XTRACT - Standardised protocols for automated tractography and connectivity blueprints in the human and macaque brain

Warrington S[1], Bryant KL[2,3], Khrapitchev AA[4], Sallet J[5], Charquero-Ballester M[6], Douaud G[3], Jbabdi S[3*], Mars RB[2,3*], Sotiropoulos SN[1,3,7*]

[1]Sir Peter Mansfield Imaging Centre, School of Medicine, University of Nottingham, UK
[2]Donders Institute for Brain, Cognition, & Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands
[3]FMRIB, Wellcome Centre for Integrative Neuroimaging, Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, University of Oxford, Oxford, UK
[4]CRUK and MRC Oxford Institute for Radiation Oncology, Department of Oncology, University of Oxford, UK
[5]Wellcome Centre for Integrative Neuroimaging - Department of Experimental Psychology, University of Oxford, UK
[6]Department of Psychiatry, University of Oxford, UK
[7]National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre, Queens Medical Centre, Nottingham, UK

*Equal contribution

Corresponding author: Stamatios Sotiropoulos (stamatios.sotiropoulos@nottingham.ac.uk)

Abstract

We present a new toolbox and library of standardised tractography protocols devised for the robust automated extraction of white matter tracts both in the human and the macaque brain. Using in vivo data from the Human Connectome Project (HCP) and the UK Biobank and ex vivo data for the macaque brain datasets, we obtain white matter atlases, as well as atlases for tract endpoints on the white-grey matter boundary, for both species. We illustrate that our protocols are robust against data quality, generalisable across two species and reflect the known anatomy. We further demonstrate that they capture inter-subject variability by preserving tract lateralisation in humans and tract similarities stemming from twinship in the HCP cohort. Our results demonstrate that the presented toolbox will be useful for generating imaging-derived features in large cohorts, and in facilitating comparative neuroanatomy studies. The software, tractography protocols, and atlases are publicly released through FSL, allowing users to define their own tractography protocols in a standardised manner, further contributing to open science.

Introduction

Diffusion tractography is a unique tool for extracting white matter (WM) pathways non-invasively and in vivo. The virtual dissection of major WM tracts enables the study of brain organisation (Catani et al., 2013; Jbabdi et al., 2015) and offers a probe to brain development (Huppi and Dubois, 2006) and WM pathology (Ciccarelli et al., 2008; Griffa et al., 2013). It further allows explorations of individual variations (Assaf et al., 2017) and cross-species variations (Mars et al., 2018b) in anatomy and connectivity. This information has functional relevance, as the pattern of *extrinsic* WM connections of each functional brain subunit to the rest of the brain are unique (Mars et al., 2018a; Passingham et al., 2002).

To be able to reliably study individual variability in WM pathways, tractography approaches often utilise protocols to extract a pre-defined set of WM tracts. Such protocols must be robust and reproducible, allowing reconstruction of WM tracts in a consistent manner across subjects, while respecting the underlying anatomical variation and individual differences. Tractography protocols are typically comprised of masks and rules used to impose prior anatomical knowledge to guide and constrain curve propagation, reducing the chance of false positives (Catani et al., 2002; Wakana et al., 2004). One approach that may be used is to define subject-specific tractography protocols (Conturo et al., 1996), considering the specific variations in individual anatomy. However, defining masks on a subject-wise basis is time-consuming and subjective (Jones, 2008; Nucifora et al., 2012), while for large cohorts these limitations become prohibitive. The alternative to this manual approach is to define a set of standardised masks in template space, which are then registered to the individual geometry and used in a consistent and *automated* manner for each subject.

This approach has proven powerful in the extraction of a range of tracts (Catani and Thiebaut de Schotten, 2008; de Groot et al., 2013; Thiebaut de Schotten et al., 2011b; Wakana et al., 2007; Wassermann et al., 2016; Zhang et al., 2008) (see Supplementary Table 1 for a summary). (Wakana et al., 2007) developed a set of standard masks for the extraction of 20 tracts (9 left/right, 2 commissural). They reported high inter- and intra-rater reproducibility and suggested that some tracts may display left-right asymmetry. Similarly, (Catani and Thiebaut de Schotten, 2008) defined standard space masks for the reconstruction of 19 tracts (7 left/right, 5 commissural). They demonstrated how the use of standardised mask-based protocols may aid in improving the reproducibility of tractography results and produced tract atlases. This work was furthered by (Thiebaut de Schotten et al., 2011b) through the extension of the tractography protocols to 31 tracts (14 left/right, 3 commissural), where good correspondence between their automated tractography technique and histological atlases was reported. (de Groot et al., 2013) compiled a library of tractography protocols, adapted from the literature, and used standardised mask-based automated probabilistic tractography to reconstruct 27 tracts (12 left/right, 3 commissural) in two datasets with varying quality. (Wassermann et al., 2016) proposed a framework for describing WM anatomy and tracts which uses subject-wise anatomical segmentation, clustering and a query language to extract 57 (25 left/right, 7 commissural) tracts from whole-brain tractography. This approach reduces the definition of tracts to sets of logical rules with reference to whether a tract is present or terminates in given brain parcels.

More recently the power of automated tractography has been illustrated within a different context. (Mars et al., 2018b) defined a set of tracts, derived from the same tractography protocols as in (de Groot et al., 2013), in both humans and macaques, in order to perform comparative anatomy. They defined *connectivity blueprints* as (Cortex) x (Tracts)

maps, which can be used to identify functionally equivalent cortical regions between the two different species, in the absence of *any* geometrical similarity.

This paper extends these previous efforts in devising an extended set of ROI-based tractography protocols and an automated tractography toolbox[1]. The contribution of the presented work is as follows: 1) we design tractography protocols for 42 tracts and we illustrate their robustness against data quality, using high-resolution data from the Human Connectome Project (HCP) (Sotiropoulos et al., 2013) and more typical data from the UK Biobank (Miller et al., 2016), 2) we illustrate generalisability of the tractography protocols to the macaque brain, 3) we derive high-quality tract atlases using these protocols both for the human brain (1000 HCP subjects) and the macaque brain (high-resolution, ex vivo datasets from 6 animals), 4) we perform indirect validation by assessing lateralisation of the extracted tracts (in humans), 5) we illustrate that, despite being template-driven, reconstructed tracts preserve individual variability as assessed via twinship analysis, and 6) we offer an open-source flexible framework for publicly exchanging tractography protocols available within FSL. New standard space WM tractography protocols may be defined and "plugged into" the toolbox, allowing for further expansion and tract exchange, contributing to open science and reproducibility of results.

Tractography Protocol Definition

We devised tractography protocols for 42 WM tracts (19 bilateral and 4 commissural), in a generalisable manner that allows equivalent mask definitions to apply to *both the human and the macaque brain*. The full list of tracts that are currently supported is presented in Table 1. We further implemented a new cross-species tractography (XTRACT) toolbox, capable of reading the standard space tractography protocols and performing probabilistic tractography (Behrens et al., 2007), with the option of GPU acceleration (Hernandez-Fernandez et al., 2019).

Figure 1 illustrates the main stages for a single tractography protocol. Each tract is reconstructed using a unique combination of masks, defined in standard space (MNI152 for humans and F99 for macaques). The masks include seeds (starting points of the tractography streamlines), targets/waypoints (regions through which a streamline should pass in order to be valid), exclusions (regions that serve to reject any streamline running through them) and stop/termination masks (regions that serve to stop any streamline running through them). Seeding strategies support: a) a standard single-ROI seed and b) a "reverse-seeding" approach, where a pair of seed-target masks exchange roles and the final path distributions are added. The protocol masks are transformed from standard space to subject's native space using a non-linear registration warp field. Tractography is performed in native space and results are directly resampled to standard space, allowing between-subject geometrical correspondence, necessary in certain contexts (e.g. atlasing).

[1] XTRACT toolbox and human/macaque tractography protocols are available in FSL version 6.0.2 and later (www.fmrib.ox.ac.uk/fsl/XTRACT). Tract atlases will be made available in the next release of FSL and are currently available via GitHub (https://github.com/SPMIC-UoN/XTRACT_atlases).

The sections below describe in detail the protocol for each tract in consideration in the case of human tractography and any adjustments for the macaque brain. With the exception of the brainstem and commissural tracts all protocols include the midline sagittal plane as an exclusion mask to restrict fibres to the ipsilateral hemisphere.

Association fibres

Superior Longitudinal Fasciculus (SLF) 1/2/3: The three branches of the superior longitudinal fasciculus are reconstructed using an extension of the approach taken by (Thiebaut de Schotten et al., 2011a). In each case a coronal plane in the region of the central sulcus within the frontal/parietal cortex is used as a seed along with two target masks. Frontally, target masks for the first, second, and third branches of the SLF were coronal sections through the superior, middle, and inferior frontal gyri, respectively, placed at the level of the posterior end of the genu of the corpus callosum. Posteriorly, a large coronal target mask in the superior parietal lobule, immediately posterior to the margin of the cingulate gyrus is used for SLF1. For SLF2 and SLF3, the second target masks are placed in the angular gyrus and supramarginal gyrus respectively. An axial exclusion mask was placed underneath the parietal cortex and one blocking subcortical areas prevented leaking into ventrally oriented fibres. A final coronal exclusion mask through subcortical areas posterior to the caudal end of the genu of the corpus callosum prevented leaking into ventral longitudinal tracts.

Arcuate Fasciculus (AF): The arcuate fasciculus forms part of the system of dorsal longitudinal fibres, but in the human brain is distinguished by its posterior curve ventrally into the temporal cortex. The human AF was reconstructed with a seed in the supramarginal gyrus (SMG), a temporal target mask was in the WM encompassing the superior temporal gyrus (STG) and middle temporal gyrus (MTG), and an anterior target at the level of the ventral premotor cortex, posterior to the inferior frontal gyrus (IFG) and anterior to the precentral sulcus. Following the observation in the macaque that this tract runs along the fundus of the circular insular sulcus (Petrides et al., 2012) we placed a seed mask there, just posterior to the level of the central sulcus. An axial target mask was placed in the parietal-temporal WM posterior to the caudal end of the Sylvian fissure. An additional axial plane was placed in the IFG. This protocol was validated by (Eichert et al., 2019b).

Middle/Inferior Longitudinal Fasciculus (MdLF, ILF): Three tracts that course along the temporal lobe were reconstructed (MdLF, ILF, IFO). The middle and inferior longitudinal tracts stay within the lateral posterior cortex. MdLF was seeded in the anterior part of the SFG (Makris et al., 2009); ILF in the middle and inferior temporal gyri to account for the expansion of the temporal cortex in the human brain compared to the macaque (Latini et al., 2017; Roumazeilles et al., submitted). For the MdLF, large axial and coronal planes covering the WM in the temporo-parietal-occipital junction were used as targets, based on anatomical descriptions from (Makris et al., 2013). For ILF, an axial plane in middle and inferior temporal gyrus is used as a target. Exclusion masks where placed axially through the brainstem, coronally through the fornix, axially through the cingulum bundle posterior to the corpus callosum and through the entire frontal cortex. In addition, seed and target masks of MdLF served as exclusion masks for ILF and vice versa.

Inferior Fronto-Occipital Fasciculus (IFO): In contrast to MdLF and ILF, the inferior fronto-occipital fasciculus, also termed the extreme capsule fibre complex (Mars et al., 2016), runs more medially and courses into the frontal cortex through the extreme capsule. Extending the recipe of (Wakana et al., 2007), the seed was a coronal plane through the anterior part of the occipital cortex, the target a coronal plane through the frontal cortex anterior to the genu of the corpus callosum. An exclusion mask just behind the anterior commissure excluded all fibres except those running through the extreme capsule.

Uncinate Fasciculus (UF): The bottom part of the extreme capsule contains fibres belonging to the uncinate fasciculus, curving from the inferior frontal cortex to the anterior temporal cortex. The tract was reconstructed using a seed in the STG at the first location where temporal and frontal cortex are separated, a target through the ventral part of the extreme capsule, and exclusion masks through the rest of the extreme capsule and a layer between the seed and the target to force the curve. An additional coronal exclusion mask prevented accidental leaking into the fibres running longitudinally through the temporal lobe.

Frontal Aslant Tract (FA): The frontal aslant is a short tract running in the frontal lobe between the posterior part of the inferior and superior frontal gyri (Catani et al., 2012). The seed was placed sagittally in the WM of the IFG, the target axially in that of the SFG. A posterior coronal exclusion mask prevented leakage into longitudinal fibres.

Vertical Occipital Fasciculus (VOF): The vertical occipital fasciculus (VOF) runs in a predominantly dorsal-ventral orientation in the occipital lobe. We used an adapted version of the recipe described by (Takemura et al., 2017). An axial seed mask was placed in the lateral part of the ventral occipital WM posterior to the anterior occipital sulcus (Petrides et al., 2012). A larger axial target mask was placed dorsally at the level of the lateral occipital sulcus. A coronal plane just posterior to the corpus callosum served as an exclusion mask to prevent leakage into anterior-posterior tracts.

Commissural fibres

Middle Cerebellar Peduncle (MCP): The middle cerebellar peduncle (MCP) was seeded in the cerebellar WM with a target in the opposite hemisphere (and their inverses). Exclusion masks were placed sagitally along the cerebellar midline and axially through the thalamus.

Corpus Callosum Splenium (FMA) & Genu (FMI): We reconstructed callosal connections to the occipital lobe via the splenium of corpus callosum (forceps major, FMA) and to the frontal lobe via the genu of corpus callosum (forceps minor, FMI) using recipes based on those defined by (Wakana et al., 2007). Seed and target masks (and their inverse) for the FMA were defined as coronal sections through occipital lobe at the posterior end of the parietal-occipital sulcus. The sagittal exclusion mask was confined to the occipital cortex and the subcortex. Additional exclusion masks though the inferior fronto-occipital WM and a coronal plane through the pons prevented leakages to longitudinal fibres. Seed and target masks (and their inverse) for the FMI were defined as coronal sections through the frontal lobe at the anterior end of the pregenual cingulate sulcus. The sagittal exclusion mask was interrupted at the level of the anterior third of the corpus callosum.

Anterior Commissure (AC): The anterior commissure connects the temporal lobes of the two hemispheres across the midline. It was seeded in the left-right oriented fibres on the midline, with a target mask covering the WM lateral to the globus pallidae. Stop masks were placed directly underneath and lateral to the two amygdalae. A large axial exclusion mask was placed dorsal to the seed through the entire subcortex.

Limbic fibres

Cingulum subsections (CBT, CBP, CBD): Recently, (Heilbronner and Haber, 2014) proposed a segmentation of the cingulum bundle into distinct sections based on the presence of fibres connecting specific cingulate, non-cingulate frontal, and subcortical targets. We therefore created protocols for three distinct subsections of the cingulum bundle. The temporal part (CBT) was seeded in the posterior part of the temporal lobe at a section where the fibres of the cingulum are mostly oriented in the anterior-posterior direction. The target was placed posteriorly to the amygdala and stop masks were placed posteriorly and anteriorly to the seed and target masks, respectively. An exclusion mask prevented leaking into the fornix. The dorsal segment (CBD) was seeded just above the posterior part of the corpus callosum and had a target at the start of the genu of the corpus callosum. A sagittal exclusion mask in the anterior limb of the internal capsule prevented leakage into the temporal lobe. Finally, the peri-genual part of the cingulum bundle (CBP) was seeded anteriorly above the corpus callosum and a target placed below the sub-genual callosum with a stop mask placed inferior and anterior to the target. A callosal plane at the level of the rostral end of the Sylvian fissure prevented leakage into the CBD.

Fornix (FX): The fornix connects the hippocampus with the mammillary bodies, the anterior thalamic nuclei, and the hypothalamus (Catani et al., 2013). The tract was reconstructed using a seed in the body of the fornix at the level of the middle of the corpus callosum and a target in the hippocampus. A callosal plane at the anterior end of the occipital cortex prevented leakage into posterior tracts and bilateral sagittal planes around the midline, at the level of the anterior tip of the thalamus prevented lateral propagation to the anterior limb of the internal capsule. We should point out that due to the relatively small size of the stria terminalis and its close proximity to the fornix, the fornix tracking may leak into the stria terminalis. This is a common issue in diffusion tractography and is yet to be overcome using approaches in line with those used in the current study (Kamali et al., 2015; Mori et al., 2017; Mori and Aggarwal, 2014; Pascalau et al., 2018).

Projections fibres

Corticospinal Tract (CST): The corticospinal, or pyramidal, tract extends from the spinal cord through the midbrain and distributes to motor cortex, premotor cortex and somatosensory cortex. The tract is seeded from the pons with a large target covering the motor, premotor and somatosensory cortices. An axial exclusion mask is used to restrict tracking to the cerebral peduncle of the midbrain. In addition, the exclusion mask includes two coronal planes, anterior and posterior to the target, to exclude tracking to the prefrontal cortex and occipital cortex respectively.

Anterior and Superior Thalamic Radiations (ATR, STR): The anterior and superior thalamic radiations connect the thalamus to the frontal lobe and pre-/post-central gyrus respectively. The anterior thalamic radiation is seeded using a coronal mask through the anterior part of the thalamus (Wakana et al., 2007) with coronal target mask at the anterior thalamic peduncle. In addition, the exclusion mask contains an axial plane covering the base of the midbrain, a coronal plane preventing leakage via the posterior thalamic peduncle and a coronal plane preventing leakage via the cingulum. A coronal stop mask covers the posterior part of the thalamus, extending from the base of the midbrain to the callosal sulcus. The superior thalamic radiation is seeded using a mask covering the whole thalamus and a target axial plane covering the superior thalamic peduncle. An axial plane is used as a stop mask ventrally to the thalamus. The exclusion mask includes two coronal planes, anterior and posterior to the target, to exclude tracking to the prefrontal cortex and occipital cortex respectively.

Acoustic Radiation (AR): The acoustic radiation connects the medial geniculate nucleus (MGN) of the thalamus to the auditory cortex. It was seeded from the transverse temporal gyrus with a target covering the MGN of the thalamus. The exclusion mask consists of two coronal planes, anterior and posterior to the thalamus, and an axial plane superior to the thalamus. In addition, the exclusion mask contains the brainstem and a horizontal region covering the optic tract.

Optic Radiation (OR): The optic radiation consists of fibres from the lateral geniculate nucleus (LGN) of the thalamus to the primary visual cortex. It was seeded in the LGN and the target mask consisted of a coronal plane through the anterior part of the calcarine fissure. Exclusion masks consisted of an axial block of the brainstem, a coronal block of fibres directly posterior to the LGN to select fibres that curl around dorsally, and a coronal plane anterior to the seed to prevent leaking into longitudinal fibres.

Adjustments for the macaque brain

Although the protocols described above are such that they allow for equivalent definitions in the macaque brain, some adjustments were required to ensure anatomical accuracy. For all macaque protocols, the reverse-seeding method was used, as this was found to increase robustness in the resulting tracts. In addition, the AF and MdLF protocols were adjusted to reflect the macaque brain. In the case of the AF, a seed is placed in the caudal STG, a target directly above the principal sulcus extending posterior to 8Ad (based on the tract-tracing data of (Schmahmann and Pandya, 2009)). In addition, a target placed in the caudal STG, immediately inferior and posterior to the seed ensured tracking occurred via caudal end of the lateral fissure. For the MdLF, a single axial plane in the posterior part of the STG was used as a target.

Materials and Methods

Data and Preprocessing

To assess robustness across varying data quality, we utilised very high-quality diffusion MRI data from the Human Connectome Project (HCP) (Sotiropoulos et al., 2013; Van Essen et al., 2013) and data from the UK Biobank (Miller et al., 2016), which have overall quality closer to the one typically available through clinical scanners. To ensure generalisability of the protocols across species, we also utilised diffusion MRI data from the macaque brain. This data consisted of an extended set of animals used in (Eichert et al., 2019a; Mars et al., 2018b). In total, the datasets we considered consist of 1065 subjects from the HCP (all available HCP S1200 subjects that had diffusion MRI data), 1000 subjects from the UK Biobank and ex vivo high-resolution datasets from 6 macaques. For the HCP data, we removed 44 subjects with identified anatomical abnormalities from the statistical comparisons and group atlases (see the HCP quality control website for details[2]), leaving us with a total of 1021 subjects from the 1065 subjects with diffusion data.

For both the HCP and UK Biobank, we utilised the pre-processed dMRI data (see (Glasser et al., 2013; Sotiropoulos et al., 2013) and (Alfaro-Almagro et al., 2018; Miller et al., 2016) for full descriptions respectively). Briefly, the HCP data have been acquired in a bespoke 3T Connectom Skyra (Siemens, Erlangen) with a monopolar diffusion-weighted (Stejskal-Tanner) spin-echo EPI sequence, an isotropic spatial resolution of 1.25mm, three shells (b-values=1000, 2000 and 3000 s/mm^2) and 90 unique diffusion directions per shell, acquired twice. The UK Biobank data have been acquired in a clinical 3T Skyra (Siemens, Erlangen), consist of two shells (b-values=1000 and 2000 s/mm^2) and 50 diffusion directions per shell, with an isotropic spatial resolution of 2mm. In both cases, data were motion, susceptibility distortion and eddy current distortion corrected (Andersson et al., 2003; Andersson and Sotiropoulos, 2016). Nonlinear transformations to standard space (MNI152) were obtained using the respective T1-weighted images, to which the distortion-corrected diffusion MRI data were also linearly registered. Concatenation of the diffusion-to-T1 and T1-to-MNI transforms allowed diffusion-to-MNI warp fields to be obtained.

For the macaque data, we combined data previously used in (Mars et al., 2018b) with newly acquired data. Data were acquired locally on a 7T magnet with an Agilent DirectDrive console (Agilent Technologies, Santa Clara, CA, USA) using a 2D diffusion-weighted spin-echo protocol with single line readout (DW-SEMS, TE/TR: 25 ms/10 s; matrix size: 128 x 128; resolution: 0.6 x 0.6 mm; number of slices: 128; slice thickness: 0.6 mm; diffusion data were acquired over the course of 53 hours). 16 non-diffusion-weighted (b = 0 s/mm^2) and 128 diffusion-weighted (b = 4000 s/mm^2) volumes were acquired with diffusion directions distributed over the whole sphere. The brains were soaked in PBS before scanning and placed in fomblin or fluorinert during the scan. These data will be made available via PRIME-DE (Milham et al., 2018)[§].

Using FSL's FNIRT (Andersson et al., 2007; Jenkinson et al., 2012), estimations of the nonlinear transformations to standard space (F99) (Van Essen, 2002) were obtained based on the fractional anisotropy (FA) maps.

** https://wiki.humanconnectome.org/pages/viewpage.action?pageId=88901591
§ PRIMatE Data Exchange - http://fcon_1000.projects.nitrc.org/indi/indiPRIME.html

Fibre orientation estimation and tractography

The crossing fibre model described in (Jbabdi et al., 2012) was applied to the diffusion data and used to estimate orientations to inform tractography. This is a parametric spherical deconvolution model that accounts for the non-monoexponential decay of the dMRI signal with higher b-values. Up to three fibre orientations were estimated in each voxel along with their uncertainty. The "XTRACT" toolbox read the standard space tractography protocols and performed probabilistic tractography (Behrens et al., 2007). As discussed before, tractography protocols were defined for each bundle using a unique combination of seed, target, exclusion and stop masks, along with a seeding strategy (see Table 1). A number of default tractography termination criteria were also used in all protocols (curvature threshold: ±80 degrees, max streamline steps: 2000, subsidiary fibre volume threshold: 1%, loop-checking and termination). A step size of 0.5mm and 0.2mm were used for human and macaque tractography respectively. As shown in Figure 1, the masks were warped to the subject's native space and after tractography, the tractography results are directly resampled to standard space. The resultant distributions are normalised with respect to the total number of valid streamlines generated.

In order to obtain tract atlases, in the form of population percentage overlap, we binarised each normalised path distribution at a threshold value. The binary masks were then cohort-averaged to give the percentage of subjects for which a given tract is present at a given voxel.

Connectivity blueprints

The estimated bundles were further used to estimate maps of "cortical termination" for each tract in consideration, using connectivity blueprints (Mars et al., 2018b). Specifically, a white/grey matter boundary (WGB) x tracts matrix **CB** was reconstructed for each subject. This was achieved by seeding from every WGB location and counting the number of visitations to the whole WM, giving a WGB x WM connectivity $\mathbf{C_1}$ matrix. The tracts obtained using the tractography protocols were vectorised and concatenated into a single WM x tracts $\mathbf{C_2}$ matrix. Multiplying the two matrices provides a connectivity "blueprint", i.e. a $\mathbf{CB}=\mathbf{C_1}\mathrm{x}\mathbf{C_2}$ (WGB x tracts) matrix. Columns of this matrix represent the termination points of the corresponding tract on the WGB surface, while rows illustrate the connectivity pattern of each cortical location (i.e. how each tract contributes to the overall connectivity of each cortical location). This process was performed for the HCP subjects and the macaque datasets. The results were then cohort-averaged to produce connectivity blueprint atlases.

Assessing tract lateralisation

In order to demonstrate whether our protocols produce tracts representative of the anatomical expectations, we investigate tract lateralisation using a large number of subjects. Based on the literature, it is expected that AF is left-lateralised (Eichert et al., 2019b; Nowell et al., 2016; Panesar et al., 2018), IFO, MdLF and SLF3 are right-lateralised (Hau et al., 2016; Howells et al., 2018; Menjot de Champfleur et al., 2013; Thiebaut de Schotten et al., 2011a; Zhao et al., 2016), while SLF1 is expected to be non-lateralised (Thiebaut de Schotten et al.,

2011a). The literature suggests that the SLF2 is right-lateralised, however, findings are less conclusive as in some cases the reported lateralisation does not reach significance (Hecht et al., 2015; Thiebaut de Schotten et al., 2011a).

We assessed tract lateralisation using tract volume. Specifically, lateralisation (*L*) was calculated as the relative right-left volume (V_r and V_l) difference, after binarising the normalised tracts at 0.5% and taking the voxel count, i.e. $L=(V_r-V_l)/(V_r+V_l)$.

Furthermore, we explored inter-hemispheric differences in the cortical termination maps. A connectivity blueprint $\mathbf{CB_L}$, including only the left-hemisphere tracts/columns and a $\mathbf{CB_R}$**,** including only the right-hemisphere tracts/columns, were obtained. Both matrices were row-normalised, so that the sum of all elements in each row was equal to 1. Subsequently, we calculated the Kullback-Leibler (KL) divergence (a measure of dissimilarity) between every pair of ($\mathbf{CB_R}$, $\mathbf{CB_L}$) rows. We assessed the right-left similarity in connectivity patterns in every hemispheric location *i* using the minimum KL-divergence value obtained between all possible pairs, i.e. min($\mathbf{CB_{Ri}}$, $\mathbf{CB_{Lj}}$), with *j* spanning all WGB locations.

Respecting similarities stemming from twinship

Whilst we aimed for the automated tractography protocols to be robust against data quality, be reproducible and generalisable between species, we further tested whether they could respect features stemming from the inherent individual variability in WM anatomy across subjects. To demonstrate this, we explored the similarity of tract reconstructions within twin and non-twin sub-groups in the HCP cohort. We anticipate that monozygotic twin pairs will illustrate larger similarities than dizygotic twins and non-twin siblings, and subsequently than unrelated subject pairs, in line with the literature on the heritability of structural connections (Bohlken et al., 2014; Jansen et al., 2015; Shen et al., 2014) and as may be expected from the literature on sulcal similarities in twinship (Amiez et al., 2019). Using the 72 pairs of monozygotic twins available in the HCP cohort, 72 randomly chosen pairs of dizygotic twins, 72 randomly chosen pairs of non-twin siblings and 72 randomly chosen pairs of unrelated subjects, correlations between tract reconstructions were performed and used to assess whether our automated protocols respect the underlying tract variability across individuals.

Respecting individual differences due to atypical anatomy

We excluded a small number of subjects from the HCP cohort-derived atlases, due to identified anatomical abnormalities, following the HCP quality control recommendations. However, we explored the performance of our tractography protocols in a number of these cases, and particularly the ability to handle atypical anatomical features.

Results

WM Tract and WGB termination atlases of the human brain

We applied the prescribed tractography protocols to ~1000 HCP and 1000 UK Biobank subjects. Figure 2 presents the tract atlases, obtained from the HCP datasets. To obtain these atlases, the subject-specific MNI-transformed tracts were binarised and subsequently averaged to produce the population percent coverage for every tract, shown in the Figure. The atlases obtained using the UK Biobank data are shown in Supplementary Figure 1.

Connectivity blueprints were also derived for each HCP subject and averaged to obtain an atlas. Examples of columns of this average connectivity blueprint across all HCP subjects are shown in Figure 3, representing atlases of termination points of each tract on the white-grey matter boundary (WGB) surface.

We also investigated the effect of sample size on tract atlas creation. For each of the HCP and UK Biobank cohorts, we produced tract atlases with increasing numbers of subjects and correlated, tract-wise, each set of atlases to a 1000-subject atlas set. Supplementary Figure 2 shows the distributions of the tract-wise correlations for each of the sample size atlases. The top plot includes an atlas set with a sample size of 10 subjects which, whilst already showing high correlations, performs relatively poorly compared to using a sample size of 100 or greater.

Robustness against datasets

To explore robustness against varying data quality, we compared tract atlases and inter-subject variability of the tract reconstructions within and across cohorts. To compare atlases, each tract from the HCP atlas set was correlated with its corresponding UK Biobank tract atlas (population threshold of 30% applied to each tract atlas). The average correlation across tracts was 0.80 (standard deviation = 0.07).

Inter-subject correlations were obtained by correlating random subject pairs within and across cohorts. To avoid possible family structure-induced bias in the HCP, we restricted our subjects to the 339 unrelated subjects. We matched the number of subjects in the UK Biobank data by randomly selecting 339 (gender matched) subjects. The across-cohort comparison was made by correlating a random subject in the unrelated HCP subject pool with a random UK Biobank subject, giving a distribution of 339 correlations per tract. Within-cohort comparisons gave average correlation values of 0.52 (standard deviation = 0.09) and 0.54 (standard deviation = 0.09) for the HCP and UK Biobank respectively, with no significant difference in the within-cohort correlations for the two cohorts (p = 0.06 - Mann-Whitney U test) (Figure 4). Across-cohort comparison gave an average correlation of 0.41 (standard deviation = 0.10) which is lower than within-cohort comparisons (p = 2x10^-5 and 3x10^-7), yet it is comparable enough, particularly given the age difference of subjects in the two cohorts (HCP:22-35 years old, UK Biobank: 40-69 years old, mean age for our chosen subjects in HCP = 28.6 (standard deviation = 3.7) and UK Biobank = 62.6 (standard deviation = 7.5)).

Generalisation across species - WM Tract and WGB termination atlases of the macaque brain

We repeated tractography as described above and obtained atlases from the macaque cohort. Our tractography protocol definitions are such that they allow the extraction of homologous tracts in both the human brain and macaque brain.

Tract atlases in the form of population percent were generated and are shown in Figure 5, while Figure 6 shows the averages of the connectivity blueprints derived using the macaque data.

Tract Lateralisation

Figure 7 illustrates tract-lateralisation estimates for a subset of tracts in the human brain, for which lateralisation has been previously reported in the literature. As shown in Figure 7, the AF is left-lateralised and IFO, MdLF and SLF3 are right-lateralised in both the HCP and the UK Biobank cohorts. SLF1 is symmetric in the HCP cohort but reaches rightward significance in the UK Biobank cohort. SLF2 is also variable across cohorts with left-lateralisation in the HCP and right-lateralisation in the UK Biobank.

In addition to volume-based measures of lateralisation, inter-hemispheric differences on connectivity patterns were also assessed on the WGB surface using the tracts-derived connectivity blueprints. In an approach similar to (Mars et al., 2018b), Kullback-Leibler (KL) divergence was calculated to explore connectivity similarity between the two hemispheres of the human brain. For every location on the right hemisphere surface, the minimum KL divergence value assesses the most similar connectivity pattern on the left hemisphere. In doing so, we can probe cortical locations that demonstrate dissimilar connection patterns between left and right hemispheres and assess which tracts are contributing to these dissimilarities. In areas of high minimum KL divergence, we would expect to observe differences in the tract contribution profiles between the corresponding vertices. Figure 8a shows the minimum KL divergence values obtained for all WGB surface locations, overlaid with a subset of the Glasser parcellation (Glasser et al., 2016). Regions of high left-right dissimilarity in connectivity patterns were generally confined to frontal and temporo-parietal junction (TPJ) regions. Figure 8b shows examples of the tract contribution to the connection pattern of specific WGB locations on the right hemisphere and how these compare with the connection patterns of the best matching location on the left hemisphere. Three examples are shown corresponding to varying degrees of inter-hemispheric dissimilarity. It can be seen that high dissimilarity was mediated by tracts that were found to be lateralised. For example, inter-hemispheric differences for a selected voxel in IFSa were primarily driven by differences in how the lateralised SLF3 contributes to its connectivity pattern. For a mid-range dissimilarity vertex, selected in the temporo-parietal-occipital junction (TPOJ1), we can see that small inter-hemispheric divergence was driven primarily by the AF and SLF3. Conversely, regions of low dissimilarity, such as the fourth visual area (V4), show little inter-hemispheric difference in the tracts contributing to the connectivity profile.

Preserving individual variability and twinship-induced similarities

To explore whether the tractography protocols preserved individual variability in WM anatomy, we compared tract reconstructions for 72 pairs of monozygotic twins, dizygotic twins, non-twin siblings and unrelated subjects from the HCP. Figure 9 shows the distributions of the average tract-wise correlations (i.e. average across subjects for each tract) for each group. As shown in Figure 9, monozygotic twin pairs, on average (median), have a higher correlation (0.598, standard deviation = 0.108) with their corresponding twin compared to dizygotic twin pairs (0.550, standard deviation = 0.099), non-twin sibling pairs (0.544, standard deviation = 0.098), and unrelated subject pairs (0.512, standard deviation = 0.093). A Kruskal-Wallis test demonstrates statistically significant differences between subgroup medians: $\chi^2 = 13.2$, $p = 0.0043$, mean ranks = 104.1 (monozygotic), 85.3 (dizygotic), 82.9 (non-twin siblings) and 65.7 (unrelated).

In addition to exploring the similarity of tract reconstructions in twins, we further investigated whether the automated tractography respected individual variability in the case of anatomical abnormalities. A subset of subjects with gross anatomical abnormalities were identified using the HCP quality control. Figure 10 gives examples of these subjects and highlights the difference between the average tracts (as provided by the atlas) and the individual subject tractography results, which reflect the presence of cavernomas and cysts in WM.

Discussion

We have presented a new toolbox (XTRACT) for automated probabilistic tractography along with standardised protocols for extracting white matter bundles in the human and the macaque brain. We have demonstrated that the protocols are robust when applied to data of varying image quality and to data from a non-human primate species. We have generated human WM tract atlases using an order of magnitude more data than previous efforts, as well as macaque atlases using a small number of, however high-quality ex vivo, datasets. We have performed indirect validation illustrating that reconstructed tracts are left/right asymmetric, when they are expected to be based on prior literature. We have also shown that despite automatically generating tracts using standard-space protocols, the protocols respect the underlying individual variability, as reflected in twinship-induced similarities and in respecting anatomical abnormalities. The toolbox, tractography protocols and atlases are freely and openly available as a part of FMRIB's software library (FSL) (version 6.0.2 and later).

A current issue in the field of tractography is that protocol definitions in journal publications often lack detail or are designed without data-sharing in mind. Solutions have been proposed to resolve this issue, for example, the white matter query language (Wassermann et al., 2016). Here, we offer a platform for direct sharing of standardised protocol masks and tract atlases. Moreover, we made the protocol definitions generalisable across species to directly facilitate comparative anatomy studies.

We reconstructed tracts using imaging datasets of different quality and we generated atlases for both the Human Connectome Project (HCP) and the UK Biobank cohorts. Comparisons of tract reconstructions within and across the human cohorts demonstrate that

the method and protocols are robust across subjects and against data quality. The HCP and UK Biobank cohorts provide examples of high-quality data and more typical quality data respectively. Within cohort comparisons reveal similar inter-subject tract correlations across the varying quality data, with greater inter-subject tract correlations observed in the UK Biobank. This may reflect a reduced level of detail in the lower resolution UK Biobank data compared to the HCP data, but also differences in the mean age of subjects in the two cohorts. In addition, we have generated atlases using a smaller cohort of macaques. To compensate for the small number of subjects, we used high-quality and high-resolution ex vivo data. The respective results demonstrate the generalisability of our method to the macaque brain. Recent efforts to obtain macaque data from larger cohorts (HCP-style protocols) are ongoing (Autio et al., 2019; Milham et al., 2018) and our tools will be a useful resource for these new initiatives for the non-human primate brain (Thiebaut de Schotten et al., 2019).

As a means to indirectly validate our results, we investigated left-right tract lateralisation. We compared our lateralisation results to *a priori* knowledge from the literature. For both human cohorts (HCP and UK Biobank), we found that reconstructed AF is strongly left-lateralised, while SLF3, IFO and MDLF were right-lateralised, as expected from the literature (Eichert et al., 2019b; Hau et al., 2016; Hecht et al., 2015; Nowell et al., 2016; Panesar et al., 2018; Thiebaut de Schotten et al., 2011a; Zhao et al., 2016). Results were less clear-cut for SLF1 and SLF2, where prior studies (with much fewer numbers of subjects) are inconclusive (Hecht et al., 2015; Howells et al., 2018; Thiebaut de Schotten et al., 2011a). This may be due to the large variance observed (in the case of SLF2), perhaps reflecting some underlying interaction, such as handedness (Howells et al., 2018).

We performed further sanity checks by investigating lateralisation using the connectivity blueprint that we obtained from the reconstructed tracts. By using the KL divergence between connectivity patterns to assess inter-hemispheric dissimilarity, we identified that regions associated with language, known to be lateralised (Hiscock and Kinsbourne, 2008), have dissimilar connectivity patterns across the two hemispheres (Figure 8a).

Whilst being a robust automated method for the consistent reconstruction of tracts, our method also respected the underlying anatomical variation. We demonstrated this by assessing inter-subject tract similarity in monozygotic twins, dizygotic twins, non-twin siblings and unrelated subject pairs. Our results show greater similarity in twin pairs compared to unrelated pairs, as would be expected from the heritability literature (Bohlken et al., 2014; Shen et al., 2014). We further demonstrated that the automated method respects underlying anatomical variation by exploring how tractography results differ from the cohort-averaged results in the case of subjects with anatomical abnormalities.

Conclusions

In conclusion, we have developed and demonstrated a set of robust and standardised tractography protocols for cross-species automated delineation of white matter bundles, along with a platform to use them. The demonstrated toolbox (XTRACT) is freely available along with the tractography protocols and human/macaque tract atlases as a part of FMRIB's software library (FSL version 6.0.2 and later). Given the benefits with regards to data and protocol sharing, we expect that this toolbox will aid reproducibility in the field of tractography and facilitate comparative neuroanatomy studies.

Acknowledgements

S.W. was supported by a Medical Research Council PhD Studentship (MR/N013913/1). K.L.B. was supported by a Marie Skłodowska-Curie Individual Fellowship Grant MSCA-IF [750026]. R.B.M. is supported by the Biotechnology and Biological Sciences Research Council (BBSRC) UK [BB/N019814/1] and the Netherlands Organization for Scientific Research NWO [452-13-015]. J.S. was supported by a Sir Henry Dale Wellcome Trust Fellowship (105651/Z/14/Z). G.D. is supported by an MRC Career Development Fellowship (MR/K006673/1). The work was also supported by grant EP/L023067/1 from the UK Engineering and Physical Sciences Research Council (EPSRC). Human datasets were provided in part by a) The Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University and b) The UK Biobank Resource under Application Number 43822. The computations described in this paper were performed using the University of Nottingham's Augusta HPC service and the Precision Imaging Beacon Cluster, which provide High Performance Computing service to the University's research community. The Wellcome Centre for Integrative Neuroimaging is supported by core funding from the Wellcome Trust [203139/Z/16/Z].

References

Alfaro-Almagro, F., Jenkinson, M., Bangerter, N.K., Andersson, J.L.R., Griffanti, L., Douaud, G., Sotiropoulos, S.N., Jbabdi, S., Hernandez-Fernandez, M., Vallee, E., Vidaurre, D., Webster, M., McCarthy, P., Rorden, C., Daducci, A., Alexander, D.C., Zhang, H., Dragonu, I., Matthews, P.M., Miller, K.L., Smith, S.M., 2018. Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. Neuroimage 166, 400–424. https://doi.org/10.1016/j.neuroimage.2017.10.034

Amiez, C., Wilson, C.R.E., Procyk, E., 2019. Variations of cingulate sulcal organization and link with cognitive performance. Sci. Rep. 32. https://doi.org/10.1038/s41598-018-32088-9

Andersson, J.L., Skare, S., Ashburner, J., 2003. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. Neuroimage 20, 870–888. https://doi.org/10.1016/S1053-8119(03)00336-7

Andersson, J.L.R., Jenkinson, M., Smith, S.M., 2007. Non-linear optimisation. FMRIB technical report TR07JA1. Pr. https://doi.org/10.1109/EMOBILITY.2010.5668100

Andersson, J.L.R., Sotiropoulos, S.N., 2016. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. Neuroimage 125, 1063–1078. https://doi.org/10.1016/j.neuroimage.2015.10.019

Assaf, Y., Johansen-Berg, H., Thiebaut de Schotten, M., 2017. The role of diffusion MRI in neuroscience. NMR Biomed. https://doi.org/10.1002/nbm.3762

Autio, J.A., Glasser, M.F., Ose, T., Donahue, C.J., Bastiani, M., Ohno, M., Kawabata, Y., Urushibata, Y., Murata, K., Nishigori, K., Yamaguchi, M., Hori, Y., Yoshida, A., Go, Y., Coalson, T.S., Jbabdi, S., Sotiropoulos, S.N., Smith, S., Van Essen, D.C., Hayashi, T., 2019. Towards HCP-Style Macaque Connectomes: 24-Channel 3T Multi-Array Coil, MRI Sequences and Preprocessing. bioRxiv. https://doi.org/https://doi.org/10.1101/602979

Behrens, T.E., Johansen-Berg, H., Jbabdi, S., Rushworth, M.F., Woolrich, M.W., 2007. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? Neuroimage 34, 144–155. https://doi.org/10.1016/j.neuroimage.2006.09.018

Bohlken, M.M., Mandl, R.C., Brouwer, R.M., van den Heuvel, M.P., Hedman, A.M., Kahn, R.S., Hulshoff Pol, H.E., 2014. Heritability of structural brain network topology: A DTI study of 156 twins. Hum. Brain Mapp. 35, 5295–5305. https://doi.org/10.1002/hbm.22550

Catani, M., Dell'Acqua, F., Thiebaut de Schotten, M., 2013. A revised limbic system model for memory, emotion and behaviour. Neurosci. Biobehav. Rev. 37, 1724–1737. https://doi.org/https://doi.org/10.1101/645234

Catani, M., Dell'Acqua, F., Vergani, F., Malik, F., Hodge, H., Roy, P., Valabregue, R., Thiebaut de Schotten, M., 2012. Short frontal lobe connections of the human brain. Cortex 48, 273–291. https://doi.org/10.1016/j.cortex.2011.12.001

Catani, M., Howard, R.J., Pajevic, S., Jones, D.K., 2002. Virtual in Vivo interactive dissection of white matter fasciculi in the human brain. Neuroimage 17, 77–94. https://doi.org/10.1006/nimg.2002.1136

Catani, M., Thiebaut de Schotten, M., 2008. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. Cortex 44, 1105–1132.

https://doi.org/10.1016/j.cortex.2008.05.004

Ciccarelli, O., Catani, M., Johansen-Berg, H., Clark, C., Thompson, A., 2008. Diffusion-based tractography in neurological disorders: concepts, applications, and future developments. Lancet Neurol. 7, 715–727. https://doi.org/10.1016/S1474-4422(08)70163-7

Conturo, T.E., Lori, N.F., Cull, T.S., Akbudak, E., Snyder, A.Z., Shimony, J.S., McKinstry, R.C., Burton, H., Raichle, M.E., 1996. Tracking neuronal fiber pathways in the living human brain. Proc. Natl. Acad. Sci. U. S. A. 96, 10422–10427. https://doi.org/10.1073/pnas.96.18.10422

de Groot, M., Vernooij, M.W., Klein, S., Ikram, M.A., Vos, F.M., Smith, S.M., Niessen, W.J., Andersson, J.L., 2013. Improving alignment in Tract-based spatial statistics: Evaluation and optimization of image registration. Neuroimage 76, 400–411. https://doi.org/10.1016/j.neuroimage.2013.03.015

Eichert, N., Robinson, E.C., Bryant, K.L., Jbabdi, S., Li, L., Krug, K., Watkins, K.E., Mars, R.B., 2019a. Cross-species cortical alignment identifies different types of neuroanatomical reorganization in higher primates. Biorxiv.

Eichert, N., Verhagen, L., Folloni, D., Jbabdi, S., Khrapitchev, A.A., Sibson, N.R., Mantini, D., Sallet, J., Mars, R.B., 2019b. What is special about the human arcuate fasciculus? Lateralization, projections, and expansion. Cortex 118, 107–115. https://doi.org/https://doi.org/10.1016/j.cortex.2018.05.005

Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., Smith, S.M., Van Essen, D.C., 2016. A multi-modal parcellation of human cerebral cortex. Nature 536, 171–178. https://doi.org/10.1038/nature18933

Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., Van Essen, D.C., Jenkinson, M., WU-Minn HCP Consortium, 2013. The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage 80, 105–124. https://doi.org/10.1016/j.neuroimage.2013.04.127

Griffa, A., Baumann, P.S., Thiran, J.P., Hagmann, P., 2013. Structural connectomics in brain diseases. Neuroimage 80, 515–526. https://doi.org/10.1016/j.neuroimage.2013.04.056

Hau, J., Sarubbo, S., Perchey, G., Crivello, F., Zago, L., Mellet, E., Jobard, G., Joliot, M., Mazoyer, B.M., Tzourio-Mazoyer, N., Petit, L., 2016. Cortical terminations of the inferior fronto-occipital and uncinate fasciculi: Anatomical stem-based virtual dissection. Front. Neuroanat. https://doi.org/10.3389/fnana.2016.00058

Hecht, E.E., Gutman, D.A., Bradley, B.A., Preuss, T.M., Stout, D., 2015. Virtual dissection and comparative connectivity of the superior longitudinal fasciculus in chimpanzees and humans. Neuroimage 108, 124–137. https://doi.org/10.1016/j.neuroimage.2014.12.039

Heilbronner, S.R., Haber, S.N., 2014. Frontal Cortical and Subcortical Projections Provide a Basis for Segmenting the Cingulum Bundle: Implications for Neuroimaging and Psychiatric Disorders. J. Neurosci. 34, 10041–10054. https://doi.org/10.1523/jneurosci.5459-13.2014

Hernandez-Fernandez, M., Reguly, I., Jbabdi, S., Giles, M., Smith, S., Sotiropoulos, S.N., 2019.

Using GPUs to accelerate computational diffusion MRI: From microstructure estimation to tractography and connectomes. Neuroimage 188, 598–615. https://doi.org/10.1016/j.neuroimage.2018.12.015

Hiscock, M., Kinsbourne, M., 2008. Lateralization of language across the life span, in: Handbook of the Neuroscience of Language. Elsevier, pp. 247–255.

Howells, H., Thiebaut de Schotten, M., Dell'Acqua, F., Beyh, A., Zappalà, G., Leslie, A., Simmons, A., Murphy, D.G., Catani, M., 2018. Frontoparietal tracts linked to lateralized hand preference and manual specialization. Cereb. Cortex 28, 2482–2494. https://doi.org/10.1093/cercor/bhy040

Huppi, P.S., Dubois, J., 2006. Diffusion tensor imaging of brain development. Semin Fetal Neonatal Med 11, 489–497.

Jansen, A.G., Mous, S.E., White, T., Posthuma, D., Polderman, T.J., 2015. What Twin Studies Tell Us About the Heritability of Brain Development, Morphology, and Function: A Review. Neuropsychol. Rev. 25, 27–46. https://doi.org/10.1007/s11065-015-9278-9

Jbabdi, S., Sotiropoulos, S.N., Haber, S.N., Van Essen, D.C., Behrens, T.E., 2015. Measuring macroscopic brain connections in vivo. Nat. Neurosci. 18, 1546–1555. https://doi.org/10.1038/nn.4134

Jbabdi, S., Sotiropoulos, S.N., Savio, A.M., Graña, M., Behrens, T.E., 2012. Model-based analysis of multishell diffusion MR data for tractography: How to get over fitting problems. Magn. Reson. Med. 68, 1846–1855. https://doi.org/10.1002/mrm.24204

Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL - Review. Neuroimage 62, 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Jones, D.K., 2008. Studying connections in the living human brain with diffusion MRI. Cortex 44, 936–952. https://doi.org/10.1016/j.cortex.2008.05.002

Kamali, A., Yousem, D.M., Lin, D.D., Sair, H.I., Jasti, S.P., Keser, Z., Riascos, R.F., Hasan, K.M., 2015. Mapping the trajectory of the stria terminalis of the human limbic system using high spatial resolution diffusion tensor tractography. Neurosci. Lett. 608, 45–50. https://doi.org/10.1016/j.neulet.2015.09.035

Latini, F., Mårtensson, J., Larsson, E.M., Fredrikson, M., Åhs, F., Hjortberg, M., Aldskogius, H., Ryttlefors, M., 2017. Segmentation of the inferior longitudinal fasciculus in the human brain: A white matter dissection and diffusion tensor tractography study. Brain Res. 1675, 102–115. https://doi.org/10.1016/j.brainres.2017.09.005

Makris, N., Papadimitriou, G.M., Kaiser, J.R., Sorg, S., Kennedy, D.N., Pandya, D.N., 2009. Delineation of the middle longitudinal fascicle in humans: A quantitative, in vivo, DT-MRI study. Cereb. Cortex 19, 777–785. https://doi.org/10.1093/cercor/bhn124

Makris, N., Preti, M.G., Wassermann, D., Rathi, Y., Papadimitriou, G.M., Yergatian, C., Dickerson, B.C., Shenton, M.E., Kubicki, M., 2013. Human middle longitudinal fascicle: Segregation and behavioral-clinical implications of two distinct fiber connections linking temporal pole and superior temporal gyrus with the angular gyrus or superior parietal lobule using multi-tensor tractography. Brain Imaging Behav. 7, 335–352. https://doi.org/10.1007/s11682-013-9235-2

Mars, R.B., Foxley, S., Verhagen, L., Jbabdi, S., Sallet, J., Noonan, M.P., Neubert, F.X., Andersson, J.L., Croxson, P.L., Dunbar, R.I., Khrapitchev, A.A., Sibson, N.R., Miller, K.L.,

Rushworth, M.F., 2016. The extreme capsule fiber complex in humans and macaque monkeys: a comparative diffusion MRI tractography study. Brain Struct. Funct. 221, 4059–4071. https://doi.org/10.1007/s00429-015-1146-0

Mars, R.B., Passingham, R.E., Jbabdi, S., 2018a. Connectivity Fingerprints: From Areal Descriptions to Abstract Spaces. Trends Cogn. Sci. 22, 1026–1037. https://doi.org/https://doi.org/10.1016/j.tics.2018.08.009

Mars, R.B., Sotiropoulos, S.N., Passingham, R.E., Sallet, J., Verhagen, L., Krapitchev, A.A., Sibson, N., Jbabdi, S., 2018b. Whole brain comparative anatomy using connectivity blueprints. Elife 7. https://doi.org/10.7554/eLife.35237

Menjot de Champfleur, N., Maldonado, I.L., Moritz-Gasser, S., Machi, P., Le Bars, E., Bonafé, A., Duffau, H., 2013. Middle longitudinal fasciculus delineation within language pathways: a diffusion tensor imaging study in human. Eur. J. Radiol. 82, 151–157.

Milham, M.P., Ai, L., Koo, B., Xu, T., Amiez, C., Balezeau, F., Baxter, M.G., Blezer, E.L.A., Brochier, T., Chen, A., Croxson, P.L., Damatac, C.G., Dehaene, S., Everling, S., Fair, D.A., Fleysher, L., Freiwald, W., Froudist-Walsh, S., Griffiths, T.D., Guedj, C., Hadj-Bouziane, F., Ben Hamed, S., Harel, N., Hiba, B., Jarraya, B., Jung, B., Kastner, S., Klink, P.C., Kwok, S.C., Laland, K.N., Leopold, D.A., Lindenfors, P., Mars, R.B., Menon, R.S., Messinger, A., Meunier, M., Mok, K., Morrison, J.H., Nacef, J., Nagy, J., Rios, M.O., Petkov, C.I., Pinsk, M., Poirier, C., Procyk, E., Rajimehr, R., Reader, S.M., Roelfsema, P.R., Rudko, D.A., Rushworth, M.F.S., Russ, B.E., Sallet, J., Schmid, M.C., Schwiedrzik, C.M., Seidlitz, J., Sein, J., Shmuel, A., Sullivan, E.L., Ungerleider, L., Thiele, A., Todorov, O.S., Tsao, D., Wang, Z., Wilson, C.R.E., Yacoub, E., Ye, F.Q., Zarco, W., Zhou, Y. di, Margulies, D.S., Schroeder, C.E., 2018. An Open Resource for Non-human Primate Imaging. Neuron 100, 61–74. https://doi.org/10.1016/j.neuron.2018.08.039

Miller, K.L., Alfaro-Almagro, F., Bangerter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L., Griffanti, L., Douaud, G., Okell, T.W., Weale, P., Dragonu, I., Garratt, S., Hudson, S., Collins, R., Jenkinson, M., Matthews, P.M., Smith, S.M., 2016. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat. Neurosci. 19, 1523–1536. https://doi.org/10.1038/nn.4393

Mori, S., Aggarwal, M., 2014. In vivo magnetic resonance imaging of the human limbic white matter. Front. Aging Neurosci. https://doi.org/10.3389/fnagi.2014.00321

Mori, S., Kageyama, Y., Hou, Z., Aggarwal, M., Patel, J., Brown, T., Miller, M.I., Wu, D., Troncoso, J.C., 2017. Elucidation of white matter tracts of the human amygdala by detailed comparison between high-resolution postmortem magnetic resonance imaging and histology. Front. Neuroanat. https://doi.org/10.3389/fnana.2017.00016

Nowell, M., Vos, S.B., Sidhu, M., Wilcoxen, K., Sargsyan, N., Ourselin, S., Duncan, J.S., 2016. Meyer's loop asymmetry and language lateralisation in epilepsy. J. Neurol. Neurosurg. Psychiatry 87, 836–842.

Nucifora, P.G.P., Wu, X., Melhem, E.R., Gur, R.E., Gur, R.C., Verma, R., 2012. Automated Diffusion Tensor Tractography. Implementation and Comparison to User-driven Tractography. Acad. Radiol. 19, 622–629. https://doi.org/10.1016/j.acra.2012.01.002

Panesar, S.S., Yeh, F.C., Jacquesson, T., Hula, W., Fernandez-Miranda, J.C., 2018. A Quantitative Tractography Study into the Connectivity, Segmentation and Laterality of

the Human Inferior Longitudinal Fasciculus. Front. Neuroanat. 12.

Pascalau, R., Popa Stănilă, R., Sfrângeu, S., Szabo, B., 2018. Anatomy of the Limbic White Matter Tracts as Revealed by Fiber Dissection and Tractography. World Neurosurg. 113, e672–e689. https://doi.org/10.1016/j.wneu.2018.02.121

Passingham, R.E., Stephan, K.E., Kötter, R., 2002. The anatomical basis of functional localization in the cortex. Nat. Rev. Neurosci. 3, 606–616. https://doi.org/10.1038/nrn893

Petrides, M., Tomaiuolo, F., Yeterian, E.H., Pandya, D.N., 2012. The prefrontal cortex: Comparative architectonic organization in the human and the macaque monkey brains. Cortex 48, 46–57. https://doi.org/10.1016/j.cortex.2011.07.002

Roumazeilles, L., Eichert, N., Bryant, K.L., Folloni, D., Sallet, J., Vijayakumar, S., Foxley, S., Tendler, B.C., Jbabdi, S., Reveley, C., Verhagen, L., Dershowitz, L.B., Guthrie, M., Flach, E., Miller, K.L., Mars, R.B., n.d. Longitudinal connections and the organization of the temporal cortex in macaques, great apes, and humans. Submitted.

Schmahmann, J.D., Pandya, D.N., 2009. Fiber Pathways of the Brain, Fiber Pathways of the Brain. https://doi.org/10.1093/acprof:oso/9780195104233.001.0001

Shen, K.-K., Rose, S., Fripp, J., McMahon, K.L., de Zubicaray, G.I., Martin, N.G., Thompson, P.M., Wright, M.J., Salvado, O., 2014. Investigating brain connectivity heritability in a twin study using diffusion imaging data. Neuroimage 100, 628–641. https://doi.org/10.1016/j.neuroimage.2014.06.041

Sotiropoulos, S.N., Jbabdi, S., Xu, J., Andersson, J.L., Moeller, S., Auerbach, E.J., Glasser, M.F., Hernandez, M., Sapiro, G., Jenkinson, M., Feinberg, D.A., Yacoub, E., Lenglet, C., Ven Essen, D.C., Ugurbil, K., Behrens, T.E., WU-Minn HCP Consortium, 2013. Advances in diffusion MRI acquisition and processing in the Human Connectome Project. Neuroimage 80, 125–143. https://doi.org/10.1016/j.neuroimage.2013.05.057

Takemura, H., Pestilli, F., Weiner, K.S., Keliris, G.A., Landi, S.M., Sliwa, J., Ye, F.Q., Barnett, M.A., Leopold, D.A., Freiwald, W.A., Logothetis, N.K., Wandell, B.A., 2017. Occipital White Matter Tracts in Human and Macaque. Cereb. Cortex 27, 3346–3359. https://doi.org/10.1093/cercor/bhx070

Thiebaut de Schotten, M., Croxson, P.L., Mars, R.B., 2019. Large-scale comparative neuroimaging: Where are we and what do we need? Cortex 118, 188–202. https://doi.org/10.1016/j.cortex.2018.11.028

Thiebaut de Schotten, M., Dell'Acqua, F., Forkel, S.J., Simmons, A., Vergani, F., Murphy, D.G.M., Catani, M., 2011a. A lateralized brain network for visuospatial attention. Nat. Neurosci. 14, 1245–1246. https://doi.org/10.1038/nn.2905

Thiebaut de Schotten, M., Ffytche, D.H., Bizzi, A., Dell'Acqua, F., Allin, M., Walshe, M., Murray, R., Williams, S.C., Murphy, D.G., Catani, M., 2011b. Atlasing location, asymmetry and inter-subject variability of white matter tracts in the human brain with MR diffusion tractography. Neuroimage 54, 49–59. https://doi.org/10.1016/j.neuroimage.2010.07.055

Van Essen, D.C., 2002. Windows on the brain: The emerging role of atlases and databases in neuroscience. Curr. Opin. Neurobiol. 12, 574–579. https://doi.org/10.1016/S0959-4388(02)00361-6

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., WU-Minn HCP Consortium, 2013. The WU-Minn Human Connectome Project: an overview. Neuroimage 80, 62–79. https://doi.org/10.1016/j.neuroimage.2013.05.041

Wakana, S., Caprihan, A., Panzenboeck, M.M., Fallon, J.H., Perry, M., Gollub, R.L., Hua, K., Zhang, J., Jiang, H., Dubey, P., Blitz, A., van Zijl, P., Mori, S., 2007. Reproducibility of quantitative tractography methods applied to cerebral white matter. Neuroimage 36, 630–644. https://doi.org/10.1016/j.neuroimage.2007.02.049

Wakana, S., Jiang, H., Nagae-Poetscher, L.M., van Zijl, P.C., Mori, S., 2004. Fiber Tract–based Atlas of Human White Matter Anatomy. Radiology 230. https://doi.org/10.1148/radiol.2301021640

Wassermann, D., Makris, N., Rathi, Y., Shenton, M., Kikinis, R., Kubicki, M., Westin, C.F., 2016. The white matter query language: a novel approach for describing human white matter anatomy. Brain Struct. Funct. 221, 4705–4721. https://doi.org/10.1007/s00429-015-1179-4

Zhang, W., Olivi, A., Hertig, S.J., van Zijl, P., Mori, S., 2008. Automated fiber tracking of human brain white matter using diffusion tensor imaging. Neuroimage 42, 771–777. https://doi.org/10.1016/j.neuroimage.2008.04.241

Zhao, J., Thiebaut de Schotten, M., Altarelli, I., Dubois, J., Ramus, F., 2016. Altered hemispheric lateralization of white matter pathways in developmental dyslexia: Evidence from spherical deconvolution tractography. Cortex 76, 51–62. https://doi.org/10.1016/j.cortex.2015.12.004

	Tract	Abbreviation	Left/Right?	Seeding Strategy
Association Fibres	Arcuate Fasciculus	AF	Yes	Reverse-seeding
	Frontal Aslant Tract	FA	Yes	Single-ROI
	Inferior Longitudinal Fasciculus	ILF	Yes	Reverse-seeding
	Inferior Fronto-Occipital fasciculus	IFO	Yes	Reverse-seeding
	Middle Longitudinal Fasciculus	MdLF	Yes	Reverse-seeding
	Superior Longitudinal Fasciculus I	SLF1	Yes	Single-ROI
	Superior Longitudinal Fasciculus II	SLF2	Yes	Single-ROI
	Superior Longitudinal Fasciculus III	SLF3	Yes	Single-ROI
	Uncinate Fasciculus	UF	Yes	Single-ROI
	Vertical Occipital Fasciculus	VOF	Yes	Reverse-seeding
Commissural Fibres	Anterior Commissure	AC	No	Reverse-seeding
	Forceps Major	FMA	No	Reverse-seeding
	Forceps Minor	FMI	No	Reverse-seeding
	Middle Cerebellar Peduncle	MCP	No	Reverse-seeding
Limbic Fibres	Cingulum subsection: Dorsal	CBD	Yes	Single-ROI
	Cingulum subsection: Parahippocampal	CBP	Yes	Single-ROI
	Cingulum subsection: Temporal	CBT	Yes	Single-ROI
	Fornix	FX	Yes	Single-ROI
Projection Fibres	Acoustic Radiation	AR	Yes	Reverse-seeding
	Anterior Thalamic Radiation	ATR	Yes	Single-ROI
	Corticospinal Tract	CST	Yes	Single-ROI
	Optic Radiation	OR	Yes	Reverse-seeding
	Superior Thalamic Radiation	STR	Yes	Single-ROI

Table *1*. The list of reconstructed WM tracts, their correspodning tract type and the seeding strategy used.

Figures

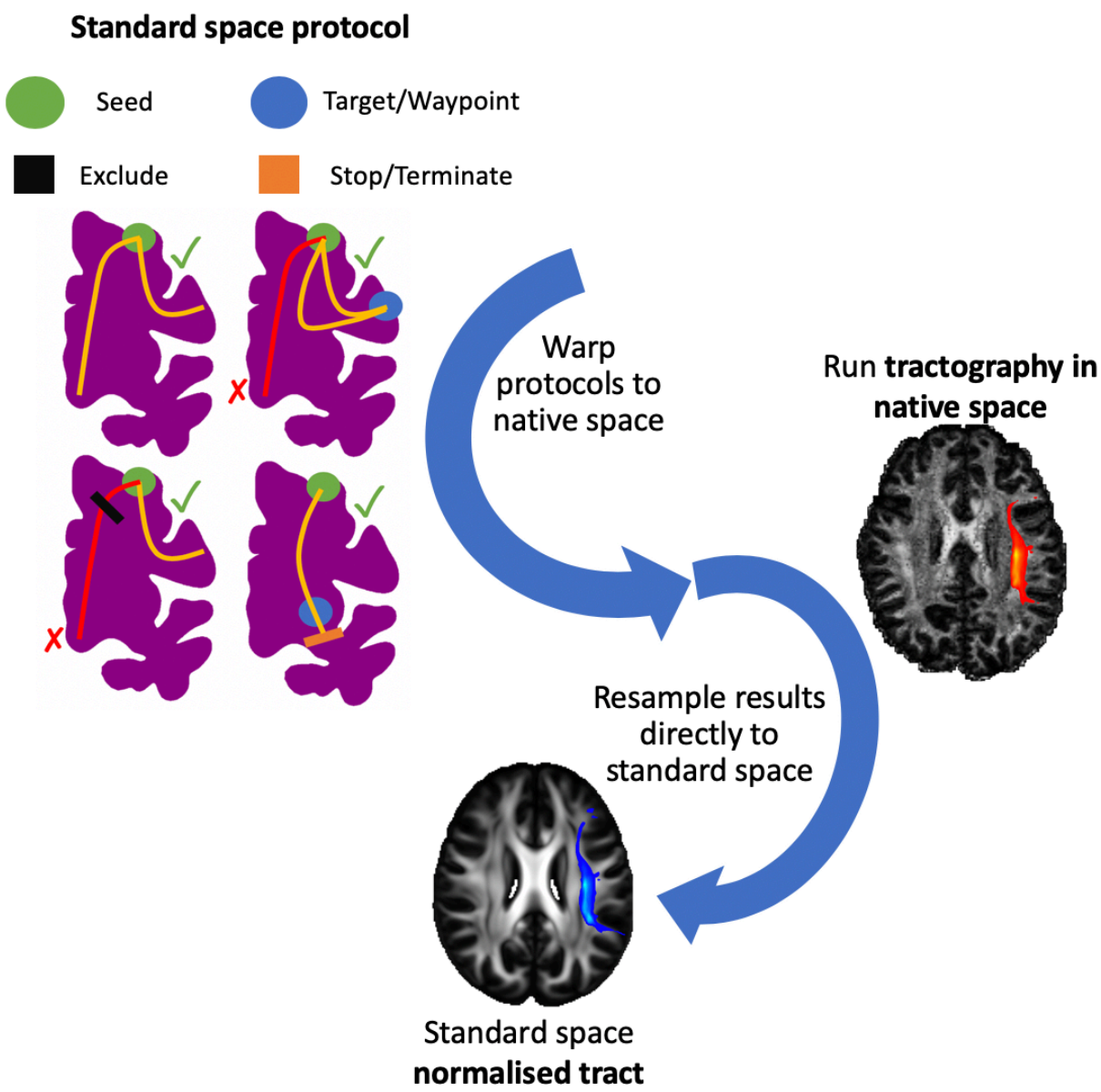


Figure 1. Schematic of the stages for automated tractography, as implemented in the "XTRACT" toolbox, with an example of the left arcuate fasciculus (AF) for the human brain. 1) Tractography protocol masks are defined in standard space with seed (green), exclusion (black), waypoint (blue) and termination (orange) masks (see the "Protocols" section for full details of definitions). 2) The protocol masks are warped to the subject's native space using the subject-specific non-linear warp fields. 3) Probabilistic tractography is performed in the subject's native space using the crossing fibre modelled diffusion data. Notice that results are mapped directly into standard space, leading to a single interpolation step. 4) The resultant tract stored in standard space, overlaid on the FSL_HCP1065 FA atlas.

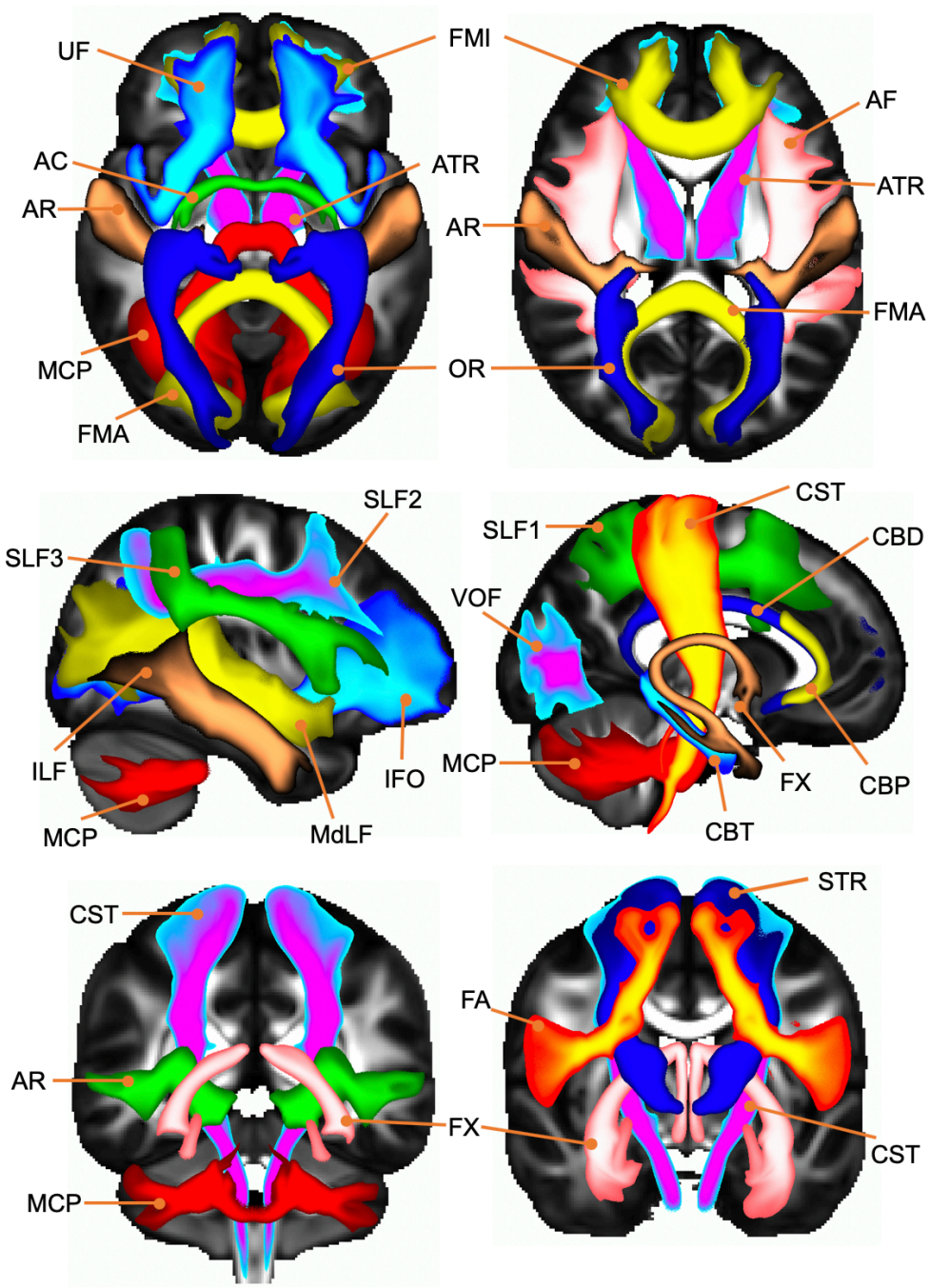


Figure 2. Axial, sagittal and coronal maximal intensity projections of the population percentage tract atlases (varying maximal intensity projection window lengths are applied to different tracts for visualisation purposes, display range = 0.3-1.0). Association fibre bundles: Arcuate Fasciculus (AF), Frontal Aslant Tract (FA), Inferior Longitudinal Fasciculus (ILF), Inferior Fronto-Occipital Fasciculus (IFO), Middle Longitudinal Fasciculus (MdLF), Superior Longitudinal Fasciculus I, II and III (SLF), Uncinate Fasciculus (UF) and Vertical Occipital Fasciculus (VOF). Projection fibre bundles: Acoustic Radiation (AR), Anterior Thalamic Radiation (ATR), Corticospinal Tract (CST), Optic Radiation (OR) and Superior Thalamic Radiation (STR). Limbic fibre bundles: Cingulum Bundle: Parahippocampal (CBP), Cingulum Bundle: Temporal (CBT), Cingulum Bundle: Dorsal (CBD) and Fornix (FX). Commissural fibre bundles: Anterior Commissure (AC), Forceps Major (FMA) and Forceps Minor (FMI). Tract atlases are created by averaging binarised (threshold of 0.1%) normalised tract density maps across subjects.

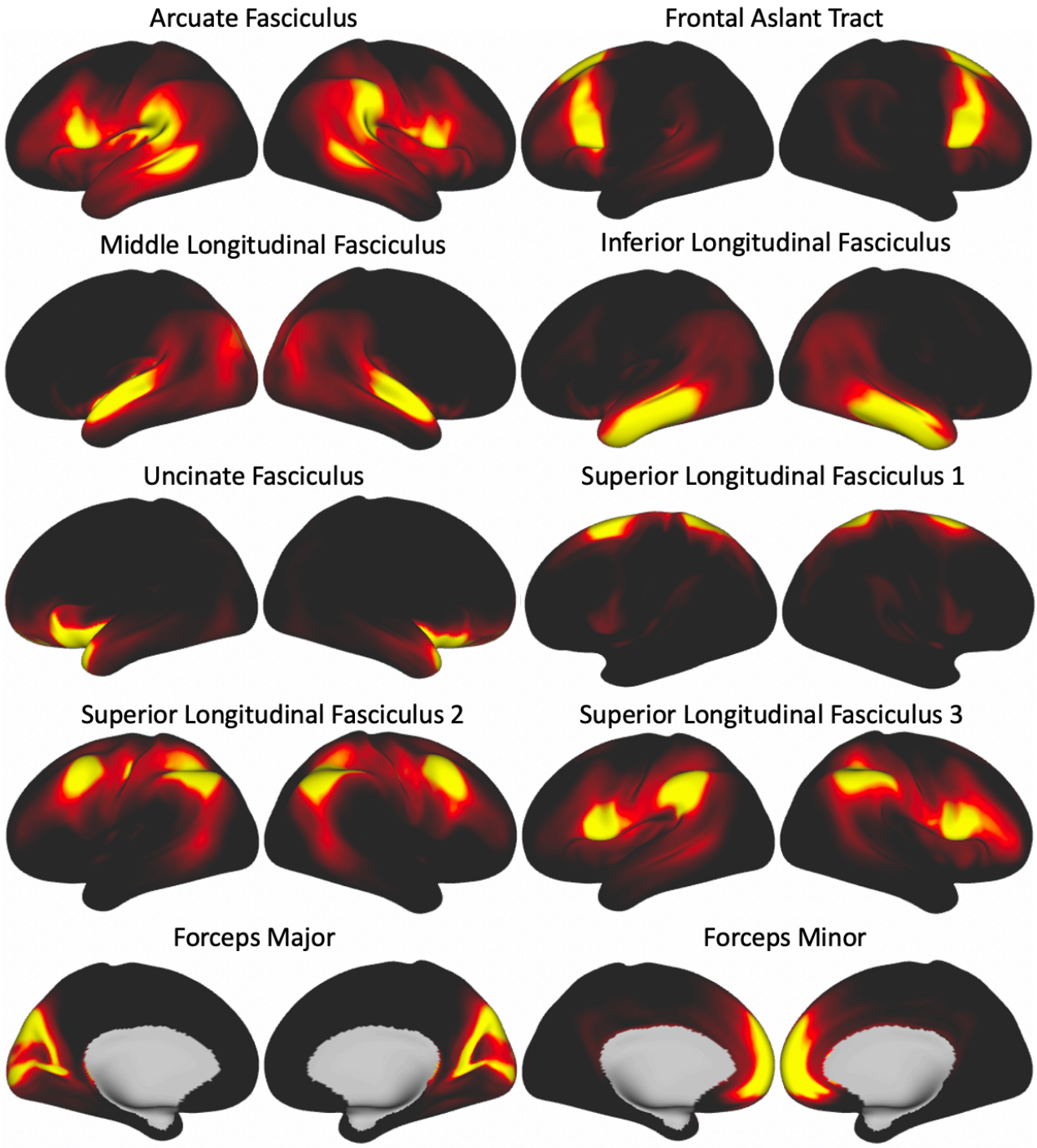


Figure 3. WGB endpoints for a subset of tracts (i.e. columns of the average connectivity blueprint) derived from the HCP cohort.

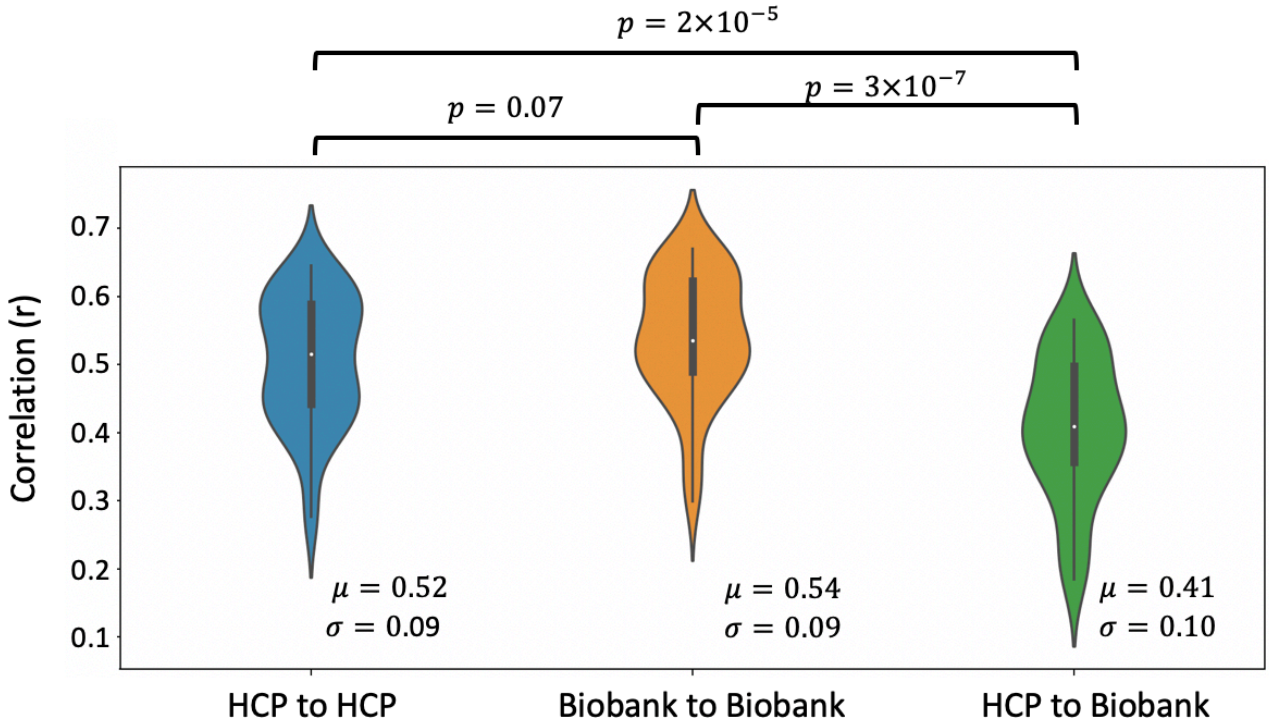


Figure 4. Summary of inter-cohort robustness. Plots of the correlations between 339 subject pairs within and across cohorts. Correlations are performed on normalised tract density maps with a threshold of 0.5%. μ is the median of the correlations across tracts and subject pairs and σ is the standard deviation. Significance is obtained via Mann-Whitney U test. Corrected p-value is 0.05/3 = 0.017.

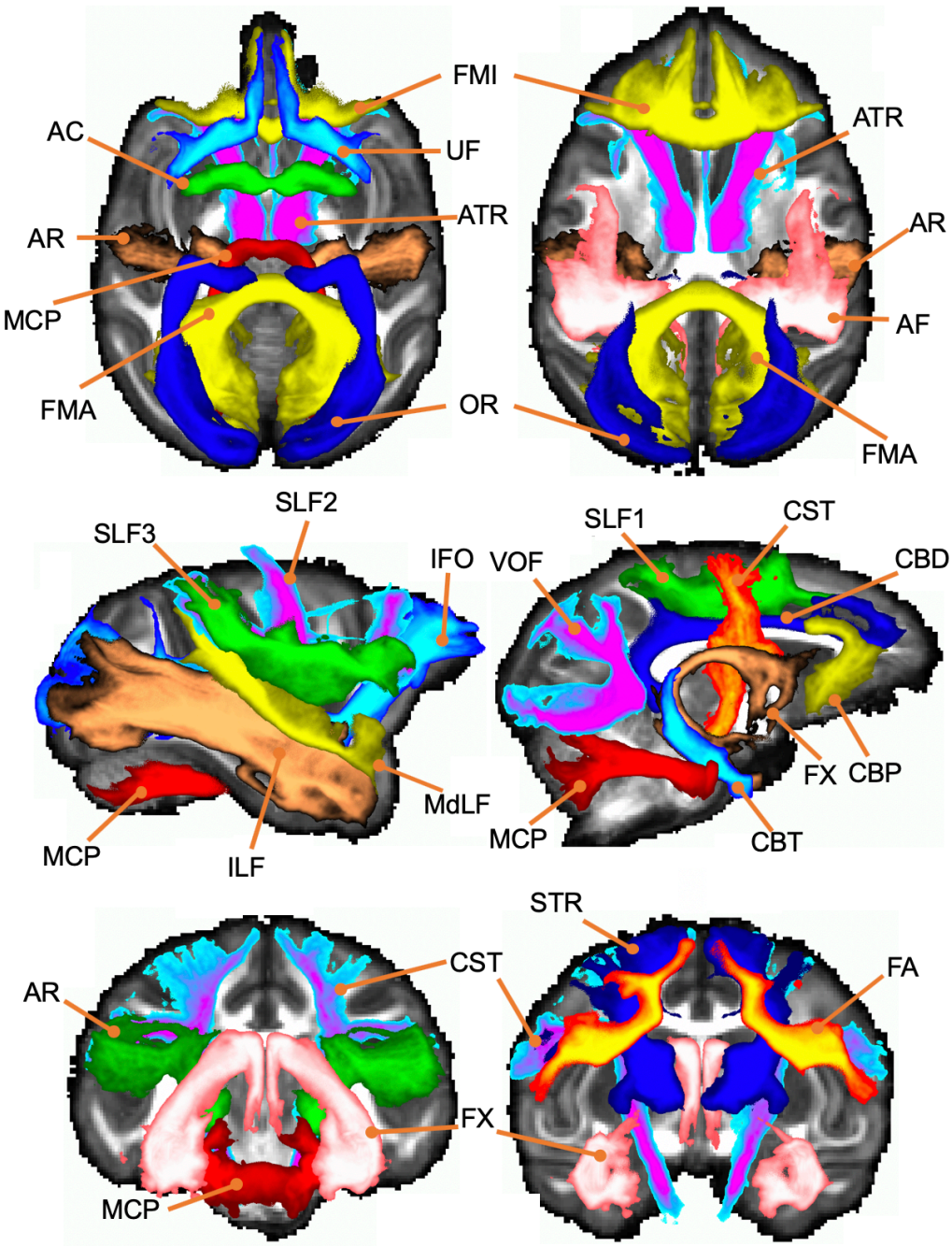


Figure 5. Axial, sagittal and coronal maximal intensity projections of the population percentage tract atlases (display range = 0.3-1.0) for the macaque subjects. Association fibre bundles: Arcuate Fasciculus (AF), Frontal Aslant Tract (FA), Inferior Longitudinal Fasciculus (ILF), Inferior Fronto-Occipital Fasciculus (IFO), Middle Longitudinal Fasciculus (MdLF), Superior Longitudinal Fasciculus I, II and III (SLF), Uncinate Fasciculus (UF) and Vertical Occipital Fasciculus (VOF). Projection fibre bundles: Acoustic Radiation (AR), Anterior Thalamic Radiation (ATR), Corticospinal Tract (CST), Optic Radiation (OR) and Superior Thalamic Radiation (STR). Limbic fibre bundles: Cingulum Bundle: Parahippocampal (CBP), Cingulum Bundle: Temporal (CBT), Cingulum Bundle: Dorsal (CBD) and Fornix (FX). Commissural fibre bundles: Anterior Commissure (AC), Forceps Major (FMA) and Forceps Minor (FMI). Tract atlases are created by averaging binarised (threshold of 0.1%) normalised tract density maps across subjects.

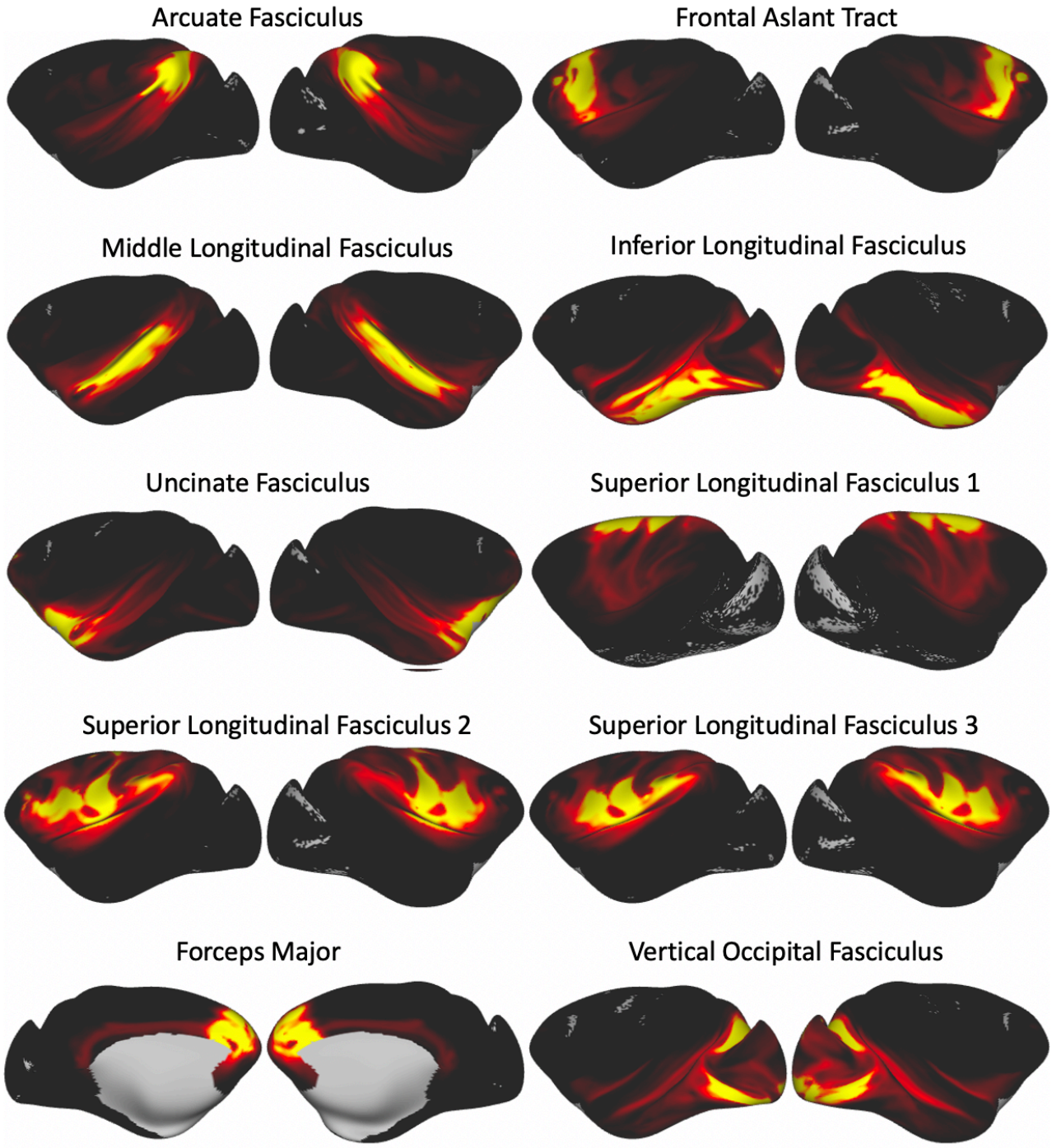


Figure 6. WGB endpoints for a subset of tracts (i.e. columns of the average connectivity blueprint) derived from the macaque subjects.

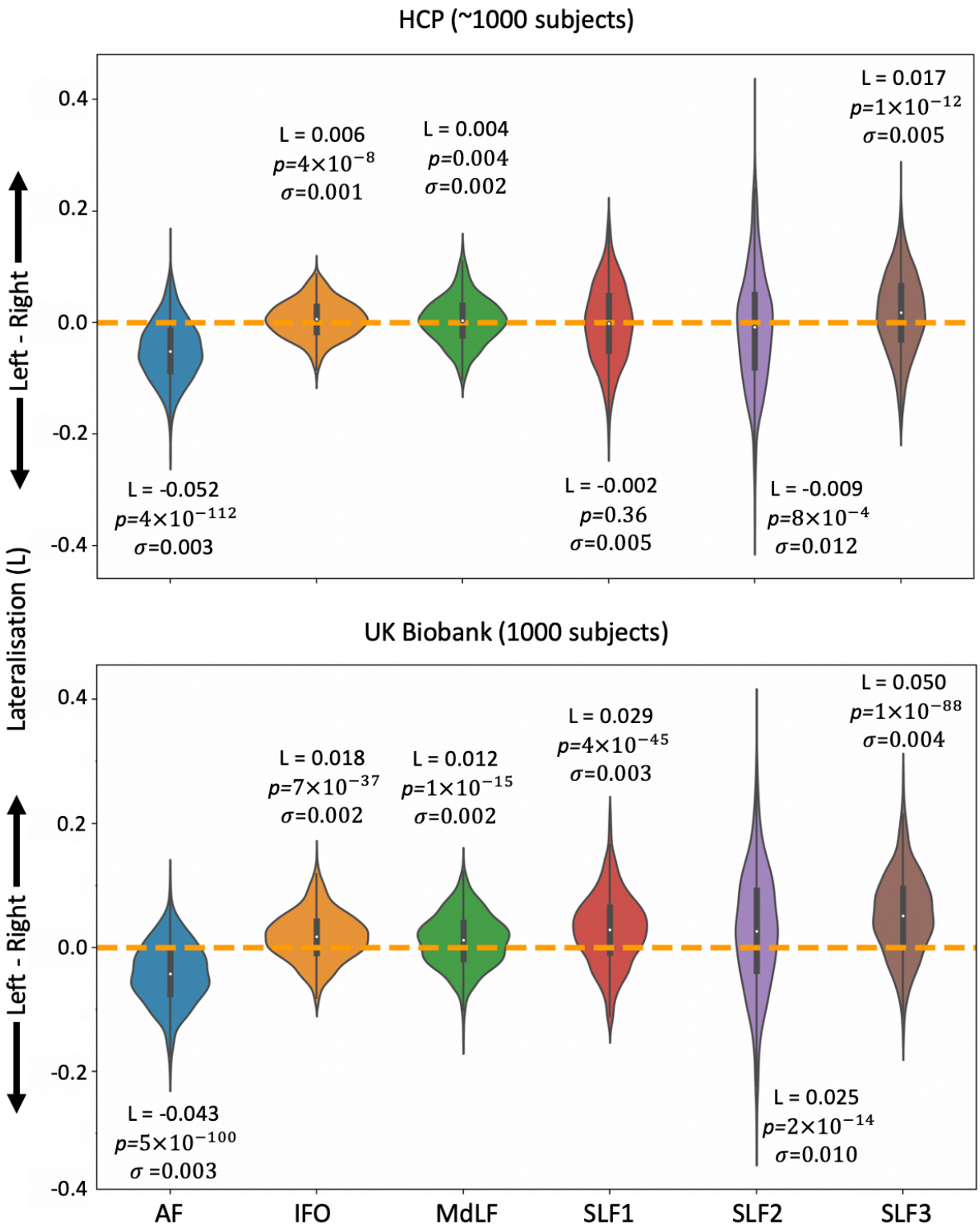


Figure 7. Summary of WM tract lateralisation for the arcuate fasciculus (AF), inferior fronto-occipital fasciculus (IFO), middle longitudinal fasciculus (MdLF) and the superior longitudinal fasciculi (SLFs) using the HCP (top) and UK Biobank (bottom) data. L is the cohort median WM tract lateralisation, p is the p-value obtained from the Mann-Whitney *U* test and σ is the variance for the given WM tract lateralisation. A threshold value of 0.5% has been used to binarise tracts and obtain their volume. Corrected p-value is 0.05/12 = 0.0042.

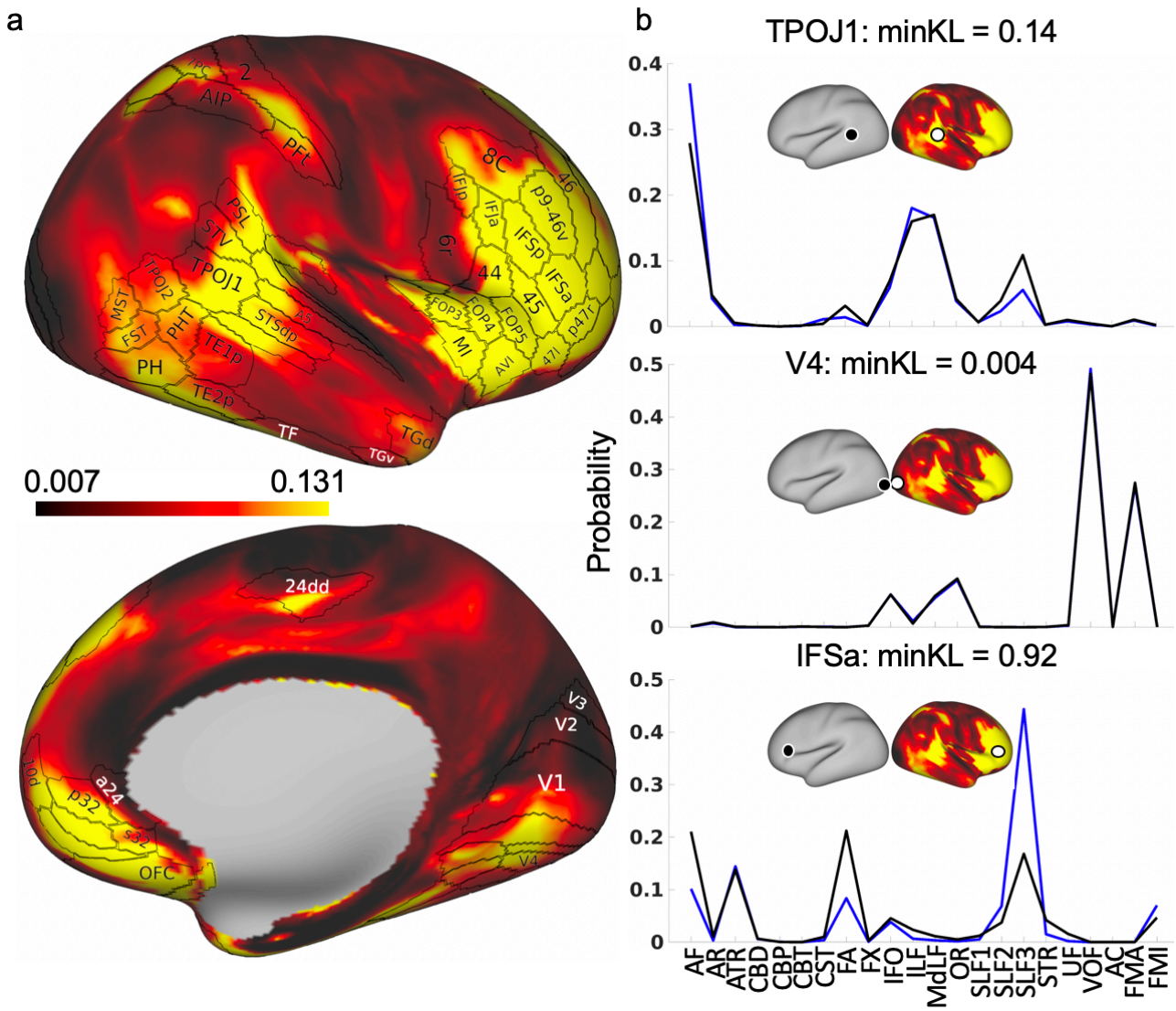


Figure 8. The minimum Kullback-Leibler (KL) divergence between the right and left hemispheres. (a) The minimum KL divergence on the right hemisphere with a subset of the Glasser parcellation (Glasser et al., Nature 2016) highlighting regions of dissimilarity. (b) By selecting vertices of interest from the right hemisphere (white circle) in (a) and extracting the vertex in the left hemisphere (black circle) with the greatest similarity, it is possible to investigate how differences in tract contribution to location connectivity contribute to divergence. Black lines correspond to the tract contributions to the vertex on the right hemisphere and blue lines for the left hemisphere. For example, in V4 (middle), the minimum KL-divergence is small which is reflected by the almost identical underlying tract contributions. Regions with mid- to high-range dissimilarity – TPOJ1 (top) and IFSa (bottom) – are seen to have greater differences in their underlying tract contributions, primarily driven by differences in AF and SLFs.

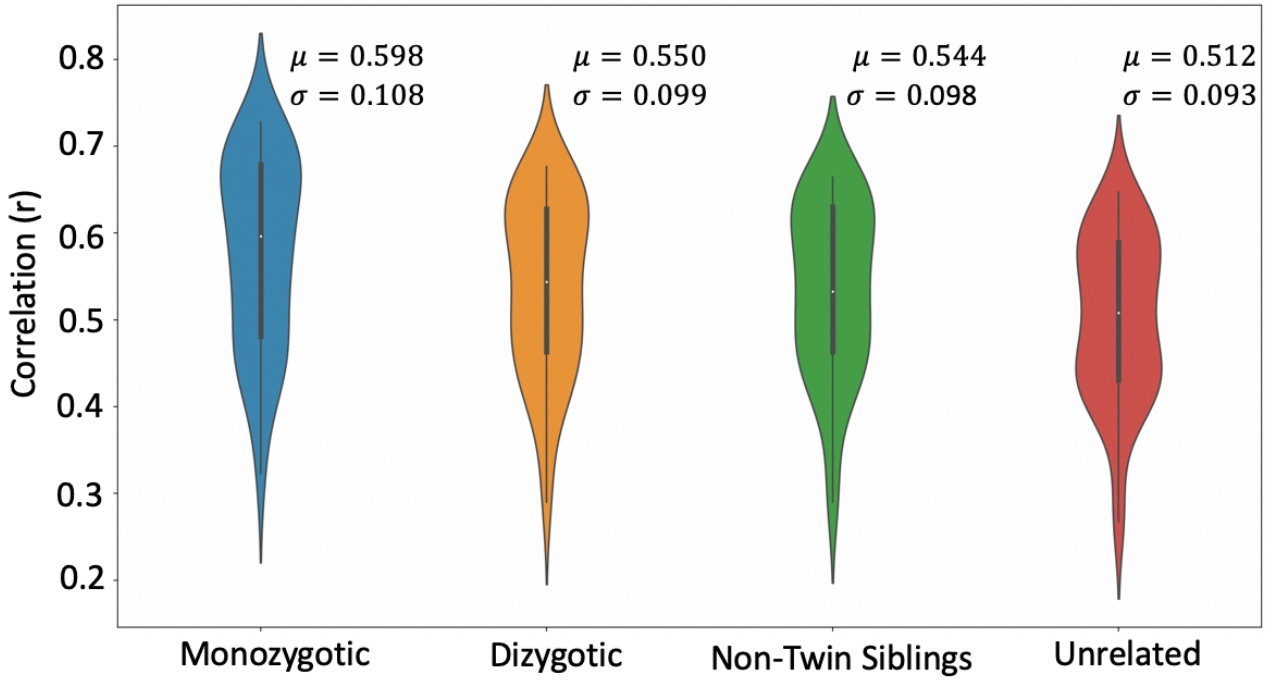


Figure 9. Twin/non-twin WM tract similarity using 72 subject pairs per group. Correlations are performed on normalised tract density maps with a threshold of 0.5%. μ is the group median across tracts and subjects and σ is the standard deviation. A Kruskal-Wallis test is used to determine whether the groups come from the same median: $\chi^2 = 13.1$, $p = 0.0043$, mean ranks = 104.1 (monozygotic), 85.3 (dizygotic), 82.9 (non-twin siblings) and 65.7 (unrelated).

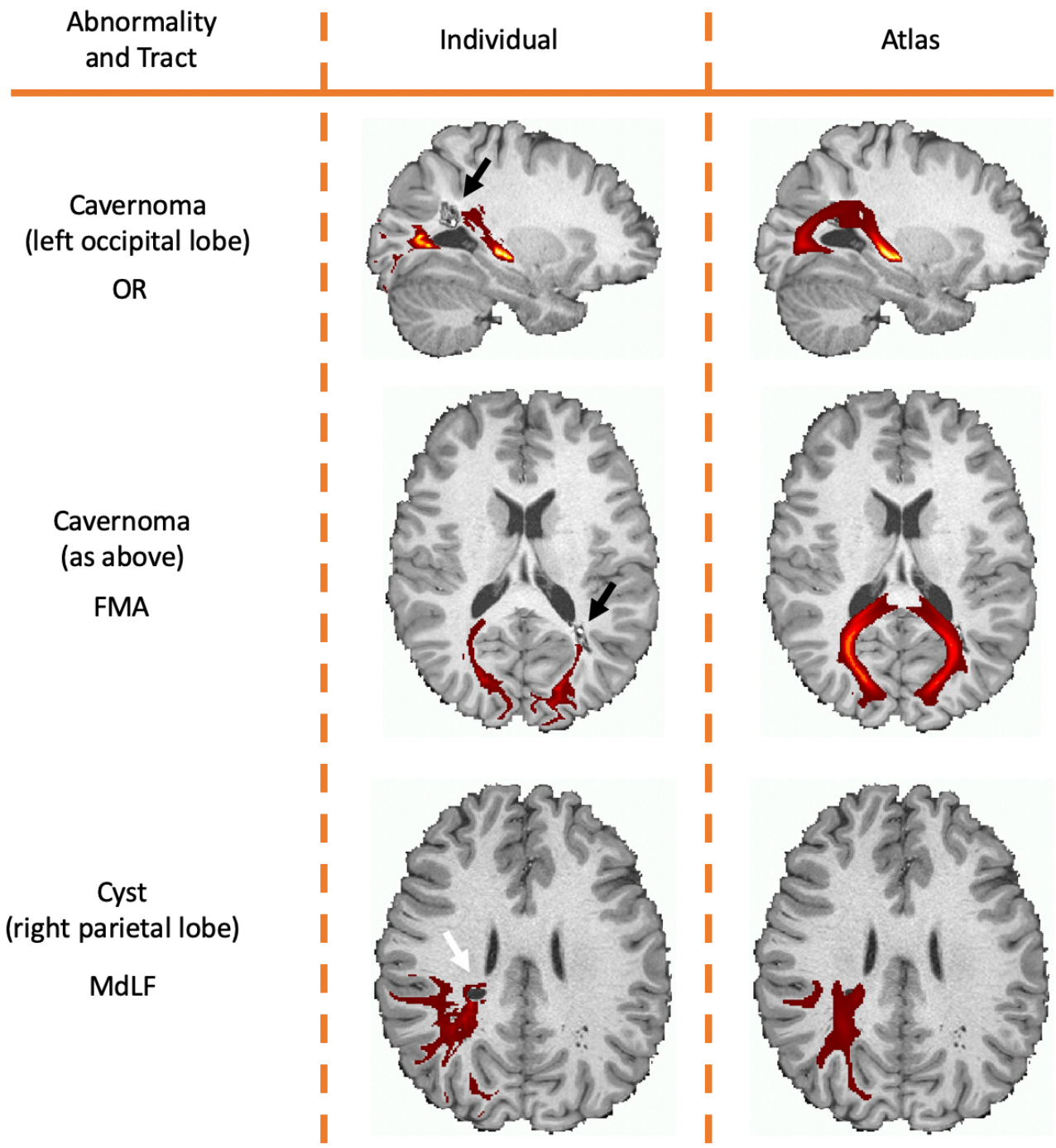


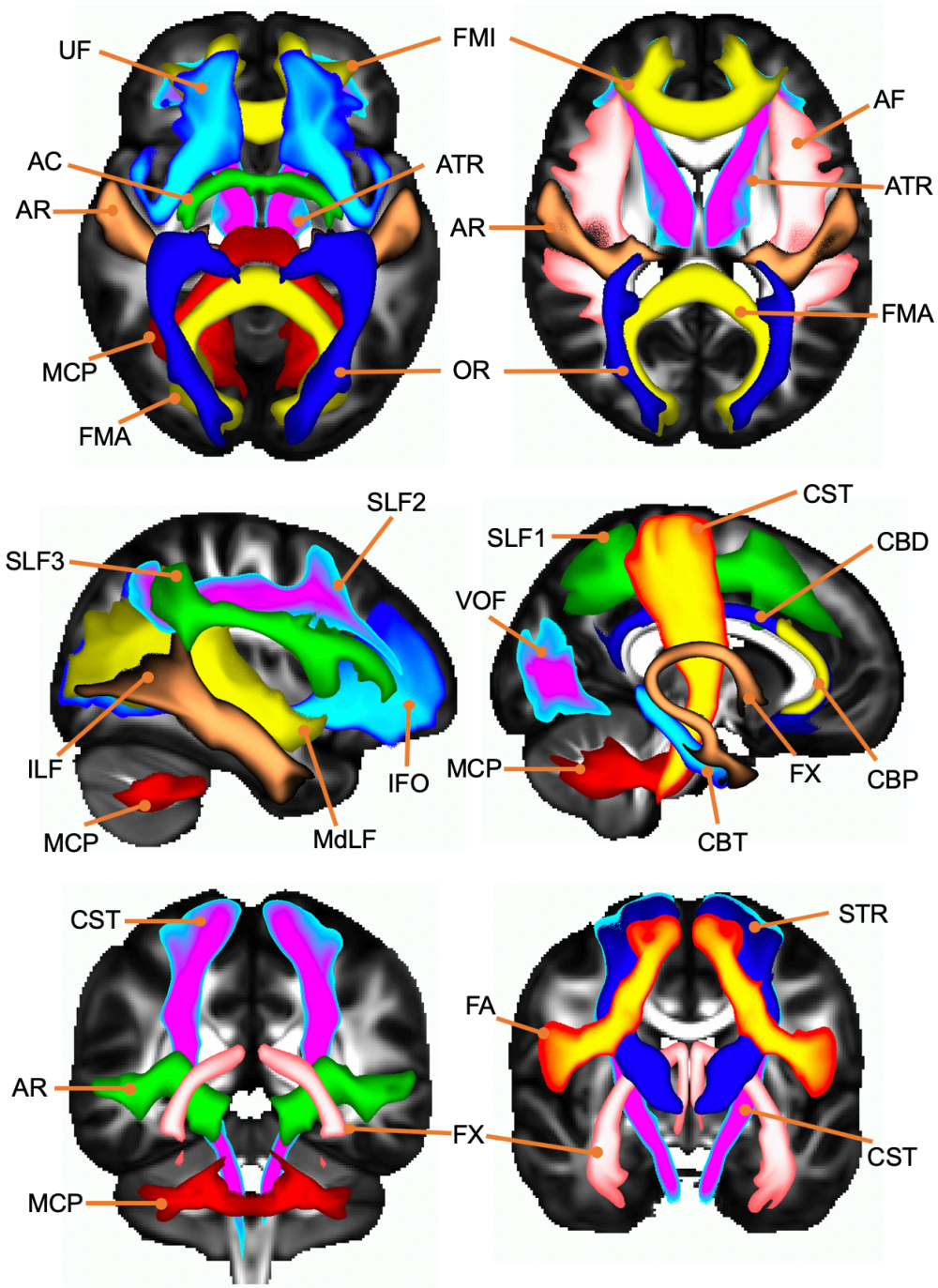
Figure 10. Examples of tractography results for a subset of the subjects found to have anatomical abnormalities. The left column describes the anatomical abnormality (as described on the HCP quality control website) and the affected tract (OR: Optic Radiation, FMA: Forceps Major, MdLF: Middle Longitudinal Fasciculus). The middle column shows the individual subject's extracted tract overlaid on the individual T1-weighted scan. The right column shows the subject's scan with the tract atlas overlaid, demonstrating that the cohort-averaged tract does not respect the observed variation. Tracts are displayed with a threshold of 0.1.

Supplementary

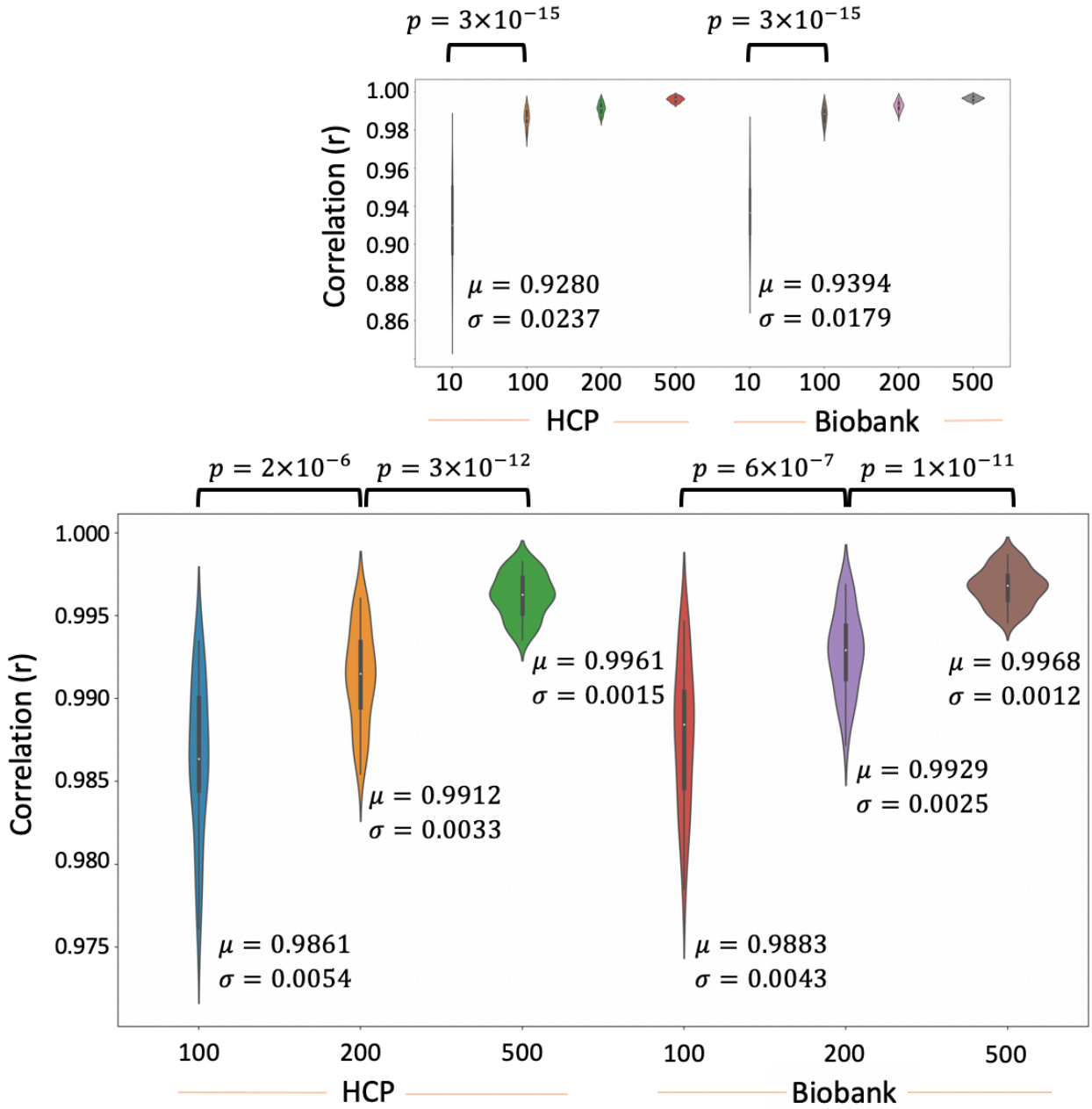
Tract	Wakana[1]	Catani[2]	Thiebaut de Schotten[3]	de Groot[4]	Wassermann[5]
Acoustic Radiation				✓	
Anterior Commissure		✓	✓		
Anterior Thalamic Radiation	✓			✓	
Arcuate Fasciculus		✓	✓ (Arcuate and anterior, long and posterior segments)		✓
Cingulum subsection: Dorsal	✓	✓	✓	✓	✓
Cingulum subsection: Parahippocampal	✓			✓	
Cingulum subsection: Temporal					
Corpus Collosum		✓	✓		✓ (Rostrum, rostral body, anterior midbody, posterior midbody, isthmus)
Corpus Collosum: Forceps Major	✓			✓	✓
Corpus Collosum: Forceps Minor	✓			✓	✓
Cortico-Ponto-Cerebellar			✓		
Corticospinal Tract	✓		✓	✓	✓
Extreme Capsule					✓
Fornix		✓	✓		
Inferior Cerebellar Peduncle		✓			
Inferior Fronto-Occipital Fasciculus	✓	✓	✓	✓	✓
Inferior Longitudinal Fasciculus	✓	✓	✓	✓	✓
Internal Capsule/Corona Radiata		✓	✓		
Medial Lemniscus				✓	
Middle Cerebellar Peduncle		✓		✓	
Middle Longitudinal Fasciculus					✓
Optic Radiation/Posterior Thalamic Radiation			✓	✓	✓
Spino-Cerebellar Tract			✓		
Striato-Fronto-Orbital					✓
Striato-Prefrontal					✓

Striato-Premotor					✓
Striato-Precentral					✓
Striato-Postcentral					✓
Striato-Parietal					✓
Striato-Occipital					✓
Superior Cerebellar Peduncle		✓	✓		
Superior Longitudinal Fasciculus I	✓ (SLF and the temporal component of the SLF)			✓	✓
Superior Longitudinal Fasciculus II					✓
Superior Longitudinal Fasciculus III					✓
Superior Thalamic Radiation				✓	
Thalamo-Prefrontal					✓
Thalamo-Premotor					✓
Thalamo-Precentral					✓
Thalamo-Postcentral					✓
Thalamo-Parietal					✓
Thalamo-Occipital					✓
Uncinate Fasciculus	✓	✓	✓	✓	✓
Total:	20	19	31	27	57

Supplementary Table *1*. A brief review of the protocols previously defined in the literature. [1] - Wakana, S., Jiang, H., Nagae-Poetscher, L.M., van Zijl, P.C., Mori, S., 2004. Fiber Tract–based Atlas of Human White Matter Anatomy. Radiology. [2] - Catani, M., Thiebaut de Schotten, M., 2008. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. Cortex. [3] - Thiebaut de Schotten, M., Ffytche, D.H., Bizzi, A., Dell'Acqua, F., Allin, M., Walshe, M., Murray, R., Williams, S.C., Murphy, D.G., Catani, M., 2011. Atlasing location, asymmetry and inter-subject variability of white matter tracts in the human brain with MR diffusion tractography. NeuroImage. [4] - de Groot, M., Vernooij, M.W., Klein, S., Ikram, M.A., Vos, F.M., Smith, S.M., Niessen, W.J., Andersson, J.L., 2013. Improving alignment in Tract-based spatial statistics: Evaluation and optimization of image registration. NeuroImage. [5] - Wassermann, D., Makris, N., Rathi, Y., Shenton, M., Kikinis, R., Kubicki, M., Westin, C.F., 2016. The white matter query language: a novel approach for describing human white matter anatomy. Brain Struct. Funct.



Supplementary Figure 1. Axial, sagittal and coronal maximal intensity projections of the population percentage tract atlases (display range = 0.3-1.0) for the UK Biobank subset. Association fibre bundles: Arcuate Fasciculus (AF), Frontal Aslant Tract (FA), Inferior Longitudinal Fasciculus (ILF), Inferior Fronto-Occipital Fasciculus (IFO), Middle Longitudinal Fasciculus (MdLF), Superior Longitudinal Fasciculus I, II and III (SLF), Uncinate Fasciculus (UF) and Vertical Occipital Fasciculus (VOF). Projection fibre bundles: Acoustic Radiation (AR), Anterior Thalamic Radiation (ATR), Corticospinal Tract (CST), Optic Radiation (OR) and Superior Thalamic Radiation (STR). Limbic fibre bundles: Cingulum Bundle: Parahippocampal (CBP), Cingulum Bundle: Temporal (CBT), Cingulum Bundle: Dorsal (CBD) and Fornix (FX). Commissural fibre bundles: Anterior Commissure (AC), Forceps Major (FMA) and Forceps Minor (FMI). Tract atlases are created by averaging binarised (threshold of 0.1%) normalised tract density maps across subjects.



Supplementary Figure 2. Summary of the tract-wise correlations of varying sample-size atlases to 1000-subject atlas for the HCP and the UK Biobank datasets. Each atlas is created by binarising (threshold of 0.1%) each of the subject's normalised tract density maps and then averaging across subjects. The full range of sample sizes used (10, 100, 200 and 500) is shown at the top and a zoomed-in version highlighting the differences between the larger sample sizes (100, 200 and 500) is shown at the bottom. μ is the group mean across tracts and σ is the standard deviation. Significance is obtained via Mann-Whitney U test. Corrected p-value is 0.05/6 = 0.0083.