1	ICVF Frontal inferior longitudinal Right	1.01E-13	-7.439879
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	L3 Frontal Aslant Tract Right	2.17E-08	5.597846
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	FA Inferior Frontal Occipital fasciculus Right	3.98E-09	-5.884952
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	FA Superior Longitudinal Fasciculus III Right	8.41E-	

Genomic Locus	Lead SNP	chr	Position	Functional category	Non-effect allele	Effect allele	MAF	mvqwasP (discovery -British-)	mvqwasP (replication -non British-)	Nearest gene	Central endophenotypes	univariate P value	univariate Z
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	MD Superior Longitudinal Fasciculus III Left	1.66E-08	5.643807
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	ICVF Inferior Fronto Occipital fasciculus Left	8.09E-10	-6.143095
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	ICVF Frontal Aslant Tract Right	3.51E-08	-5.514008
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	L3 Superior Longitudinal Fasciculus III Left	5.07E-10	-6.216850
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	L3 Inferior Fronto Occipital fasciculus Right	3.00E-12	6.977807
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	L3 Inferior Fronto Occipital fasciculus Left	8.96E-13	7.145681
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	L3 Arcuate Long Segment Left	1.02E-14	7.737274
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	ICVF Superior Longitudinal Fasciculus III Right	7.81E-09	-5.772468
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	FA Frontal Aslant Tract Right	3.95E-08	-5.492995
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	ICVF Arcuate Posterior Segment Left	2.73E-08	-5.557851
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	ICVF Arcuate Long Segment Right	5.24E-11	-6.504042
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	L3 Inferior Longitudinal Left	1.87E-08	5.623294
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	L3 Corpus callosum	1.11E-08	5.722612
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	ICVF Inferior Longitudinal Right	8.87E-09	-5.751019
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	L3 Arcuate Long Segment Right	7.02E-10	6.165690
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	FA Corpus callosum	9.70E-11	-6.471490
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	ICVF Superior Longitudinal Fasciculus I Right	4.24E-09	-5.874402
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	FA Inferior Longitudinal Right	4.47E-08	-5.471067
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	MD Arcuate Long Segment Right	2.41E-09	-5.967378
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	ICVF Superior Longitudinal Fasciculus I Left	2.70E-09	-5.949066
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	FA Superior Longitudinal Fasciculus I Right	2.34E-11	-6.682975
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	L1 Arcuate Posterior Segment Right	3.97E-08	5.491966
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	MD Arcuate Long Segment Left	8.29E-12	6.833361
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	L3 Superior Longitudinal Fasciculus I Right	4.42E-10	6.238563
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	ICVF Inferior Longitudinal Left	1.58E-08	-5.652328
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	ICVF Superior Longitudinal Fasciculus III Left	3.98E-09	-5.885161
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	L3 Inferior Longitudinal Right	3.59E-10	6.271020
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	L3 Frontal Inferior Occipital Right	2.43E-08	5.578298
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	MD Inferior Fronto Occipital fasciculus Left	6.63E-10	6.174699
283	rs738443	22	38458681	ncRNA intronic	A	G	0.381	3.31E-18	4.80E-01	PICK1:RP5-1039K5.13	L3 Arcuate Anterior Segment Right	1.75E-09	6.019113
285	rs61748567	22	50685332	exonic	A	G	0.017	1.39E-22	3.23E-01	HDAC10:MAPK12	OD Superior Longitudinal Fasciculus II Right	1.17E-10	6.442761
285	rs61748567	22	50685332	exonic	A	G	0.017	1.39E-22	3.23E-01	HDAC10:MAPK12	OD Arcuate Posterior Segment Right	7.86E-09	5.771396
285	rs61748567	22	50685332	exonic	A	G	0.017	1.39E-22	3.23E-01	HDAC10:MAPK12	OD Superior Longitudinal Fasciculus III Right	4.33E-11	6.592241
285	rs61748567	22	50685332	exonic	A	G	0.017	1.39E-22	3.23E-01	HDAC10:MAPK12	FA Superior Longitudinal Fasciculus II Left	4.07E-09	-5.881400
285	rs61748567	22	50685332	exonic	A	G	0.017	1.39E-22	3.23E-01	HDAC10:MAPK12	OD Superior Longitudinal Fasciculus II Left	1.93E-15	7.945964

Résumé en français

Abstract in French

Sujet : L'architecture génétique du connectome du langage dans le cerveau humain

Motivations et contextes

Le langage est une composante essentielle de la vie humaine en raison de son rôle central dans les interactions sociales, la transmission culturelle, et la structuration de nos pensées. De part son omniprésence, le langage peut sembler avoir un caractère simple, mais celui-ci n'est pas suffisamment compris en tant que phénomène et processus. Dans différents contextes et sous différentes formes, les chercheurs soulignent le caractère particulier du langage par rapport aux autres capacités humaines. En conséquence, de nombreuses hypothèses entourent ce sujet dont celle relative à son origine. Le langage possède une dimension inhérente d'apprentissage, mais aussi une aspect innée qui se manifeste par le babillage des nourrissons, par exemple. Il est intéressant de noter que ce phénomène ne s'observe pas pour d'autres pratiques culturelles et qu'il est spécifique au langage. Le langage existe donc dans un espace unique, à la conjonction entre l'inné et l'acquis, entre individu et société.

Le langage est un phénomène à la fois biologique et social, il peut être abordé avec une grande variété de méthodes qui incluent les méthodes et les modèles explicatifs des sciences dites dures, ainsi que les outils offerts par les sciences humaines et sociales. La première approche se concentre sur les bases cognitives du langage et les différents aspects de son origine biologique qui permettent à l'espèce humaine d'apprendre et d'utiliser le langage. Elle tisse des relations étroites entre les neurosciences cognitives, la biologie humaine et aussi les statistiques qui fournissent des instruments précis pour sa modélisation. La deuxième approche, quant à elle, se concentre sur les relations entre le langage et la société, l'histoire, la philologie et la littérature. La nature même du langage, sa complexité, et son statut hybride justifient une pluralité d'approches et de méthodes. La présente thèse s'inscrit dans la première approche présentée ci-dessus et son objectif principal est de contribuer à l'étude de la neurobiologie du langage.

L'étude des causes biologiques des troubles du langage a suscité un vif intérêt pour la cartographie des gènes sous-jacents à cette capacité cognitive. Jusqu'à récemment, les racines biologiques du langage étaient principalement investiguées via des études de lésions, où les patients présentaient des troubles spécifiques du langage tels que les troubles du développement du langage et la dyslexie. L'imagerie-génomique est un domaine scientifique émergent qui bénéficie des propriétés non-invasives des techniques d'imagerie comme l'imagerie par résonance magnétique, qui permettent de "voir" le cerveau en fonctionnement, et du génotypage à haut débit par puces à ADN ou séquençage, qui permet de mesurer la variabilité génétique individuelle. Ces techniques permettent l'acquisition systématique des caractéristiques anatomiques et fonctionnelles du cerveau d'un individu ainsi que de son génotype. La possibilité d'acquérir ces données dans la population générale permet de mener des études sur des questions ouvertes de neurosciences cliniques et fondamentales. Les outils de la science des données nous permettent d'étudier comment la génétique et l'environnement modulent la fonction ou l'anatomie du cerveau humain ainsi que le comportement humain qui en découle. Avec la disponibilité récente de cohortes à grande échelle, il est devenu possible d'étudier le langage dans la population générale. L'objectif principal de cette thèse est de mettre en exergue les gènes qui sous-tendent le langage humain, améliorant ainsi notre compréhension des fondements génétiques du langage humain. À cette fin, nous avons étudié les variants génétiques communs qui s'associent aux traits quantitatifs, dérivés d'IRM fonctionnel de repos, liés au langage chez l'humain. Nous avons mis en avant les variations inter-individuelles de la fonction et de la structure du cerveau qui reflètent des traits neurocognitifs spécifiques du langage et avons testé les associations potentielles entre cette variabilité inter-individuelle observée et le profil génotypique. Cette étude a tiré parti de deux cohortes à grande échelle bien connues et adaptées aux études d'imagerie génétique : "UK Biobank" et "Human Connectome Project". Ces deux cohortes comprennent un génotypage en génome entier et une grande variété d'endophénotypes allant de données psychométriques et démographiques, en passant par la neuro-imagerie.

Déroulement du manuscrit de thèse

Le langage humain est un trait complexe. On s'attend à ce que différents types de variants génétiques affectent sa variabilité. Cette thèse vise à étudier les associations entre variations génétiques inter-individuelles et variations des endophénotypes dérivés des IRM cérébraux et liés au langage, afin de proposer des gènes jouant un rôle dans le fondement du langage chez l'humain.

Ce manuscrit de thèse est organisé comme suit :

Chapitre 4 : *Heritability of human language*

Dans ce chapitre, nous proposons une étude approfondie de l'héritabilité du langage humain caractérisé par des phénotypes intermédiaires soigneusement choisis. Plus particulièrement, nous avons évalué l'héritabilité de phénotypes dérivés de la psychométrie et de l'imagerie en tirant profit de deux ressources à grande échelle actuellement disponibles : le Human Connectome Project (HCP) et la UK Biobank (UKB) et ce, en utilisant des approches à l'état de l'art. Ces

(endo)phénotypes comprennent la connectivité fonctionnelle (CF) à l'état de repos, la connectivité structurelle et l'activation neural relative au langage. Nous avons observé une hiérarchie des estimations de l'héritabilité du langage avec 19,4-71,7 % pour les scores cognitifs du langage, 11,4-48,7 % pour la CF au repos et 14,4-34,9 % pour les activations neuronal du langage. La connectivité fonctionnelle au repos présente de faibles niveaux d'héritabilité par rapport aux scores cognitifs et aux caractéristiques structurelles, tant dans la population générale que dans les études sur les jumeaux. Ces observations sont cohérentes avec les héritabilités obtenues sur la population générale par [Elliott et al., 2018] qui a rapporté de faibles niveaux d'héritabilité pour la CF à l'état de repos par rapport aux endophénotypes structurels et de diffusion. En ce qui concerne la variabilité inter-cohorte, nous avons observé, et ce en accord avec la littérature, des valeurs d'héritabilité plus élevées dans les études généalogiques par rapport aux études d'héritabilité basées sur le génotypage. Les origines possibles de ces divergences ont été étudiées selon un plan expérimental qui prend en compte les dissimilarités entre les deux cohortes.

Chapitre 5 : *Shared genetic variance between language-related performance, language-brain activations and language functional connectivity*

Nous nous sommes intéressés en particulier aux données d'IRM fonctionnelle à l'état de repos et à leur pertinence pour l'étude de fonctions cognitives telles que le langage. Ce type de données offre de multiples avantages. Contrairement aux données d'IRM fonctionnelle basé sur une tâche, il est exempt de paradigme et, de ce fait, est plus facile à mettre en œuvre dans de très grandes cohortes comme UK Biobank. Par ailleurs l'IRM fonctionnelle à l'état de repos suscite un vif intérêt par la communauté scientifique du fait que certaines populations ont des difficultés à s'engager dans des tâches spécifiques. Bien que les données d'IRMf au repos soient considérées comme héritable, peu de connaissance existent sur leur relation avec le processus cognitif lié au langage.

Dans ce chapitre, nous abordons la question de l'héritabilité partagée en explorant la variabilité génétique commune qui expliquerait à la fois la variabilité des données d'IRMf et celles des mesures plus directes du langage en tant que processus cognitif. Pour ce faire, nous avons estimé les corrélations phénotypiques et génétiques entre les capacités cognitives mesurées par les tests psychométriques, les activations neuronales et les données de repos dans le langage. À cette fin, nous avons tiré parti de la cohorte du Human Connectome Project qui comprend à la fois, des scores liés au langage, des données d'IRMf basées sur des tâches de langage ainsi que des données d'IRMf en état de repos. Les résultats montrent une étiologie partagée entre la connectivité fonctionnelle à l'état de repos et les activations neuronales. Les résultats sont donc en faveur de l'existence d'un lien entre les données de l'état de repos et les données basées sur les tâches, tant sur le plan phénotypique que génétique, en ce qui concerne l'organisation cérébrale du langage humain.

Chapitre 6 : *The genetic architecture of human language functional connectome*

Dans ce chapitre, nous avons contribué à la compréhension de l'architecture génétique de la connectivité fonctionnelle du langage en mettant en évidence certains gènes liés au processus du langage. Nous avons extrait des phénotypes intermédiaire du langage - 300 paramètres de

connectivité fonctionnelle - à partir de données d'IRM à l'état de repos dans la plus grande cohorte d'imagerie-génétique sur population générale existante à ce jour (UK Biobank). Afin de s'assurer que ces endophénotypes reflètent bien la fonction du langage dans le cerveau, nous avons adopté une approche basée sur des régions d'intérêt pour l'estimation de la CF. Ces régions du langage ont été identifiées et dérivées d'une méta-analyse complète des études d'IRM fonctionnelle basées sur des expériences sollicitant explicitement le langage. Cette approche rend les endophénotypes comparables entre les individus et convient donc à une étude d'association. Nous avons appliqué un filtre sur l'héritabilité des CFs pour n'en garder qu'un sous ensemble significativement héritable. Nous avons ensuite mis en oeuvre une étude pangénomique multivariée afin de tirer parti de la structure de corrélation entre les CFs et de découvrir des loci potentiellement pléiotropes. Suite à une annotation fonctionnelle fine de ces loci, nous avons mis en évidence des gènes potentiels liés au langage, notamment :

- Le gène *EPHA3* présent dans le locus 3p11.1 avec un rôle potentiel dans le réseau *fronto-temporal sémantique*,
- Le gène *THBS1*, modulé par le locus 15q14, associé à la connectivité fonctionnelle présente dans l'interaction *perception-moteur* requise pour le langage, et
- Le gène *PLCE1* présent dans le locus 10q23.33 avec un rôle potentiel dans le *réseau bilatérale fronto-temporal auditif-moteur*.

Ces gènes pourraient être considérés comme prioritaires pour étudier le langage dans un modèle génétique approprié. Un nombre croissant de travaux avance que les études sur le langage devraient prendre en compte les données d'IRM fonctionnelle de repos en population générale afin de se concentrer sur l'organisation neurobiologique du langage dans sa globalité et sur la façon dont elle soutient le langage naturel, c'est-à-dire tel qu'il est utilisé dans la vie de tous les jours [Hasson et al., 2018]. D'un point de vue clinique, de tels conditions d'investigation sont hautement souhaitables car il est nécessaire de cartographier les zones du langage chez les patients incapables d'effectuer des tâches linguistiques [Branco et al., 2016, Klingbeil et al., 2019, Park et al., 2020, Ramage et al., 2020]. En cherchant à valider nos résultats obtenus sur les endophénotypes CF par l'étude d'autres endophénotypes dérivés de l'IRM de diffusion qui informent sur la connectivité anatomique, nous avons essayé de donner un point de vue du connectome fonctionnel et structurel de l'organisation du langage dans le cerveau. Enfin, de récentes approches d'imagerie génétique telles que les études pangénomique multivariées, nous ont permis d'augmenter la puissance statistique et de contourner les problématique lié aux très faibles tailles d'effet. Nous pensons que ces approches sont prometteuses et qu'elles seront largement diffusées dans le domaine de l'imagerie génétique et au-delà. Avec l'approche présentée, nous avons essayé de contribuer à ces tendances innovantes et d'ouvrir la voie à d'autres approches alternatives utilisant l'IRMf de repos pour étudier le langage naturel et ses fondements génétiques.

Dans ce chapitre, nous avons adopté un plan expérimental similaire à celui présenté dans le chapitre 6 pour étudier l'architecture génétique de la connectivité anatomique à l'échelle de toute la connectivité structurelle actuellement connue dans le langage. Nous avons extrait les endophénotypes de connectivité anatomique (CA) du langage à partir d'acquisitions pondérées en diffusion disponibles dans la cohorte UK Biobank. Nous avons adopté une approche basée sur régions d'intérêt pour l'estimation de la CA avec des tracts de matière blanche du langage bien connus. Après avoir filtré ces endophénotypes en considérant leur héritabilité, nous avons mis en oeuvre une étude pangénomique multivariée afin de tirer parti de la structure de corrélation entre la connectivité structurelle et de découvrir des loci pléiotropes potentiels. Nous avons ainsi mis en évidence des gènes potentiels liés au traitement du langage. Ces gènes pourraient être prioritaires pour étudier le langage dans un modèle génétique approprié. Les résultats suggèrent de nouveaux gènes candidats qui sont liés à des voies biologiques déjà présumées être importantes dans le développement du cerveau humain et le langage. Dans l'ensemble, ces résultats proposent une nouvelle vision de l'architecture génétique du langage chez l'homme.

Discussion, limitations

Une approche intégrative. Nous avons suivi une approche intégrative dans notre étude de l'architecture génétique du langage. Tout d'abord, en validant nos résultats sur les endophénotypes CF de l'IRM fonctionnelle de repos avec les endophénotypes dérivés de l'IRM de diffusion, nous avons essayé de fournir un point de vue à la fois fonctionnel et structurel du connectome de l'organisation du langage dans le cerveau. Deuxièmement, nous avons utilisé des approches récentes d'association imagerie-génétique telles que les études pangénomique multivariées, qui nous ont permis d'augmenter la puissance statistique et de contourner les petites tailles d'effet produites par des variants génétiques sur ces endophénotypes initiaux. Nous pensons que ces approches sont prometteuses et qu'elles seront largement diffusées dans le domaine de l'imagerie-génétique et au-delà. Avec l'approche présentée, nous avons essayé de contribuer à ces tendances innovantes et d'ouvrir la voie à d'autres approches alternatives d'IRMf au repos pour étudier le langage naturel et ses fondements génétiques.

Une étude en population générale. Nous avons mené notre étude sur une cohorte de la population générale en sélectionnant des (endo)phénotypes du langage à l'état de l'art. Nous pensons qu'elle offre un cadre alternatif intéressant aux études basées sur les troubles du langage. Les cohortes de la population générale fournissent une multitude de mesures qui permettent de développer des (endo)phénotypes d'intérêt, y compris pour les processus cognitifs complexes. Cette approche offre une alternative prometteuse au paradigme cas-témoins.

Conclusion

Le langage est un élément central de la vie humaine. Comprendre comment nous sommes capables d'acquérir et d'utiliser ce trait hautement complexe a nécessité le développement de nombreuses descriptions du langage à plusieurs niveaux. D'une part, nous avons exploité des scores

de traitement du langage bien caractérisés et des phénotypes intermédiaires. D'autre part, nous avons appliqué des méthodes d'association pour identifier les gènes liés aux différences comportementales et comprendre leur impact sur le développement et le fonctionnement du cerveau en examinant la fonction cellulaire des protéines que codent ces gènes et les voies biologiques dans lesquelles ils sont impliqués. Une compréhension plus complète du langage humain nécessitera d'intégrer les résultats des différentes stratégies entreprises et des différents niveaux de description.

Dans cette thèse, nous avons abordé la question de la base génétique du langage humain en utilisant des méthodes à l'état de l'art : en mettant en évidence la variabilité inter-individuelle du cerveau et en évaluant le rôle des variants génétiques communs pris dans l'ensemble du génome sans *a priori* sur des gènes candidats potentiels. Nous avons relié ces recherches aux connaissances actuelles sur les performances cognitives liées au langage et aux mécanismes moléculaires connus dans le cerveau et le développement. Les résultats de cette thèse fournissent de nouveaux points d'entrée pour de futurs travaux visant à combler le fossé entre la contribution génétique et le processus du langage humain.

* * *
* *
*

Bibliography

- The gene ontology resource: enriching a gold mine. *Nucleic Acids Research*, 49(D1):D325–D334, 2021.
- Alexandre Abraham, Fabian Pedregosa, Michael Eickenberg, Philippe Gervais, Andreas Mueller, Jean Kossaifi, Alexandre Gramfort, Bertrand Thirion, and Gaël Varoquaux. Machine learning for neuroimaging with scikit-learn. *Frontiers in neuroinformatics*, 8:14, 2014.
- Fidel Alfaro-Almagro, Mark Jenkinson, Neal K Bangerter, Jesper LR Andersson, Ludovica Griffanti, Gwenaëlle Douaud, Stamatios N Sotiropoulos, Saad Jbabdi, Moises Hernandez-Fernandez, Emmanuel Vallee, et al. Image processing and quality control for the first 10,000 brain imaging datasets from uk biobank. *Neuroimage*, 166:400–424, 2018.
- David B Allison, Bonnie Thiel, Pamela St Jean, Robert C Elston, Ming C Infante, and Nicholas J Schork. Multiple phenotype modeling in gene-mapping studies of quantitative traits: power advantages. *The American Journal of Human Genetics*, 63(4):1190–1201, 1998.
- Laura Almasy and John Blangero. Multipoint quantitative-trait linkage analysis in general pedigrees. *The American Journal of Human Genetics*, 62(5):1198–1211, 1998.
- Jesper LR Andersson and Stamatios N Sotiropoulos. Non-parametric representation and prediction of single-and multi-shell diffusion-weighted mri data using gaussian processes. *Neuroimage*, 122:166–176, 2015.
- Heidi Anthoni, Marco Zucchelli, Hans Matsson, Bertram Müller-Myhsok, Ingegerd Fransson, Johannes Schumacher, Satu Massinen, Päivi Onkamo, Andreas Warnke, Heide Griesemann, et al. A locus on 2p12 containing the co-regulated mrpl19 and c2orf3 genes is associated to dyslexia. *Human molecular genetics*, 16(6):667–677, 2007.
- John Ashburner and Karl J Friston. Voxel-based morphometry—the methods. *Neuroimage*, 11(6):805–821, 2000.
- Michael Ashburner, Catherine A Ball, Judith A Blake, David Botstein, Heather Butler, J Michael Cherry, Allan P Davis, Kara Dolinski, Selina S Dwight, Janan T Eppig, et al. Gene ontology: tool for the unification of biology. *Nature genetics*, 25(1):25–29, 2000.

- Deanna M Barch, Gregory C Burgess, Michael P Harms, Steven E Petersen, Bradley L Schlaggar, Maurizio Corbetta, Matthew F Glasser, Sandra Curtiss, Sachin Dixit, Cindy Feldd, et al. Function in the human connectome: task-fmri and individual differences in behavior. *Neuroimage*, 80:169–189, 2013.
- Christopher W Bartlett, Judy F Flax, Mark W Logue, Veronica J Vieland, Anne S Bassett, Paula Tallal, and Linda M Brzustowicz. A major susceptibility locus for specific language impairment is located on 13q21. *The American Journal of Human Genetics*, 71(1):45–55, 2002.
- Christopher W Bartlett, Judy F Flax, Mark W Logue, Brett J Smith, Veronica J Vieland, Paula Tallal, and Linda M Brzustowicz. Examination of potential overlap in autism and language loci on chromosomes 2, 7, and 13 in two independent samples ascertained for specific language impairment. *Human heredity*, 57(1):10–20, 2004.
- T Mark Beasley, Stephen Erickson, and David B Allison. Rank-based inverse normal transformations are increasingly used, but are they merited? *Behavior genetics*, 39(5):580–595, 2009.
- Christian F Beckmann and Stephen M Smith. Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE transactions on medical imaging*, 23(2):137–152, 2004.
- Emma Belton, Claire H Salmond, Kate E Watkins, Faraneh Vargha-Khadem, and David G Gadian. Bilateral brain abnormalities associated with dominantly inherited verbal and orofacial dyspraxia. *Human brain mapping*, 18(3):194–200, 2003.
- Michel Belyk and Steven Brown. The origins of the vocal brain in humans. *Neuroscience & Biobehavioral Reviews*, 77:177–193, 2017.
- Sarah E Bergen, Charles O Gardner, and Kenneth S Kendler. Age-related changes in heritability of behavioral phenotypes over adolescence and young adulthood: a meta-analysis. *Twin Research and Human Genetics*, 10(3):423–433, 2007.
- Jeffrey R Binder, Rutvik H Desai, William W Graves, and Lisa L Conant. Where is the semantic system? a critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral cortex*, 19(12):2767–2796, 2009.
- Jeffrey R Binder, William L Gross, Jane B Allendorfer, Leonardo Bonilha, Jessica Chapin, Jonathan C Edwards, Thomas J Grabowski, John T Langfitt, David W Loring, Mark J Lowe, et al. Mapping anterior temporal lobe language areas with fmri: a multicenter normative study. *Neuroimage*, 54(2):1465–1475, 2011.
- Bharat Biswal, F Zerrin Yetkin, Victor M Haughton, and James S Hyde. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic resonance in medicine*, 34(4):537–541, 1995.

- Bharat B Biswal, Maarten Mennes, Xi-Nian Zuo, Suril Gohel, Clare Kelly, Steve M Smith, Christian F Beckmann, Jonathan S Adelstein, Randy L Buckner, Stan Colcombe, et al. Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences*, 107(10):4734–4739, 2010.
- Gunnar Blom. *Statistical estimates and transformed beta-variables*. PhD thesis, Almqvist & Wiksell, 1958.
- James R Booth, Douglas D Burman, Joel R Meyer, Darren R Gitelman, Todd B Parrish, and M Marsel Mesulam. Modality independence of word comprehension. *Human brain mapping*, 16(4):251–261, 2002.
- Thomas J Bouchard. The wilson effect: the increase in heritability of iq with age. *Twin Research and Human Genetics*, 16(5):923–930, 2013.
- Alan P Boyle, Eurie L Hong, Manoj Hariharan, Yong Cheng, Marc A Schaub, Maya Kasowski, Konrad J Karczewski, Julie Park, Benjamin C Hitz, Shuai Weng, et al. Annotation of functional variation in personal genomes using regulomedb. *Genome research*, 22(9):1790–1797, 2012.
- Paulo Branco, Daniela Seixas, Sabine Deprez, Silvia Kovacs, Ronald Peeters, São L Castro, and Stefan Sunaert. Resting-state functional magnetic resonance imaging for language preoperative planning. *Frontiers in human neuroscience*, 10:11, 2016.
- Korbinian Brodmann. Beiträge zur histologischen lokalisation der grosshirnrinde: Vi. die cortexgliederung des menschen. *Journal für Psychologie und Neurologie*, 10:231–246, 1909.
- Toai Bui and J Nguyen. Neuroanatomy, cerebral hemisphere. 2019.
- Brendan K Bulik-Sullivan, Po-Ru Loh, Hilary K Finucane, Stephan Ripke, Jian Yang, Nick Patterson, Mark J Daly, Alkes L Price, and Benjamin M Neale. Ld score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics*, 47(3):291–295, 2015.
- Clare Bycroft, Colin Freeman, Desislava Petkova, Gavin Band, Lloyd T Elliott, Kevin Sharp, Allan Motyer, Damjan Vukcevic, Olivier Delaneau, Jared O’Connell, et al. The uk biobank resource with deep phenotyping and genomic data. *Nature*, 562(7726):203–209, 2018.
- Mario Cáceres, Joel Lachuer, Matthew A Zapala, John C Redmond, Lili Kudo, Daniel H Geschwind, David J Lockhart, Todd M Preuss, and Carrolee Barlow. Elevated gene expression levels distinguish human from non-human primate brains. *Proceedings of the National Academy of Sciences*, 100(22):13030–13035, 2003.
- Mario Cáceres, Carolyn Suwyn, Marcelia Maddox, James W Thomas, and Todd M Preuss. Increased cortical expression of two synaptogenic thrombospondins in human brain evolution. *Cerebral Cortex*, 17(10):2312–2321, 2007.

- Vince D Calhoun, Tulay Adali, Godfrey D Pearlson, and James J Pekar. A method for making group inferences from functional mri data using independent component analysis. *Human brain mapping*, 14(3):140–151, 2001.
- Amaia Carrion-Castillo, Antonietta Pepe, Xiang-Zhen Kong, Simon E Fisher, Bernard Mazoyer, Nathalie Tzourio-Mazoyer, Fabrice Crivello, and Clyde Francks. Genetic effects on planum temporale asymmetry and their limited relevance to neurodevelopmental disorders, intelligence or educational attainment. *cortex*, 124:137–153, 2020.
- Jillian P Casey, Tiago Magalhaes, Judith M Conroy, Regina Regan, Naisha Shah, Richard Anney, Denis C Shields, Brett S Abrahams, Joana Almeida, Elena Bacchelli, et al. A novel approach of homozygous haplotype sharing identifies candidate genes in autism spectrum disorder. *Human genetics*, 131(4):565–579, 2012.
- Marco Catani and Valentina Bambini. A model for social communication and language evolution and development (scaled). *Current opinion in neurobiology*, 28:165–171, 2014.
- Marco Catani and Michel Thiebaut De Schotten. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *cortex*, 44(8):1105–1132, 2008.
- Marco Catani and Stephanie J Forkel. Diffusion imaging methods in language sciences. *The Oxford Handbook of Neurolinguistics*, page 212, 2019.
- Marco Catani, Derek K Jones, and Dominic H Ffytche. Perisylvian language networks of the human brain. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 57(1):8–16, 2005.
- Marco Catani, Flavio Dell’Acqua, Francesco Vergani, Farah Malik, Harry Hodge, Prasun Roy, Romain Valabregue, and Michel Thiebaut De Schotten. Short frontal lobe connections of the human brain. *cortex*, 48(2):273–291, 2012.
- Christopher C Chang, Carson C Chow, Laurent CAM Tellier, Shashaank Vattikuti, Shaun M Purcell, and James J Lee. Second-generation plink: rising to the challenge of larger and richer datasets. *Gigascience*, 4(1):s13742–015, 2015.
- Sreenivas Chavali, Fredrik Barrenas, Kartiek Kanduri, and Mikael Benson. Network properties of human disease genes with pleiotropic effects. *BMC systems biology*, 4(1):1–11, 2010.
- Kulpreet Cheema, Amberley V Ostevik, Lindsey Westover, William E Hodgetts, and Jacqueline Cummine. Resting-state networks and reading in adults with and without reading impairments. *Journal of Neurolinguistics*, 60:101016, 2021.
- Karen S Christopherson, Erik M Ullian, Caleb CA Stokes, Christine E MULLowney, Johannes W Hell, Azin Agah, Jack Lawler, Deane F Mosher, Paul Bornstein, and Ben A Barres. Thrombospondins are astrocyte-secreted proteins that promote cns synaptogenesis. *Cell*, 120(3):421–433, 2005.

- Maria-Daniela Cirnaru, Sicheng Song, Kizito-Tshitoko Tshilenge, Chuhyon Corwin, Justyna Mleczko, Carlos Galicia Aguirre, Houda Benlhabib, Jaroslav Bendl, John Fullard, Pasha Apontes, et al. Transcriptional and epigenetic characterization of early striosomes identifies *foxf2* and *olig2* as factors required for development of striatal compartmentation and neuronal phenotypic differentiation. *bioRxiv*, 2020.
- Michael W Cole, Danielle S Bassett, Jonathan D Power, Todd S Braver, and Steven E Petersen. Intrinsic and task-evoked network architectures of the human brain. *Neuron*, 83(1):238–251, 2014.
- 1000 Genomes Project Consortium et al. A global reference for human genetic variation. *Nature*, 526(7571):68, 2015.
- GTEEx Consortium et al. Genetic effects on gene expression across human tissues. *Nature*, 550(7675):204–213, 2017.
- Sli Consortium et al. Highly significant linkage to the *slii* locus in an expanded sample of individuals affected by specific language impairment. *The American Journal of Human Genetics*, 74(6):1225–1238, 2004.
- R Todd Constable, Kenneth R Pugh, Ella Berroya, W Einar Mencl, Michael Westerveld, Weijia Ni, and Donald Shankweiler. Sentence complexity and input modality effects in sentence comprehension: an fmri study. *NeuroImage*, 22(1):11–21, 2004.
- Natalie Cope, John D Eicher, Haiying Meng, Christopher J Gibson, Karl Hager, Cheryl Lacadie, Robert K Fulbright, R Todd Constable, Grier P Page, and Jeffrey R Gruen. Variants in the *dyx2* locus are associated with altered brain activation in reading-related brain regions in subjects with reading disability. *Neuroimage*, 63(1):148–156, 2012.
- Jillian M Couto, Lissette Gomez, Karen Wigg, Tasha Cate-Carter, Jennifer Archibald, Barbara Anderson, Rosemary Tannock, Elizabeth N Kerr, Maureen W Lovett, Tom Humphries, et al. The *kiaa0319*-like (*kiaa0319l*) gene on chromosome 1p34 as a candidate for reading disabilities. *Journal of neurogenetics*, 22(4):295–313, 2008.
- Simon R Cox, Stuart J Ritchie, Chloe Fawns-Ritchie, Elliot M Tucker-Drob, and Ian J Deary. Structural brain imaging correlates of general intelligence in uk biobank. *Intelligence*, 76:101376, 2019.
- Alexandra M Cross, Reshma Ramdajal, Lien Peters, Matthew RJ Vandermeer, Elizabeth P Hayden, Jan C Frijters, Karen A Steinbach, Maureen W Lovett, Lisa MD Archibald, and Marc F Joanisse. Resting-state functional connectivity and reading subskills in children. *NeuroImage*, 243:118529, 2021.
- Darina Czamara, Jennifer Bruder, Jessica Becker, Jürgen Bartling, Per Hoffmann, Kerstin Urte Ludwig, Bertram Müller-Myhsok, and Gerd Schulte-Körne. Association of a rare variant with mismatch negativity in a region between *kiaa0319* and *dcdc2* in dyslexia. *Behavior genetics*, 41(1):110–119, 2011.

- Alessandro Daducci, Erick J Canales-Rodríguez, Hui Zhang, Tim B Dyrby, Daniel C Alexander, and Jean-Philippe Thiran. Accelerated microstructure imaging via convex optimization (amico) from diffusion mri data. *NeuroImage*, 105:32–44, 2015.
- Philip S Dale, Ginette Dionne, Thalia C Eley, and Robert Plomin. Lexical and grammatical development: A behavioural genetic perspective. *Journal of child language*, 27(3):619–642, 2000.
- Fahimeh Darki, Myriam Peyrard-Janvid, Hans Matsson, Juha Kere, and Torkel Klingberg. Three dyslexia susceptibility genes, *dys1c1*, *dcdc2*, and *kiaa0319*, affect temporo-parietal white matter structure. *Biological psychiatry*, 72(8):671–676, 2012.
- Fahimeh Darki, Myriam Peyrard-Janvid, Hans Matsson, Juha Kere, and Torkel Klingberg. *Dcdc2* polymorphism is associated with left temporoparietal gray and white matter structures during development. *Journal of Neuroscience*, 34(43):14455–14462, 2014.
- Gail Davies, Riccardo E Marioni, David C Liewald, W David Hill, Saskia P Hagenaars, Sarah E Harris, Stuart J Ritchie, Michelle Luciano, Chloe Fawns-Ritchie, Donald Lyall, et al. Genome-wide association study of cognitive functions and educational attainment in uk biobank (n=112 151). *Molecular psychiatry*, 21(6):758–767, 2016.
- CGF De Kovel, FA Hol, JGAM Heister, JJHT Willems, LA Sandkuijl, B Franke, and GW Padberg. Genomewide scan identifies susceptibility locus for dyslexia on xq27 in an extended dutch family. *Journal of medical genetics*, 41(9):652–657, 2004.
- Christiaan A de Leeuw, Joris M Mooij, Tom Heskes, and Danielle Posthuma. Magma: generalized gene-set analysis of gwas data. *PLoS Comput Biol*, 11(4):e1004219, 2015.
- Karen E Deffenbacher, Judith B Kenyon, Denise M Hoover, Richard K Olson, Bruce F Pennington, John C DeFries, and Shelley D Smith. Refinement of the 6p21. 3 quantitative trait locus influencing dyslexia: linkage and association analyses. *Human Genetics*, 115(2):128–138, 2004.
- Emily L Dennis, Neda Jahanshad, Jeffrey D Rudie, Jesse A Brown, Kori Johnson, Katie L McMahon, Greig I de Zubicaray, Grant Montgomery, Nicholas G Martin, Margaret J Wright, et al. Altered structural brain connectivity in healthy carriers of the autism risk gene, *cntnap2*. *Brain connectivity*, 1(6):447–459, 2011.
- Rahul S Desikan, Florent Ségonne, Bruce Fischl, Brian T Quinn, Bradford C Dickerson, Deborah Blacker, Randy L Buckner, Anders M Dale, R Paul Maguire, Bradley T Hyman, et al. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *Neuroimage*, 31(3):968–980, 2006.
- Nina F Dronkers. The pursuit of brain–language relationships. *Brain and Language*, 71(1):59–61, 2000.

- Hugues Duffau, Peggy Gatignol, Dominique Denvil, Manuel Lopes, and Laurent Capelle. The articulatory loop: study of the subcortical connectivity by electrostimulation. *Neuroreport*, 14(15):2005–2008, 2003.
- John D Eicher, Natalie R Powers, Laura L Miller, Natacha Akshoomoff, David G Amaral, Cinnamon S Bloss, Ondrej Libiger, Nicholas J Schork, Burcu F Darst, BJ Casey, et al. Genome-wide association study of shared components of reading disability and language impairment. *Genes, Brain and Behavior*, 12(8):792–801, 2013.
- John D Eicher, Angela M Montgomery, Natacha Akshoomoff, David G Amaral, Cinnamon S Bloss, Ondrej Libiger, Nicholas J Schork, Burcu F Darst, BJ Casey, Linda Chang, et al. Dyslexia and language impairment associated genetic markers influence cortical thickness and white matter in typically developing children. *Brain imaging and behavior*, 10(1):272–282, 2016.
- Jennifer Stine Elam, Matthew F Glasser, Michael P Harms, Stamatios N Sotiropoulos, Jesper LR Andersson, Gregory C Burgess, Sandra W Curtiss, Robert Oostenveld, Linda J Larson-Prior, Jan-Mathijs Schoffelen, et al. The human connectome project: A retrospective. *NeuroImage*, page 118543, 2021.
- Lloyd T Elliott, Kevin Sharp, Fidel Alfaró-Almagro, Sinan Shi, Karla L Miller, Gwenaëlle Douaud, Jonathan Marchini, and Stephen M Smith. Genome-wide association studies of brain imaging phenotypes in uk biobank. *Nature*, 562(7726):210–216, 2018.
- Maxwell L Elliott, Annchen R Knodt, Megan Cooke, M Justin Kim, Tracy R Melzer, Ross Keenan, David Ireland, Sandhya Ramrakha, Richie Poulton, Avshalom Caspi, et al. General functional connectivity: Shared features of resting-state and task fmri drive reliable and heritable individual differences in functional brain networks. *NeuroImage*, 189:516–532, 2019.
- Jason Ernst and Manolis Kellis. Chromhmm: automating chromatin-state discovery and characterization. *Nature methods*, 9(3):215–216, 2012.
- Alan C Evans, Brain Development Cooperative Group, et al. The nih mri study of normal brain development. *Neuroimage*, 30(1):184–202, 2006.
- Luciano Fadiga, Laila Craighero, Giovanni Buccino, and Giacomo Rizzolatti. Speech listening specifically modulates the excitability of tongue muscles: a tms study. *European journal of Neuroscience*, 15(2):399–402, 2002.
- Manuel AR Ferreira and Shaun M Purcell. A multivariate test of association. *Bioinformatics*, 25(1):132–133, 2009.
- LL Field, K Shumansky, J Ryan, D Truong, E Swiergala, and BJ Kaplan. Dense-map genome scan for dyslexia supports loci at 4q13, 16p12, 17q22; suggests novel locus at 7q36. *Genes, Brain and Behavior*, 12(1):56–69, 2013.

- Deborah Finkel and Matt McGue. Age differences in the nature and origin of individual differences in memory: A behavior genetic analysis. *The International Journal of Aging and Human Development*, 47(3):217–239, 1998.
- Bruce Fischl and Anders M. Dale. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences of the United States of America*, 97(20):11050–11055, 2000.
- Simon E Fisher and Sonja C Vernes. Genetics and the language sciences. *Annu. Rev. Linguist.*, 1(1):289–310, 2015.
- Simon E Fisher, Faraneh Vargha-Khadem, Kate E Watkins, Anthony P Monaco, and Marcus E Pembrey. Localisation of a gene implicated in a severe speech and language disorder. *Nature genetics*, 18(2):168–170, 1998.
- Simon E Fisher, Clyde Francks, Angela J Marlow, I Laurence MacPhie, Dianne F Newbury, Lon R Cardon, Yumiko Ishikawa-Brush, Alex J Richardson, Joel B Talcott, Javier Gayán, et al. Independent genome-wide scans identify a chromosome 18 quantitative-trait locus influencing dyslexia. *Nature genetics*, 30(1):86–91, 2002.
- Adeen Flinker, Anna Korzeniewska, Avgusta Y Shestyuk, Piotr J Franaszczuk, Nina F Dronkers, Robert T Knight, and Nathan E Crone. Redefining the role of broca’s area in speech. *Proceedings of the National Academy of Sciences*, 112(9):2871–2875, 2015.
- Stephanie J Forkel, Michel Thiebaut de Schotten, Jamie M Kawadler, Flavio Dell’Acqua, Adrian Danek, and Marco Catani. The anatomy of fronto-occipital connections from early blunt dissections to contemporary tractography. *Cortex*, 56:73–84, 2014.
- Michael D Fox and Marcus E Raichle. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature reviews neuroscience*, 8(9):700–711, 2007.
- Julius Fridriksson, Dana Moser, Jack Ryalls, Leonardo Bonilha, Chris Rorden, and Gordon Baylis. Modulation of frontal lobe speech areas associated with the production and perception of speech movements. 2009.
- Angela D Friederici. The brain basis of language processing: from structure to function. *Physiological reviews*, 91(4):1357–1392, 2011.
- Angela D Friederici. *Language in our brain: The origins of a uniquely human capacity*. MIT Press, 2017.
- Karl J Friston. Functional and effective connectivity in neuroimaging: a synthesis. *Human brain mapping*, 2(1-2):56–78, 1994.
- Karl J Friston. Modalities, modes, and models in functional neuroimaging. *Science*, 326(5951):399–403, 2009.

- Karl J Friston, Andrew P Holmes, Keith J Worsley, J-P Poline, Chris D Frith, and Richard SJ Frackowiak. Statistical parametric maps in functional imaging: a general linear approach. *Human brain mapping*, 2(4):189–210, 1994.
- Joaquin M Fuster. *Cortex and mind: Unifying cognition*. Oxford university press, 2002.
- Bruno Galantucci, Carol A Fowler, and Michael T Turvey. The motor theory of speech perception reviewed. *Psychonomic bulletin & review*, 13(3):361–377, 2006.
- Tessel E Galesloot, Kristel Van Steen, Lambertus ALM Kiemeney, Luc L Janss, and Sita H Vermeulen. A comparison of multivariate genome-wide association methods. *PLoS one*, 9(4):e95923, 2014.
- J Ganger, S Pinker, S Chawla, and A Baker. Heritability of early milestones of vocabulary and grammar: A twin study. *Unpublished manuscript*, 2002.
- Katrin Gerstmann and Geraldine Zimmer. The role of the eph/ephrin family during cortical development and cerebral malformations. *Medical Research Archives*, 6(3), 2018.
- Alessandro Gialluisi, Dianne F Newbury, Erik G Wilcutt, Richard K Olson, John C DeFries, William M Brandler, Bruce F Pennington, Shelley D Smith, Thomas S Scerri, Nuala H Simpson, et al. Genome-wide screening for dna variants associated with reading and language traits. *Genes, Brain and Behavior*, 13(7):686–701, 2014.
- Alessandro Gialluisi, Tulio Guadalupe, Clyde Francks, and Simon E Fisher. Neuroimaging genetic analyses of novel candidate genes associated with reading and language. *Brain and language*, 172:9–15, 2017.
- Richard A Gibbs, John W Belmont, Paul Hardenbol, Thomas D Willis, FL Yu, HM Yang, Lan-Yang Ch'ang, Wei Huang, Bin Liu, Yan Shen, et al. The international hapmap project. 2003.
- Daniel A Gibson and Le Ma. Developmental regulation of axon branching in the vertebrate nervous system. *Development*, 138(2):183–195, 2011.
- David C Glahn, Paul M Thompson, and John Blangero. Neuroimaging endophenotypes: strategies for finding genes influencing brain structure and function. *Human brain mapping*, 28(6):488–501, 2007.
- Matthew F Glasser, Stamatios N Sotiropoulos, J Anthony Wilson, Timothy S Coalson, Bruce Fischl, Jesper L Andersson, Junqian Xu, Saad Jbabdi, Matthew Webster, Jonathan R Polimeni, et al. The minimal preprocessing pipelines for the human connectome project. *Neuroimage*, 80:105–124, 2013.
- Gary H Glover. Overview of functional magnetic resonance imaging. *Neurosurgery Clinics*, 22(2):133–139, 2011.

- Irving I Gottesman and Todd D Gould. The endophenotype concept in psychiatry: etymology and strategic intentions. *American journal of psychiatry*, 160(4):636–645, 2003.
- Michael D Greicius, Ben Krasnow, Allan L Reiss, and Vinod Menon. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences*, 100(1):253–258, 2003.
- Ludovica Griffanti, Gholamreza Salimi-Khorshidi, Christian F Beckmann, Edward J Auerbach, Gwenaëlle Douaud, Claire E Sexton, Enikő Zsoldos, Klaus P Ebmeier, Nicola Filippini, Clare E Mackay, et al. Ica-based artefact removal and accelerated fmri acquisition for improved resting state network imaging. *Neuroimage*, 95:232–247, 2014.
- Yongtao Guan and Matthew Stephens. Practical issues in imputation-based association mapping. *PLoS genetics*, 4(12):e1000279, 2008.
- Jennifer M Gurd, Katrin Amunts, Peter H Weiss, Oliver Zafris, Karl Zilles, John C Marshall, and Gereon R Fink. Posterior parietal cortex is implicated in continuous switching between verbal fluency tasks: an fmri study with clinical implications. *Brain*, 125(5):1024–1038, 2002.
- Peter Hagoort. *Human Language: From Genes and Brains to Behavior*. The MIT Press, 2019.
- Michelle Hampson, Fuyuze Tokoglu, Zhongdong Sun, Robin J Schafer, Pawel Skudlarski, John C Gore, and R Todd Constable. Connectivity–behavior analysis reveals that functional connectivity between left ba39 and broca’s area varies with reading ability. *Neuroimage*, 31(2): 513–519, 2006.
- Yixing Han, Jianke Ren, Eunice Lee, Xiaoping Xu, Weishi Yu, and Kathrin Muegge. Lsh/hells regulates self-renewal/proliferation of neural stem/progenitor cells. *Scientific reports*, 7(1): 1–14, 2017.
- Katariina Hannula-Jouppi, Nina Kaminen-Ahola, Mikko Taipale, Ranja Eklund, Jaana Nopola-Hemmi, Helena Kääriäinen, and Juha Kere. The axon guidance receptor gene *robo1* is a candidate gene for developmental dyslexia. *PLoS genetics*, 1(4):e50, 2005.
- Nicole Harlaar, Emma L Meaburn, Marianna E Hayiou-Thomas, Oliver SP Davis, Sophia Docherty, Ken B Hanscombe, Claire MA Haworth, Thomas S Price, Maciej Trzaskowski, Philip S Dale, et al. Genome-wide association study of receptive language ability of 12-year-olds. 2014.
- Uri Hasson, Giovanna Egidi, Marco Marelli, and Roel M Willems. Grounding the neurobiology of language in first principles: The necessity of non-language-centric explanations for language comprehension. *Cognition*, 180:135–157, 2018.
- Isabelle Hesling, Loic Labache, Marc Joliot, and Nathalie Tzourio-Mazoyer. Large-scale plurimodal networks common to listening to, producing and reading word lists: an fmri study combining task-induced activation and intrinsic connectivity in 144 right-handers. *Brain Structure and Function*, 224(9):3075–3094, 2019.

- Gregory Hickok and David Poeppel. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2):67–99, 2004.
- Gregory Hickok and David Poeppel. The cortical organization of speech processing. *Nature reviews neuroscience*, 8(5):393–402, 2007.
- William G Hill, Michael E Goddard, and Peter M Visscher. Data and theory point to mainly additive genetic variance for complex traits. *PLoS genetics*, 4(2):e1000008, 2008.
- Edith Hofer, Gennady V Roshchupkin, Hieab HH Adams, Maria J Knol, Honghuang Lin, Shuo Li, Habil Zare, Shahzad Ahmad, Nicola J Armstrong, Claudia L Satizabal, et al. Genetic determinants of cortical structure (thickness, surface area and volumes) among disease free adults in the charge consortium. *BioRxiv*, page 409649, 2018.
- Sam Horng, Gabriel Kreiman, Charlene Ellsworth, Damon Page, Marissa Blank, Kathleen Millen, and Mriganka Sur. Differential gene expression in the developing lateral geniculate nucleus and medial geniculate nucleus reveals novel roles for *zic4* and *foxp2* in visual and auditory pathway development. *Journal of Neuroscience*, 29(43):13672–13683, 2009.
- Ging-Yuek R Hsiung, Bonnie J Kaplan, Tracey L Petryshen, Shao Lu, and L Leigh Field. A dyslexia susceptibility locus (*dyx7*) linked to dopamine d4 receptor (*drd4*) region on chromosome 11p15. 5. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 125(1):112–119, 2004.
- Rebecca L Jackson, Paul Hoffman, Gorana Pobric, and Matthew A Lambon Ralph. The semantic network at work and rest: differential connectivity of anterior temporal lobe subregions. *Journal of Neuroscience*, 36(5):1490–1501, 2016.
- Kaja K Jasińska, Peter J Molfese, Sergey A Kornilov, W Einar Mencl, Stephen J Frost, Maria Lee, Kenneth R Pugh, Elena L Grigorenko, and Nicole Landi. The *bdnf* val66met polymorphism influences reading ability and patterns of neural activation in children. *PloS one*, 11(8): e0157449, 2016.
- Mark Jenkinson, Peter Bannister, Michael Brady, and Stephen Smith. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, 17(2):825–841, 2002.
- Heidi Johansen-Berg and Timothy EJ Behrens. *Diffusion MRI: from quantitative measurement to in vivo neuroanatomy*. Academic Press, 2013.
- Matthew B Johnson, Yuka Imamura Kawasawa, Christopher E Mason, Željka Krsnik, Giovanni Coppola, Darko Bogdanović, Daniel H Geschwind, Shrikant M Mane, Matthew W State, and Nenad Šestan. Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron*, 62(4):494–509, 2009.
- Marc Joliot, Gaël Jobard, Mikaël Naveau, Nicolas Delcroix, Laurent Petit, Laure Zago, Fabrice Crivello, Emmanuel Mellet, Bernard Mazoyer, and Nathalie Tzourio-Mazoyer. Aicha: An

- atlas of intrinsic connectivity of homotopic areas. *Journal of neuroscience methods*, 254: 46–59, 2015.
- O Parker Jones, NL Voets, JE Adcock, R Stacey, and S Jbabdi. Resting connectivity predicts task activation in pre-surgical populations. *NeuroImage: Clinical*, 13:378–385, 2017.
- Hanna Julienne, Pierre Lechat, Vincent Guillemot, Carla Lasry, Chunzi Yao, Robinson Araud, Vincent Laville, Bjarni Vilhjalmsson, Hervé Ménager, and Hugues Aschard. Jass: Command line and web interface for the joint analysis of gwas results. *NAR genomics and bioinformatics*, 2(1):lqaa003, 2020.
- Nina Kaminen, Katariina Hannula-Jouppi, Marjo Kestilä, Päivi Lahermo, Kurt Muller, M Kaaranen, B Myllyluoma, Arja Voutilainen, H Lyytinen, Jaana Nopola-Hemmi, et al. A genome scan for developmental dyslexia confirms linkage to chromosome 2p11 and suggests a new locus on 7q32. *Journal of medical genetics*, 40(5):340–345, 2003.
- DE Kaplan, J Gayan, J Ahn, T-W Won, D Pauls, RK Olson, JC DeFries, F Wood, BF Pennington, GP Page, et al. Evidence for linkage and association with reading disability, on 6p21.3-22. *The American Journal of Human Genetics*, 70(5):1287–1298, 2002.
- Clare Kelly, Lucina Q Uddin, Zarrar Shehzad, Daniel S Margulies, F Xavier Castellanos, Michael P Milham, and Michael Petrides. Broca’s region: linking human brain functional connectivity data and non-human primate tracing anatomy studies. *European Journal of Neuroscience*, 32(3):383–398, 2010.
- Martin Kircher, Daniela M Witten, Preti Jain, Brian J O’Roak, Gregory M Cooper, and Jay Shendure. A general framework for estimating the relative pathogenicity of human genetic variants. *Nature genetics*, 46(3):310–315, 2014.
- Lambertus Klei, Diana Luca, Bernie Devlin, and Kathryn Roeder. Pleiotropy and principal components of heritability combine to increase power for association analysis. *Genetic Epidemiology: The Official Publication of the International Genetic Epidemiology Society*, 32(1): 9–19, 2008.
- Julian Klingbeil, Max Wawrzyniak, Anika Stockert, and Dorothee Saur. Resting-state functional connectivity: An emerging method for the study of language networks in post-stroke aphasia. *Brain and Cognition*, 131:22–33, 2019.
- Chiung-Yuan Ko, Yu-Yi Chu, Shuh Narumiya, Jih-Ying Chi, Tomoyuki Furuyashiki, Tomohiro Aoki, Shao-Ming Wang, Wen-Chang Chang, and Ju-Ming Wang. The ccaat/enhancer-binding protein delta/mir135a/thrombospondin 1 axis mediates pge2-induced angiogenesis in alzheimer’s disease. *Neurobiology of aging*, 36(3):1356–1368, 2015.
- Michihiko Koeda, Atsushi Watanabe, Kumiko Tsuda, Miwako Matsumoto, Yumiko Ikeda, Woochan Kim, Amane Tateno, Banyar Than Naing, Hiroyuki Karibe, Takashi Shimada, et al. Interaction effect between handedness and cntnap2 polymorphism (rs7794745 genotype) on

- voice-specific frontotemporal activity in healthy individuals: an fmri study. *Frontiers in behavioral neuroscience*, 9:87, 2015.
- Maki S Koyama, Clare Kelly, Zarrar Shehzad, Deepak Penesetti, F Xavier Castellanos, and Michael P Milham. Reading networks at rest. *Cerebral cortex*, 20(11):2549–2559, 2010.
- Maki S Koyama, Adriana Di Martino, Xi-Nian Zuo, Clare Kelly, Maarten Mennes, Devika R Jutagir, F Xavier Castellanos, and Michael P Milham. Resting-state functional connectivity indexes reading competence in children and adults. *Journal of Neuroscience*, 31(23):8617–8624, 2011.
- Veena Kumar, Paula L Croxson, and Kristina Simonyan. Structural organization of the laryngeal motor cortical network and its implication for evolution of speech production. *Journal of Neuroscience*, 36(15):4170–4181, 2016.
- Loic Labache, M Joliot, Jérôme Saracco, Gaël Jobard, I Hesling, Laure Zago, Emmanuel Mellet, Laurent Petit, Fabrice Crivello, B Mazoyer, et al. A sentence supramodal areas atlas (sensaas) based on multiple task-induced activation mapping and graph analysis of intrinsic connectivity in 144 healthy right-handers. *Brain Structure and Function*, 224(2):859–882, 2019.
- Satu Lamminmäki, Satu Massinen, Jaana Nopola-Hemmi, Juha Kere, and Riitta Hari. Human robo1 regulates interaural interaction in auditory pathways. *Journal of Neuroscience*, 32(3):966–971, 2012.
- Nicole Landi and Meaghan V Perdue. Neuroimaging genetics studies of specific reading disability and developmental language disorder: A review. *Language and linguistics compass*, 13(9):e12349, 2019.
- Nicole Landi, Stephen J Frost, W Einar Mencl, Jonathan L Preston, Leslie K Jacobsen, Maria Lee, Carolyn Yrigollen, Kenneth R Pugh, and Elena L Grigorenko. The comt val/met polymorphism is associated with reading-related skills and consistent patterns of functional neural activation. *Developmental science*, 16(1):13–23, 2013.
- Ellen F Lau, Colin Phillips, and David Poeppel. A cortical network for semantics:(de) constructing the n400. *Nature reviews neuroscience*, 9(12):920–933, 2008.
- Isobel D Lawrenson, Sabine H Wimmer-Kleikamp, Peter Lock, Simone M Schoenwaelder, Michelle Down, Andrew W Boyd, Paul F Alewood, and Martin Lackmann. Ephrin-a5 induces rounding, blebbing and de-adhesion of epha3-expressing 293t and melanoma cells by crkii and rho-mediated signalling. *Journal of cell science*, 115(5):1059–1072, 2002.
- Yann Le Guen, Marie Amalric, Philippe Pinel, Christophe Pallier, and Vincent Frouin. Shared genetic aetiology between cognitive performance and brain activations in language and math tasks. *Scientific reports*, 8(1):1–11, 2018.
- Yann Le Guen, Slim Karkar, Antoine Grigis, Cathy Philippe, Jean-François Mangin, and Vincent Frouin. Heritability of surface area and cortical thickness: a comparison between the human

- connectome project and the uk biobank dataset. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 1887–1890. IEEE, 2019.
- Yann Le Guen, François Leroy, Cathy Philippe, IMAGEN Consortium, Jean-François Mangin, Ghislaine Dehaene-Lambertz, and Vincent Frouin. Enhancer Locus in ch14q23.1 Modulates Brain Asymmetric Temporal Regions Involved in Language Processing. *Cerebral Cortex*, 05 2020.
- Olivier Ledoit and Michael Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of multivariate analysis*, 88(2):365–411, 2004.
- Megan H Lee, Christopher D Smyser, and Joshua S Shimony. Resting-state fmri: a review of methods and clinical applications. *American Journal of neuroradiology*, 34(10):1866–1872, 2013.
- Sang Hong Lee, Jian Yang, Michael E Goddard, Peter M Visscher, and Naomi R Wray. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics*, 28(19):2540–2542, 2012.
- Jean-Michel Lemée, David Hassanein Berro, Florian Bernard, Eva Chinier, Louis-Marie Leiber, Philippe Menei, and Aram Ter Minassian. Resting-state functional magnetic resonance imaging versus task-based activity for language mapping and correlation with perioperative cortical mapping. *Brain and behavior*, 9(10):e01362, 2019.
- Jason P Lerch, André JW van der Kouwe, Armin Raznahan, Tomáš Paus, Heidi Johansen-Berg, Karla L Miller, Stephen M Smith, Bruce Fischl, and Stamatis N Sotiropoulos. Studying neuroanatomy using mri. *Nature neuroscience*, 20(3):314–326, 2017.
- Zhi-Pei Liang and Paul C Lauterbur. *Principles of magnetic resonance imaging*. SPIE Optical Engineering Press Bellingham, WA, 2000.
- Alvin M Liberman and Ignatius G Mattingly. The motor theory of speech perception revised. *Cognition*, 21(1):1–36, 1985.
- Frédérique Liégeois, Torsten Baldeweg, Alan Connelly, David G Gadian, Mortimer Mishkin, and Faraneh Vargha-Khadem. Language fmri abnormalities associated with foxp2 gene mutation. *Nature neuroscience*, 6(11):1230–1237, 2003.
- Jake Lin, Rubina Tabassum, Samuli Ripatti, and Matti Pirinen. Metaphat: Detecting and decomposing multivariate associations from univariate genome-wide association statistics. *Frontiers in genetics*, 11:431, 2020.
- Nikos K Logothetis. What we can do and what we cannot do with fmri. *Nature*, 453(7197):869–878, 2008.

- Junfeng Lu, Han Zhang, NU Farrukh Hameed, Jie Zhang, Shiwen Yuan, Tianming Qiu, Dinggang Shen, and Jinsong Wu. An automated method for identifying an independent component analysis-based language-related resting-state network in brain tumor subjects for surgical planning. *Scientific reports*, 7(1):1–16, 2017.
- Lina Lu, Hui Guo, Yu Peng, Guanglei Xun, Yanling Liu, Zhimin Xiong, Di Tian, Yalan Liu, Wei Li, Xiaojuan Xu, et al. Common and rare variants of the *thbs1* gene associated with the risk for autism. *Psychiatric genetics*, 24(6):235–240, 2014.
- Zhenjie Lu and Jonathan Kipnis. Thrombospondin 1—a key astrocyte-derived neurogenic factor. *The FASEB Journal*, 24(6):1925–1934, 2010.
- Michelle Luciano, David M Evans, Narelle K Hansell, Sarah E Medland, Grant W Montgomery, Nicholas G Martin, Margaret J Wright, and Timothy C Bates. A genome-wide association study for reading and language abilities in two population cohorts. *Genes, Brain and Behavior*, 12(6):645–652, 2013.
- Rizzi Luigi. *Complexité des structures linguistiques, simplicité des mécanismes du langage*. Fayard, 2021.
- Reedik Mägi and Andrew P Morris. Gwama: software for genome-wide association meta-analysis. *BMC bioinformatics*, 11(1):1–6, 2010.
- Teri A Manolio, Francis S Collins, Nancy J Cox, David B Goldstein, Lucia A Hindorff, David J Hunter, Mark I McCarthy, Erin M Ramos, Lon R Cardon, Aravinda Chakravarti, et al. Finding the missing heritability of complex diseases. *Nature*, 461(7265):747–753, 2009.
- Jonathan Marchini, Lon R Cardon, Michael S Phillips, and Peter Donnelly. The effects of human population structure on large genetic association studies. *Nature genetics*, 36(5):512–517, 2004.
- Jonathan Marchini, Bryan Howie, Simon Myers, Gil McVean, and Peter Donnelly. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nature genetics*, 39(7):906–913, 2007.
- Cecilia Marino, Paola Scifo, Pasquale A Della Rosa, Sara Mascheretti, Andrea Facoetti, Maria L Lorusso, Roberto Giorda, Monica Consonni, Andrea Falini, Massimo Molteni, et al. The *dc2*/intron 2 deletion and white matter disorganization: focus on developmental dyslexia. *Cortex*, 57:227–243, 2014.
- Guillaume Marrelec, Alexandre Krainik, Hugues Duffau, Mélanie Péligrini-Issac, Stéphane Lehericy, Julien Doyon, and Habib Benali. Partial correlation for functional brain interactivity investigation in functional mri. *Neuroimage*, 32(1):228–237, 2006.
- Shashwath A Meda, Joel Gelernter, Jeffrey R Gruen, Vince D Calhoun, Haiying Meng, Natalie A Cope, and Godfrey D Pearlson. Polymorphism of *dc2* reveals differences in cortical morphology of healthy individuals—a preliminary voxel based morphometry study. *Brain imaging and behavior*, 2(1):21–26, 2008.

- Arthur Mensch, Gaël Varoquaux, and Bertrand Thirion. Compressed online dictionary learning for fast resting-state fmri decomposition. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 1282–1285. IEEE, 2016.
- M-Marsel Mesulam. Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 28(5):597–613, 1990.
- M-Marsel Mesulam, Cynthia K Thompson, Sandra Weintraub, and Emily J Rogalski. The wernicke conundrum and the anatomy of language comprehension in primary progressive aphasia. *Brain*, 138(8):2423–2437, 2015.
- Jeremy A Miller, Song-Lin Ding, Susan M Sunkin, Kimberly A Smith, Lydia Ng, Aaron Szafer, Amanda Ebbert, Zackery L Riley, Joshua J Royall, Kaylynn Aiona, et al. Transcriptional landscape of the prenatal human brain. *Nature*, 508(7495):199–206, 2014.
- Karla L Miller, Fidel Alfaro-Almagro, Neal K Bangerter, David L Thomas, Essa Yacoub, Junqian Xu, Andreas J Bartsch, Saad Jbabdi, Stamatiou N Sotiropoulos, Jesper LR Andersson, et al. Multimodal population brain imaging in the uk biobank prospective epidemiological study. *Nature neuroscience*, 19(11):1523–1536, 2016.
- Grigorios Nasios, Efthymios Dardiotis, and Lambros Messinis. From broca and wernicke to the neuromodulation era: insights of brain language networks for neurorehabilitation. *Behavioural neurology*, 2019, 2019.
- Dianne F Newbury, Laura Winchester, Laura Addis, Silvia Paracchini, Lyn-Louise Buckingham, Ann Clark, Wendy Cohen, Hilary Cowie, Katharina Dworzynski, Andrea Everitt, et al. Cmp1 and atp2c2 modulate phonological short-term memory in language impairment. *The American Journal of Human Genetics*, 85(2):264–272, 2009.
- Guiyan Ni, Gerhard Moser, Stephan Ripke, Benjamin M Neale, Aiden Corvin, James TR Walters, Kai-How Farh, Peter A Holmans, Phil Lee, Brendan Bulik-Sullivan, et al. Estimation of genetic correlation via linkage disequilibrium score regression and genomic restricted maximum likelihood. *The American Journal of Human Genetics*, 102(6):1185–1194, 2018.
- Nobuyuki Nishitani and Riitta Hari. Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences*, 97(2):913–918, 2000.
- Uta Noppeney and Cathy J Price. Retrieval of abstract semantics. *Neuroimage*, 22(1):164–170, 2004.
- Ron Nudel, Nuala H Simpson, Gillian Baird, Anne O’Hare, Gina Conti-Ramsden, Patrick F Bolton, Elizabeth R Hennessy, SLI Consortium, SM Ring, G Davey Smith, et al. Genome-wide association analyses of child genotype effects and parent-of-origin effects in specific language impairment. *Genes, Brain and Behavior*, 13(4):418–429, 2014.

- Dale R Nyholt. Seca: Snp effect concordance analysis using genome-wide association summary results. *Bioinformatics*, 30(14):2086–2088, 2014.
- Lauren J O’Donnell and Carl-Fredrik Westin. An introduction to diffusion tensor image analysis. *Neurosurgery Clinics*, 22(2):185–196, 2011.
- Paul F O’Reilly, Clive J Hoggart, Yotsawat Pomyen, Federico CF Calboli, Paul Elliott, Marjo-Riitta Jarvelin, and Lachlan JM Coin. Multiphen: joint model of multiple phenotypes can increase discovery in gwas. *PLoS one*, 7(5):e34861, 2012.
- Christophe Pallier, Anne-Dominique Devauchelle, and Stanislas Dehaene. Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences*, 108(6):2522–2527, 2011.
- Anand S Pandit, Gareth Ball, A David Edwards, and Serena J Counsell. Diffusion magnetic resonance imaging in preterm brain injury. *Neuroradiology*, 55(2):65–95, 2013.
- Hae Jeong Park, Su Kang Kim, Jong Woo Kim, Won Sub Kang, and Joo-Ho Chung. Association of thrombospondin 1 gene with schizophrenia in korean population. *Molecular biology reports*, 39(6):6875–6880, 2012.
- Ki Yun Park, John J Lee, Donna Dierker, Laura M Marple, Carl D Hacker, Jarod L Roland, Daniel S Marcus, Mikhail Milchenko, Michelle M Miller-Thomas, Tammie L Benzinger, et al. Mapping language function with task-based vs. resting-state functional mri. *PLoS one*, 15(7):e0236423, 2020.
- Elena B Pasquale. Eph-ephrin bidirectional signaling in physiology and disease. *Cell*, 133(1):38–52, 2008.
- Meaghan V Perdue, Sara Mascheretti, Sergey A Kornilov, Kaja K Jasińska, Kayleigh Ryherd, W Einar Mencl, Stephen J Frost, Elena L Grigorenko, Kenneth R Pugh, and Nicole Landi. Common variation within the setbp1 gene is associated with reading-related skills and patterns of functional neural activation. *Neuropsychologia*, 130:44–51, 2019.
- Robin L Peterson and Bruce F Pennington. Developmental dyslexia. *Annual review of clinical psychology*, 11:283–307, 2015.
- Tracey L Petryshen, Bonnie J Kaplan, Ming Fu Liu, Norma Schmill de French, Rose Tobias, Martha L Hughes, and L Leigh Field. Evidence for a susceptibility locus on chromosome 6q influencing phonological coding dyslexia. *American journal of Medical genetics*, 105(6):507–517, 2001.
- KC Sreedharan Pillai. Some new test criteria in multivariate analysis. *The Annals of Mathematical Statistics*, pages 117–121, 1955.
- Philippe Pinel, Fabien Fauchereau, Antonio Moreno, Alexis Barbot, Mark Lathrop, Diana Zelenika, Denis Le Bihan, Jean-Baptiste Poline, Thomas Bourgeron, and Stanislas Dehaene.

- Genetic variants of *foxp2* and *kiaa0319/ttrap/them2* locus are associated with altered brain activation in distinct language-related regions. *Journal of Neuroscience*, 32(3):817–825, 2012.
- Nicholas A Popp, Dianke Yu, Bridgett Green, Emily Y Chew, Baitang Ning, Chi-Chao Chan, and Jingsheng Tuo. Functional single nucleotide polymorphism in *il-17a* 3' untranslated region is targeted by *mi-r-4480* in vitro and may be associated with age-related macular degeneration. *Environmental and molecular mutagenesis*, 57(1):58–64, 2016.
- Heather F Porter and Paul F O'Reilly. Multivariate simulation framework reveals performance of multi-trait gwas methods. *Scientific reports*, 7(1):1–12, 2017.
- Jonathan D Power, Alexander L Cohen, Steven M Nelson, Gagan S Wig, Kelly Anne Barnes, Jessica A Church, Alecia C Vogel, Timothy O Laumann, Fran M Miezin, Bradley L Schlaggar, et al. Functional network organization of the human brain. *Neuron*, 72(4):665–678, 2011.
- Leslie Pray. Discovery of dna structure and function: Watson and crick. *Nature Education*, 1(1), 2008.
- Shaun Purcell. Variance components models for gene–environment interaction in twin analysis. *Twin Research and Human Genetics*, 5(6):554–571, 2002.
- Shaun Purcell, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel AR Ferreira, David Bender, Julian Maller, Pamela Sklar, Paul IW De Bakker, Mark J Daly, et al. Plink: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics*, 81(3):559–575, 2007.
- Ting Qi, Yang Wu, Jian Zeng, Futao Zhang, Angli Xue, Longda Jiang, Zhihong Zhu, Kathryn Kemper, Loic Yengo, Zhili Zheng, et al. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. *Nature communications*, 9(1):1–12, 2018.
- Marcus E Raichle, Ann Mary MacLeod, Abraham Z Snyder, William J Powers, Debra A Gusnard, and Gordon L Shulman. A default mode of brain function. *Proceedings of the National Academy of Sciences*, 98(2):676–682, 2001.
- Amy E Ramage, Semra Aytur, and Kirrie J Ballard. Resting-state functional magnetic resonance imaging connectivity between semantic and phonological regions of interest may inform language targets in aphasia. *Journal of Speech, Language, and Hearing Research*, 63(9):3051–3067, 2020.
- Vasily Ramensky, Peer Bork, and Shamil Sunyaev. Human non-synonymous snps: server and survey. *Nucleic acids research*, 30(17):3894–3900, 2002.
- J Steven Reznick, Robin Corley, JoAnn Robinson, and Adam P Matheny Jr. A longitudinal twin study of intelligence in the second year. *Monographs of the society for research in child development*, pages i–160, 1997.
- Sally-Ann Rhea, Josh B Bricker, Sally J Wadsworth, and Robin P Corley. The colorado adoption project. *Twin Research and Human Genetics*, 16(1):358–365, 2013.

- Stuart J Ritchie, David Alexander Dickie, Simon R Cox, Maria del C Valdes Hernandez, Janie Corley, Natalie A Royle, Alison Pattie, Benjamin S Aribisala, Paul Redmond, Susana Muñoz Maniega, et al. Brain volumetric changes and cognitive ageing during the eighth decade of life. *Human brain mapping*, 36(12):4910–4925, 2015.
- D Roeske, KU Ludwig, N Neuhoff, J Becker, J Bartling, J Bruder, FF Brockschmidt, A Warnke, H Remschmidt, P Hoffmann, et al. First genome-wide association scan on neurophysiological endophenotypes points to trans-regulation effects on *slc2a3* in dyslexic children. *Molecular psychiatry*, 16(1):97–107, 2011.
- K Rojkova, E Volle, M Urbanski, F Humbert, F Dell’Acqua, and M Thiebaut De Schotten. Atlasing the frontal lobe connections and their variability due to age and education: a spherical deconvolution tractography study. *Brain Structure and Function*, 221(3):1751–1766, 2016.
- Jerker Rönnerberg, Mary Rudner, and Martin Ingvar. Neural correlates of working memory for sign language. *Cognitive Brain Research*, 20(2):165–182, 2004.
- Alexander Rudov, Marco Bruno Luigi Rocchi, Augusto Accorsi, Giorgio Spada, Antonio Domenico Procopio, Fabiola Olivieri, Maria Rita Rippo, and Maria Cristina Albertini. Putative mirnas for the diagnosis of dyslexia, dyspraxia, and specific language impairment. *Epigenetics*, 8(10):1023–1029, 2013.
- Gholamreza Salimi-Khorshidi, Gwenaëlle Douaud, Christian F Beckmann, Matthew F Glasser, Ludovica Griffanti, and Stephen M Smith. Automatic denoising of functional mri data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage*, 90: 449–468, 2014.
- Jeanne E Savage, Philip R Jansen, Sven Stringer, Kyoko Watanabe, Julien Bryois, Christiaan A De Leeuw, Mats Nagel, Swapnil Awasthi, Peter B Barr, Jonathan RI Coleman, et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature genetics*, 50(7):912–919, 2018.
- Thomas S Scerri, Fahimeh Darki, Dianne F Newbury, Andrew JO Whitehouse, Myriam Peyrard-Janvid, Hans Matsson, Qi W Ang, Craig E Pennell, Susan Ring, John Stein, et al. The dyslexia candidate locus on 2p12 is associated with general cognitive ability and white matter structure. *PloS one*, 7(11):e50321, 2012.
- Benjamin J Schmiedel, Divya Singh, Ariel Madrigal, Alan G Valdovino-Gonzalez, Brandie M White, Jose Zapardiel-Gonzalo, Brendan Ha, Gokmen Altay, Jason A Greenbaum, Graham McVicker, et al. Impact of genetic polymorphisms on human immune cell gene expression. *Cell*, 175(6):1701–1715, 2018.
- Jean-Luc Schwartz, Marc Sato, and Luciano Fadiga. The common language of speech perception and action: a neurocognitive perspective. *Revue française de linguistique appliquée*, 13(2): 9–22, 2008.

- Jean-Luc Schwartz, Anahita Basirat, Lucie Ménard, and Marc Sato. The perception-for-action-control theory (pact): A perceptuo-motor theory of speech perception. *Journal of Neurolinguistics*, 25(5):336–354, 2012.
- Zhiqiang Sha, Dick Schijven, Amaia Carrion-Castillo, Marc Joliot, Bernard Mazoyer, Simon E Fisher, Fabrice Crivello, and Clyde Francks. The genetic architecture of structural left-right asymmetry of the human brain. *Nature Human Behaviour*, pages 1–14, 2021.
- Stephen T Sherry, M-H Ward, M Kholodov, J Baker, Lon Phan, Elizabeth M Smigielski, and Karl Sirotkin. dbsnp: the ncbi database of genetic variation. *Nucleic acids research*, 29(1): 308–311, 2001.
- Chet C Sherwood, Francys Subiaul, and Tadeusz W Zawidzki. A natural history of the human mind: tracing evolutionary changes in brain and cognition. *Journal of anatomy*, 212(4): 426–454, 2008.
- Michael A Skeide, Holger Kirsten, Indra Kraft, Gesa Schaadt, Bent Müller, Nicole Neef, Jens Brauer, Arndt Wilcke, Frank Emmrich, Johannes Boltze, et al. Genetic dyslexia risk variant is related to neural connectivity patterns underlying phonological awareness in children. *Neuroimage*, 118:414–421, 2015.
- Michael A Skeide, Indra Kraft, Bent Müller, Gesa Schaadt, Nicole E Neef, Jens Brauer, Arndt Wilcke, Holger Kirsten, Johannes Boltze, and Angela D Friederici. Nrsn1 associated grey matter volume of the visual word form area reveals dyslexia before school. *Brain*, 139(10): 2792–2803, 2016.
- Stephen M Smith, Mark Jenkinson, Heidi Johansen-Berg, Daniel Rueckert, Thomas E Nichols, Clare E Mackay, Kate E Watkins, Olga Ciccarelli, M Zaheer Cader, Paul M Matthews, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage*, 31(4):1487–1505, 2006.
- Stephen M Smith, Peter T Fox, Karla L Miller, David C Glahn, P Mickle Fox, Clare E Mackay, Nicola Filippini, Kate E Watkins, Roberto Toro, Angela R Laird, et al. Correspondence of the brain’s functional architecture during activation and rest. *Proceedings of the national academy of sciences*, 106(31):13040–13045, 2009.
- Stephen M Smith, Karla L Miller, Gholamreza Salimi-Khorshidi, Matthew Webster, Christian F Beckmann, Thomas E Nichols, Joseph D Ramsey, and Mark W Woolrich. Network modelling methods for fmri. *Neuroimage*, 54(2):875–891, 2011.
- Beate St Pourcain, Rolieke AM Cents, Andrew JO Whitehouse, Claire MA Haworth, Oliver SP Davis, Paul F O’Reilly, Susan Roulstone, Yvonne Wren, Qi W Ang, Fleur P Velders, et al. Common variation near robo2 is associated with expressive vocabulary in infancy. *Nature communications*, 5(1):1–9, 2014.
- Matthew Stephens. A unified framework for association analysis with multiple related phenotypes. *PloS one*, 8(7):e65245, 2013.

- W Dale Stevens, Dwight J Kravitz, Cynthia S Peng, Michael Henry Tessler, and Alex Martin. Privileged functional connectivity between the visual word form area and the language system. *Journal of Neuroscience*, 37(21):5288–5297, 2017.
- Karin Stromswold. The heritability of language: A review and metaanalysis of twin, adoption, and linkage studies. *Language*, pages 647–723, 2001.
- Aravind Subramanian, Pablo Tamayo, Vamsi K Mootha, Sayan Mukherjee, Benjamin L Ebert, Michael A Gillette, Amanda Paulovich, Scott L Pomeroy, Todd R Golub, Eric S Lander, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43):15545–15550, 2005.
- Cathie Sudlow, John Gallacher, Naomi Allen, Valerie Beral, Paul Burton, John Danesh, Paul Downey, Paul Elliott, Jane Green, Martin Landray, et al. 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(3):e1001779, 2015.
- D Szklarczyk et al. The string database in 2017: quality-controlled protein–protein association networks, made broadly accessible nucleic acids res., 45. *D362–D368*, 2017.
- Mikko Taipale, Nina Kaminen, Jaana Nopola-Hemmi, Tuomas Haltia, Birgitta Myllyluoma, Heikki Lyytinen, Kurt Muller, Minna Kaaranen, Perttu J Lindsberg, Katariina Hannula-Jouppi, et al. A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide repeat domain protein dynamically regulated in brain. *Proceedings of the National Academy of Sciences*, 100(20):11553–11558, 2003.
- Geoffrey CY Tan, Thomas F Doke, John Ashburner, Nicholas W Wood, and Richard SJ Frackowiak. Normal variation in fronto-occipital circuitry and cerebellar structure with an autism-associated polymorphism of *cntnap2*. *Neuroimage*, 53(3):1030–1042, 2010.
- I Tavor, O Parker Jones, RB Mars, SM Smith, TE Behrens, and S Jbabdi. Task-free mri predicts individual differences in brain activity during task performance. *Science*, 352(6282):216–220, 2016.
- Peter E Thijssen, Yuya Ito, Giacomo Grillo, Jun Wang, Guillaume Velasco, Hirohisa Nitta, Motoko Unoki, Minako Yoshihara, Mikita Suyama, Yu Sun, et al. Mutations in *cdca7* and *hells* cause immunodeficiency–centromeric instability–facial anomalies syndrome. *Nature communications*, 6(1):1–8, 2015.
- Bertrand Thirion, Gaël Varoquaux, Elvis Dohmatob, and Jean-Baptiste Poline. Which fmri clustering gives good brain parcellations? *Frontiers in neuroscience*, 8:167, 2014.
- Moriah E Thomason and Paul M Thompson. Diffusion imaging, white matter, and psychopathology. *Annual review of clinical psychology*, 7:63–85, 2011.

- Paul M Thompson, Kiralee M Hayashi, Rebecca A Dutton, Ming-Chang Chiang, Alex D Leow, Elizabeth R Sowell, Greig De Zubicaray, James T Becker, Oscar L Lopez, Howard J Aizenstein, et al. Tracking alzheimer's disease. *Annals of the New York Academy of Sciences*, 1097:183, 2007.
- Paul M Thompson, Tian Ge, David C Glahn, Neda Jahanshad, and Thomas E Nichols. Genetics of the connectome. *Neuroimage*, 80:475–488, 2013.
- Yanmei Tie, Laura Rigolo, Isaiah H Norton, Raymond Y Huang, Wentao Wu, Daniel Orringer, Srinivasan Mukundan Jr, and Alexandra J Golby. Defining language networks from resting-state fmri for surgical planning—a feasibility study. *Human brain mapping*, 35(3):1018–1030, 2014.
- Pascale Tremblay and Anthony Steven Dick. Broca and wernicke are dead, or moving past the classic model of language neurobiology. *Brain and language*, 162:60–71, 2016.
- Patrick Turley, Raymond K Walters, Omeed Maghzian, Aysu Okbay, James J Lee, Mark Alan Fontana, Tuan Anh Nguyen-Viet, Robbee Wedow, Meghan Zacher, Nicholas A Furlotte, et al. Multi-trait analysis of genome-wide association summary statistics using mtag. *Nature genetics*, 50(2):229–237, 2018.
- Travis H Turner, Julius Fridriksson, Julie Baker, David Eoute Jr, Leonardo Bonilha, and Christopher Rorden. Obligatory broca's area modulation associated with passive speech perception. *Neuroreport*, 20(5):492, 2009.
- Nathalie Tzourio-Mazoyer, Brigitte Landeau, Dimitri Papathanassiou, Fabrice Crivello, Olivier Etard, Nicolas Delcroix, Bernard Mazoyer, and Marc Joliot. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage*, 15(1):273–289, 2002.
- Julia Uddén, Tineke M Snijders, Simon E Fisher, and Peter Hagoort. A common variant of the cntnap2 gene is associated with structural variation in the left superior occipital gyrus. *Brain and language*, 172:16–21, 2017.
- Sven J van der Lee, Gennady V Roshchupkin, Hieab HH Adams, Helena Schmidt, Edith Hofer, Yasaman Saba, Reinhold Schmidt, Albert Hofman, Najaf Amin, Cornelia M van Duijn, et al. Gray matter heritability in family-based and population-based studies using voxel-based morphometry. *Human brain mapping*, 38(5):2408–2423, 2017.
- Dennis van der Meer, Oleksandr Frei, Tobias Kaufmann, Alexey A Shadrin, Anna Devor, Olav B Smeland, Wesley K Thompson, Chun Chieh Fan, Dominic Holland, Lars T Westlye, et al. Understanding the genetic determinants of the brain with mostest. *Nature communications*, 11(1):1–9, 2020.
- Sophie Van der Sluis, Danielle Posthuma, and Conor V Dolan. Tates: efficient multivariate genotype-phenotype analysis for genome-wide association studies. *PLoS genetics*, 9(1): e1003235, 2013.

- David C Van Essen, Kamil Ugurbil, Edward Auerbach, Deanna Barch, Timothy EJ Behrens, Richard Bucholz, Acer Chang, Liyong Chen, Maurizio Corbetta, Sandra W Curtiss, et al. The human connectome project: a data acquisition perspective. *Neuroimage*, 62(4):2222–2231, 2012.
- David C Van Essen, Stephen M Smith, Deanna M Barch, Timothy EJ Behrens, Essa Yacoub, Kamil Ugurbil, Wu-Minn HCP Consortium, et al. The wu-minn human connectome project: an overview. *Neuroimage*, 80:62–79, 2013.
- Faraneh Vargha-Khadem, Kate E Watkins, CJ Price, John Ashburner, Katie J Alcock, Alan Connelly, Richard SJ Frackowiak, Karl J Friston, ME Pembrey, Mortimer Mishkin, et al. Neural basis of an inherited speech and language disorder. *Proceedings of the National Academy of Sciences*, 95(21):12695–12700, 1998.
- Francesco Vergani, Luis Lacerda, Juan Martino, Johannes Attems, Christopher Morris, Patrick Mitchell, Michel Thiebaut de Schotten, and Flavio Dell’Acqua. White matter connections of the supplementary motor area in humans. *Journal of Neurology, Neurosurgery & Psychiatry*, 85(12):1377–1385, 2014.
- Sonja C Vernes, Dianne F Newbury, Brett S Abrahams, Laura Winchester, Jérôme Nicod, Matthias Groszer, Maricela Alarcón, Peter L Oliver, Kay E Davies, Daniel H Geschwind, et al. A functional genetic link between distinct developmental language disorders. *New England Journal of Medicine*, 359(22):2337–2345, 2008.
- Mathieu Vigneau, Virginie Beaucousin, Pierre-Yves Herve, Hugues Duffau, Fabrice Crivello, Olivier Houde, Bernard Mazoyer, and Nathalie Tzourio-Mazoyer. Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage*, 30(4):1414–1432, 2006.
- Mathieu Vigneau, Virginie Beaucousin, Pierre-Yves Hervé, Gael Jobard, Laurent Petit, Fabrice Crivello, Emmanuel Mellet, Laure Zago, Bernard Mazoyer, and Nathalie Tzourio-Mazoyer. What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing?: Insights from a meta-analysis. *Neuroimage*, 54(1):577–593, 2011.
- Peter M Visscher, William G Hill, and Naomi R Wray. Heritability in the genomics era—concepts and misconceptions. *Nature reviews genetics*, 9(4):255–266, 2008.
- Urmo Vösa, Annique Claringbould, Harm-Jan Westra, Marc Jan Bonder, Patrick Deelen, Biao Zeng, Holger Kirsten, Ashis Saha, Roman Kreuzhuber, Silva Kasela, et al. Unraveling the polygenic architecture of complex traits using blood eqtl metaanalysis. *BioRxiv*, page 447367, 2018.
- Louise V Wain, Nick Shrine, Suzanne Miller, Victoria E Jackson, Ioanna Ntalla, María Soler Artigas, Charlotte K Billington, Abdul Kader Kheirallah, Richard Allen, James P Cook, et al. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (uk bileve): a genetic association study in uk biobank. *The Lancet Respiratory Medicine*, 3(10):769–781, 2015.

- Daifeng Wang, Shuang Liu, Jonathan Warrell, Hyejung Won, Xu Shi, Fabio CP Navarro, Declan Clarke, Mengting Gu, Prashant Emani, Yucheng T Yang, et al. Comprehensive functional genomic resource and integrative model for the human brain. *Science*, 362(6420), 2018.
- Kai Wang, Mingyao Li, and Hakon Hakonarson. Annovar: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic acids research*, 38(16):e164–e164, 2010.
- Kyoko Watanabe, Erdogan Taskesen, Arjen Van Bochoven, and Danielle Posthuma. Functional mapping and annotation of genetic associations with fuma. *Nature communications*, 8(1):1–11, 2017.
- Kate E Watkins, Faraneh Vargha-Khadem, John Ashburner, Richard E Passingham, Alan Connelly, Karl J Friston, Richard SJ Frackowiak, Mortimer Mishkin, and David G Gadian. Mri analysis of an inherited speech and language disorder: structural brain abnormalities. *Brain*, 125(3):465–478, 2002.
- DH Whalen. The motor theory of speech perception. In *Oxford Research Encyclopedia of Linguistics*. 2019.
- Heather C Whalley, Garret O’Connell, Jessika E Sussmann, Anna Peel, Andrew C Stanfield, Marianna E Hayiou-Thomas, Eve C Johnstone, Stephen M Lawrie, Andrew M McIntosh, and Jeremy Hall. Genetic variation in cntnap2 alters brain function during linguistic processing in healthy individuals. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 156(8):941–948, 2011.
- Arndt Wilcke, Carolin Ligges, Jana Burkhardt, Michael Alexander, Christiane Wolf, Elfi Quente, Peter Ahnert, Per Hoffmann, Albert Becker, Bertram Müller-Myhsok, et al. Imaging genetics of foxp2 in dyslexia. *European Journal of Human Genetics*, 20(2):224–229, 2012.
- Cristen J Willer, Yun Li, and Gonçalo R Abecasis. Metal: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, 26(17):2190–2191, 2010.
- Stephen M Wilson, Ayşe Pinar Saygin, Martin I Sereno, and Marco Iacoboni. Listening to speech activates motor areas involved in speech production. *Nature neuroscience*, 7(7):701–702, 2004.
- Wojciech Wiszniewski, Jill V Hunter, Neil A Hanchard, Jason R Willer, Chad Shaw, Qi Tian, Anna Illner, Xueqing Wang, Sau W Cheung, Ankita Patel, et al. Tm4sf20 ancestral deletion and susceptibility to a pediatric disorder of early language delay and cerebral white matter hyperintensities. *The American Journal of Human Genetics*, 93(2):197–210, 2013.
- Chong Wu. Multi-trait genome-wide analyses of the brain imaging phenotypes in uk biobank. *Genetics*, 215(4):947–958, 2020.
- Hua-Dong Xiang, Hubert M Fonteijn, David G Norris, and Peter Hagoort. Topographical functional connectivity pattern in the perisylvian language networks. *Cerebral cortex*, 20(3): 549–560, 2010.

- Jian Yang, Beben Benyamin, Brian P McEvoy, Scott Gordon, Anjali K Henders, Dale R Nyholt, Pamela A Madden, Andrew C Heath, Nicholas G Martin, Grant W Montgomery, et al. Common snps explain a large proportion of the heritability for human height. *Nature genetics*, 42(7):565–569, 2010.
- Jian Yang, S Hong Lee, Michael E Goddard, and Peter M Visscher. Gcta: a tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, 88(1):76–82, 2011.
- Jian Yang, Andrew Bakshi, Zhihong Zhu, Gibran Hemani, Anna AE Vinkhuyzen, Sang Hong Lee, Matthew R Robinson, John RB Perry, Ilja M Nolte, Jana V van Vliet-Ostaptchouk, et al. Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nature genetics*, 47(10):1114–1120, 2015.
- Hui Zhang, Torben Schneider, Claudia A Wheeler-Kingshott, and Daniel C Alexander. Noddi: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage*, 61(4):1000–1016, 2012.
- Rui Zhang, Nan-Nan Zhong, Xiao-Gang Liu, Han Yan, Chuan Qiu, Yan Han, Wei Wang, Wei-Kun Hou, Yue Liu, Cheng-Ge Gao, et al. Is the efnb2 locus associated with schizophrenia? single nucleotide polymorphisms and haplotypes analysis. *Psychiatry research*, 180(1):5–9, 2010.
- Daria V Zhernakova, Patrick Deelen, Martijn Vermaat, Maarten Van Itersen, Michiel Van Galen, Wibowo Arindrarto, Peter Van't Hof, Hailiang Mei, Freerk Van Dijk, Harm-Jan Westra, et al. Identification of context-dependent expression quantitative trait loci in whole blood. *Nature genetics*, 49(1):139–145, 2017.
- Wensheng Zhu and Heping Zhang. Why do we test multiple traits in genetic association studies? *Journal of the Korean Statistical Society*, 38(1):1–10, 2009.
- Or Zuk, Eliana Hechter, Shamil R Sunyaev, and Eric S Lander. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences*, 109(4):1193–1198, 2012.

* * *
* *
*

Titre : L'architecture génétique du connectome du langage dans le cerveau humain

Mots clés : imagerie-génétique, IRM fonctionnelle de repos, IRM de diffusion, Langage, GWAS multivariée, UK Biobank.

Résumé : Le langage est une composante essentielle de la vie humaine en raison de son rôle central dans les interactions sociales, la transmission culturelle, et la structuration de nos pensées. Grâce à l'étude des troubles du développement du langage parlé et écrit, tel que la dyslexie, un nombre croissant de gènes ont été identifiés et étiquetés comme étant impliqués dans leurs étiologies. Des études d'association cas-témoins sont mises en œuvre, mais souffrent d'un manque de puissance statistique et peinent à produire des résultats significatifs. Parallèlement à cela, des études de neuro-imagerie ont contribué de manière significative à la compréhension des aspects structurels et fonctionnels du langage dans le cerveau. La disponibilité récente des cohortes de grande taille comportant à la fois de la neuro-imagerie et de la génétique tels que "UK Biobank" et le "Human Connectome Project" a rendu possible l'étude du langage dans les populations générales via des phénotypes intermédiaires (endophénotypes). En effet, ceux-là donnent accès à des mesures biologiquement pertinentes de la variabilité individuelle des supports structurels et fonctionnels du langage et sont par conséquent adaptés à la recherche d'associations génétiques. Dans cette thèse, nous souhaitons d'une part, mettre en évidence la variation interindividuelle de la structure et/ou de la fonction cérébrale qui reflète des traits neurocognitifs spécifiques ; d'autre part, tirer profit de cette variabilité pour mettre en évidence la base génétique de leur architecture cérébrale. Pour ce faire, nous avons étudié la neurobiologie du langage à l'aide de la connectivité fonctionnelle (CF) des régions corticales périsylviennes mesurée à l'état de repos, ainsi qu'à partir de la connectivité anatomique (CA) mesurée sur des faisceaux de matière blanche déterminée par l'IRM de diffusion et connus comme étant liés au langage. Tout d'abord, nous présentons une étude approfondie de l'héritabilité du langage en exploitant des endophénotypes du connectome du langage hu-

main extraits des deux cohortes susmentionnées. Les résultats obtenus sont conformes aux estimations de l'héritabilité rapportées dans la littérature et confirment que l'organisation cérébrale liée au langage est en partie sous-tendue par la génétique. Deuxièmement, pour les endophénotypes de CF avec une héritabilité significative, nous avons réalisé une étude d'association multivariée pangénomique sur 32.186 participants de la cohorte UK Biobank. Vingt loci sont significativement associés aux CF du langage, dont trois sont répliqués dans un échantillon indépendant (ethnicité non britannique de UKB, N=4.754). L'annotation fonctionnelle approfondie pointe notamment vers le gène EPHA3 avec un rôle dans le réseau sémantique fronto-pariéto-temporal, et le gène THBS1 avec un rôle dans l'interaction perception-moteur, requise dans le langage. Enfin, de la même manière, pour les endophénotypes de CA avec une héritabilité significative, nous avons effectué une étude d'association multivariée pangénomique sur 31.775 participants de la cohorte UK Biobank. 278 loci sont significativement associés aux CA du langage, dont ceux précédemment associés à des endophénotypes de CF. En conclusion, ce travail a permis d'étudier les bases génétiques du langage humain en utilisant des approches génomiques à l'état de l'art et ce, en développant des endophénotypes originaux associés au langage. Les gènes suggérés par l'analyse fournissent des pistes d'investigation dans des modèles animaux ou cellulaires du langage. Aussi, alors que de plus en plus de cohortes sont construites en population générale, ce travail renforce notre conviction que ce type de cohortes "multi-phénotypées" permet d'augmenter la puissance statistique et de pallier les faibles tailles d'effet en neuro-imagerie afin d'étudier des questions ouvertes en neuroscience. De nombreux autres domaines des neurosciences pourraient probablement bénéficier de ce type de méthodologie.

Title : The genetic architecture of the human language connectome

Keywords : Imaging-genetics, Resting-state functional MRI, diffusion MRI, Language, multivariate GWAS, UK Biobank.

Abstract : Language is an essential component of human life because of its central role in social interaction, cultural transmission, and for structuring our thoughts. Through developmental disorders of spoken and written language studies such as dyslexia, a growing number of genes have been identified and labelled as implicated in language etiology. Case-control studies suffering from a lack of statistical power, struggle to uncover significant variants. In parallel, neuroimaging studies have significantly contributed to the understanding of structural and functional aspects of language in the human brain. The recent availability of large-scale cohorts comprising both neuroimaging and genetics data such as the UK Biobank and the Human Connectome Project, have made it possible to study language in the general population via image-derived phenotypes (IDP). Indeed, these IDPs give access to biologically relevant measurements of individual variability and are consequently suitable for the search of genetic associations. In this thesis, we aim to first, highlight the inter-individual variation in brain structure and/or function of language; second, take advantage of this variability to study the genetic basis of the brain's infrastructure for language. For this, we used task-free functional connectivity (FC), extracted from perisylvian cortical regions as well as from anatomical connectivity (AC) measured using well-known language related white matter bundles determined from diffusion MRI. First, we present an in-depth study of language heritability using human language connectome as endophenotypes extracted from the two aforementioned large-scale cohorts. The results found are consistent with the heritability estimates reported in the literature and sup-

port that the language-related brain organization is partly underpinned by genetic. Secondly, using the significantly heritable FC endophenotypes, we performed a multivariate genome-wide association study on 32,186 participants from UK Biobank. Twenty genomic loci were found significantly associated with language FCs, out of which three were replicated in an independent sample (the UK Biobank non-British sample, N=4.754). The functional annotation highlighted noticeably the EPHA3 gene associated with FCs of the fronto-parieto-temporal semantic network, and the THBS1 gene with a potential role in the perceptual-motor interaction required for language processing. Finally, using the significantly heritable AC endophenotypes, we performed a multivariate genome-wide association study on 31.775 participants from UK Biobank. 278 genomic loci were found significantly associated with language ACs including those previously associated with FC endophenotypes. As a conclusion, this work contributed to uncovering the neurobiology of human language using state-of-the-art genomic approaches and by developing novel language endophenotypes. The genes pinpointed in this work open new leads for investigations of neurobiology of language using animal and cellular models. Furthermore, as more and more cohorts are being built in the general population, this work strengthens our belief that multi-modal data collection in large-scale cohorts allow us to increase statistical power and to circumvent the small effect sizes in neuroimaging in order to investigate open neuroscience questions. Many other neuroscience fields could probably benefit from this type of methodology.