The BigMac dataset: an open resource combining multi-contrast MRI and microscopy in the macaque brain

Amy FD Howard[a], Istvan N Huszar[a], Adele Smart[a,e], Michiel Cottaar[a], Greg Daubney[b], Taylor Hanayik[a], Alexandre A Khrapitchev[c], Rogier B Mars[a,d], Jeroen Mollink[a], Connor Scott[e], Nicola R Sibson[c], Jerome Sallet[b], Saad Jbabdi[☆a], Karla L Miller[☆a]

[a] *Wellcome Centre for Integrative Neuroimaging, FMRIB Centre, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom*

[b] *Wellcome Centre for Integrative Neuroimaging, Experimental Psychology, Medical Sciences Division, University of Oxford, Oxford, United Kingdom*

[c] *Department of Oncology, University of Oxford, Oxford, United Kingdom*

[d] *Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands*

[e] *Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom*

Abstract

Understanding brain structure and function often requires combining data across different modalities and scales to link microscale cellular structures to macroscale features of whole brain organisation. Here we introduce the BigMac dataset, a resource combining in vivo MRI, extensive postmortem MRI and multi-contrast microscopy for multimodal characterisation of a single whole macaque brain. The data spans modalities (MRI and microscopy), tissue states (in vivo and postmortem), and four orders of spatial magnitude, from microscopy images with micrometre or sub-micrometre resolution, to MRI signals on the order of millimetres. Crucially, the MRI and microscopy images are carefully co-registered together to facilitate quantitative multimodal analyses. Here we detail the acquisition, curation, and first release of the data, that together make BigMac a unique, openly-disseminated resource available to researchers worldwide. Further, we demonstrate example analyses and opportunities afforded by the data, including improvement of connectivity estimates from ultra-high angular resolution diffusion MRI, neuroanatomical insight provided by polarised light imaging and myelin-stained histology, and the joint analysis of MRI and microscopy data for reconstruction of the microscopy-inspired connectome. All data and code are made openly available.

Keywords: Multi-scale neuroscience; Multimodal data; Neuroanatomy; MRI; Diffusion MRI; Microscopy; Histology; Polarised light imaging; Data fusion;

[☆]Indicates equal contribution.

1. Introduction

Our ability to characterise brain connectivity has been greatly advanced by the scientific communitys open access to big data. Big data can employ large cohorts to examine both inter- and intra-subject variability (in e.g. the UK Biobank [1] or Human Connectome Project [2]), or aim to characterise a single brain in exquisite detail [3]. Here we introduce The BigMac Dataset - **Big** data in a **Mac**aque brain - a resource that combines in vivo MRI with extensive postmortem MRI and whole-brain, multi-contrast microscopy data in a single macaque brain. BigMac consists of in vivo data acquired over multiple sessions, over 270 hours postmortem MRI scanning, over 1000 hours of microscopy data acquisition and several terabytes of raw data.

The BigMac dataset combines multimodal data from both MRI and microscopy to explicitly address issues of sensitivity and specificity in MRI. MRI is a powerful non-invasive method that can inform on whole-brain structure and function, which can in turn be related to cognition, behaviour or medical outcomes. However, MRI also faces several limitations. In vivo signals are typically noisy and confounded by artefacts due to physiological effects and technical bottlenecks related to hardware limitations or the requirement for short scan times. Furthermore, MRI signals are often an indirect measure of the brain features of interest, making interpretation challenging. Diffusion MRI maps the microscopic motion of water molecules as they randomly move through tissue, to infer structural connectivity or changes in cellular morphology. These analyses require complex computational signal modelling with many strong assumptions. Crucially, the measurements are averaged over millimetres of tissue and so inference on brain structure or function at the micrometer scale is hard, if not ill-posed. Consequently, characterisation of the connectome using MRI alone faces significant limitations.

Alternatively, connectome data can be acquired via light microscopy, which is frequently used to study brain structure with micrometre or sub-micrometre resolution. Typically, thin tissue sections are processed to visualise specific cellular structures where this high 'specificity' approach has applications from basic neuroanatomy to disease mechanisms. Microscopy is however often limited to interrogating small, ex vivo tissue sections and thus has limited applications in vivo. Nonetheless, when microscopy is combined with MRI, it affords the opportunity for multi-scale neuroscience, interconnecting microscopic cellular processes with macroscopic MRI signals.

Such research ideally requires data that i) combines complementary MRI and microscopy data; ii) relates high-quality postmortem MRI and microscopy to in vivo MRI; iii) facilitates whole brain analysis with densely sampled data throughout; iv) has co-registered MRI and microscopy data for to facilitate meaningful voxelwise comparisons, and v) provides the above in a single specimen. The latter is essential as only when combining data from the same brain can we be sure that we are not over generalising, or ignoring important inter-subject variations. Few existing open datasets fulfil all these requirements. For example, existing data combining MRI and microscopy in the same tissue sample are often limited to small tissue sections [4, 5, 6], a single microscopy contrast or minimal MRI data [7], or compare data from different tissue samples, overlooking considerable between-brains variability. The BigMac dataset aims to address each of these goals, combining co-registered in vivo MRI, extensive postmortem MRI and whole-brain multi-contrast microscopy data in a single macaque brain.

The BigMac dataset is interesting from both an anatomical and methodological viewpoint. For those interested in microstructural neuroanatomy, the densely sampled multi-contrast microscopy can be used to examine both the myelo- and cyto-architecture in great detail and develop novel atlases or parcellations. For those interested in multi-scale neuroscience, BigMac provides whole brain multimodal data spanning four orders of magnitude with which we can link microscale cellular structures to macroscale features of brain organisation and function. For those interested in diffusion modelling, the comprehensively sampled diffusion MRI space - which is co-registered to microscopy data - can be used to drive protocol optimisation, advance computational modelling of the tissue microstructure or brain connectivity, and provide direct validation of

many current and future diffusion MRI models and analysis methods [8, 9]. Finally, for those interested in machine- or deep-learning approaches, the BigMac dataset will support the development of novel algorithms which jointly model MRI and microscopy data [10, 11], for example, to map quantitative imaging biomarkers to specific features of the tissue microstructure.

This paper documents the first release of the BigMac dataset. The open data includes in vivo and post-mortem MRI, as well as whole brain microscopy data from polarised light imaging [12, 13, 14] and myelin-stained histology [15], both of which provide detailed information about tissue myeloarchitecture [16]. Here we detail the multi-faceted data acquisition and curation, and conduct some of the first analyses which demonstrate the data quality and unique information or analyses afforded by BigMac.

2. Results & discussion

First we provide an overview of the data included in BigMac. We then explore the BigMac data from various viewpoints including i) how ultra-high angular resolution diffusion imaging affects connectivity estimates, ii) how microscopy can be used to detail the myeloarchitecture of the brain, iii) the quality of MRI-microscopy co-registration, iv) voxelwise comparisons of MRI and microscopy metrics, and v) a method for performing hybrid MRI-microscopy tractography. Here we analyse the postmortem data in BigMac, where translating our work to the in vivo domain is the focus of future work.

2.1. Data summary

Figure 1 gives an overview of the BigMac dataset, with further acquisition details provided in Table 1. Importantly, the BigMac dataset includes in vivo data combining behavioural data with diffusion MRI (1mm isotropic resolution), task fMRI (2mm), resting-state fMRI (2mm) and structural MRI (0.5mm) acquired over multiple scan sessions [17, 18]. This provides an excellent opportunity to link high-quality postmortem MRI and microscopy data (detailed structural connectivity) to in vivo data (structural and functional connectivity) within the same brain.

Ex vivo, a comprehensive MRI dataset (∼270 hours scanning time) was acquired which includes high resolution (0.3 mm isotropic) structural images and extensive diffusion MRI data at two spatial resolutions (0.6 and 1 mm isotropic). The 0.6mm data includes 128 gradient directions at $b = 4\,\text{ms}\,/\,\mu\text{m}^2$, whilst the 1mm data includes 250 gradient directions at $b = 4\,\text{ms}\,/\,\mu\text{m}^2$ and 1000 gradient directions at $b = 7$ and $10\,\text{ms}\,/\,\mu\text{m}^2$ ('ultra-HARDI data'). This ultra-HARDI data is complemented by diffusion data with spherical tensor encoding at b-values of $b = 4$, 7, & 10ms $/\ \mu\text{m}^2$ for advanced microstructural imaging. Images with negligible diffusion weighting ($b \sim 0\,\text{ms}\,/\,\mu\text{m}^2$) and a T_1 map were also acquired. The postmortem MR data in BigMac serves two specific aims. First, the postmortem data acts as a crucial intermediary between in vivo diffusion MRI and postmortem microscopy: the postmortem and in vivo MRI are similar in nature (i.e. both are MRI signals) but image the tissue in different states (postmortem, fixed tissue versus in vivo), whereas the postmortem MRI and microscopy share a similar tissue state, though the signals are different. Second, compared to in vivo MRI, the postmortem data is of particularly high quality. It benefits from being acquired on a small-bore preclinical scanner at higher field strength with exceptionally long scan times, and avoids signal instabilities from physiological movement that can fundamentally limit the signal-to-noise ratio (SNR).

After scanning, the entire brain was sectioned along the anterior-posterior axis, with consecutive sections allocated to one of six microscopy contrasts (Table 5.1). These sections were sequentially allocated to polarised light imaging, to visualise the orientation of myelinated fibres [12, 13, 14], histological staining of myelin (Gallyas silver) or Nissl bodies (Cresyl violet), or kept for complementary staining that is to be decided. These "unassigned sections" were returned to formalin and stored for longevity. The PLI and

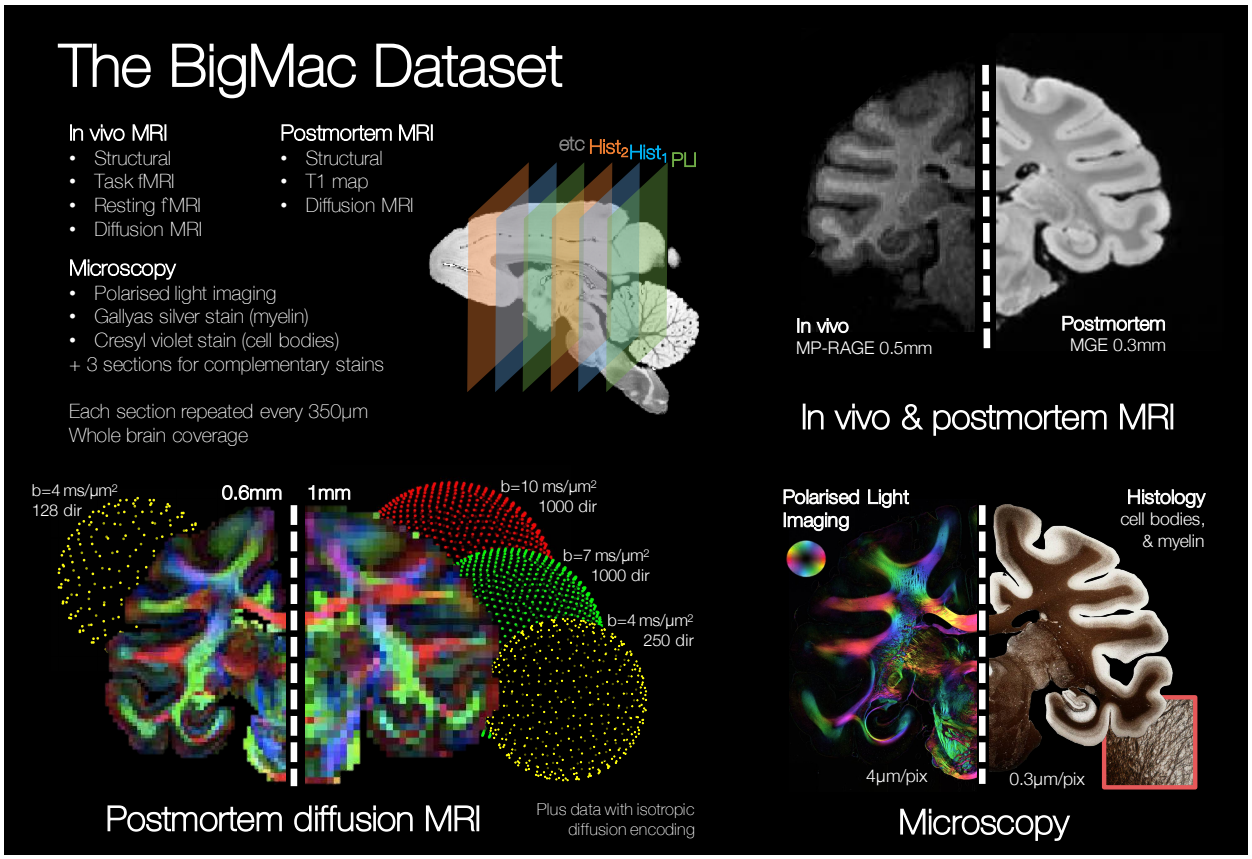


Figure 1: A summary of the data in the BigMac dataset. A single macaque monkey was scanned both in vivo and postmortem. The in vivo MRI data includes structural, task functional MRI (fMRI), resting fMRI and diffusion MRI. The postmortem diffusion MRI data contains up to 1,000 gradient directions for b-values of 4, 7 and 10 ms / µm^2 and two spatial resolutions (0.6 and 1 mm isotropic). A T_1 map and data with spherical tensor encoding were also collected. For comparison, polarised light imaging and histology provide high-resolution information about the white matter microstructure at 4 µm/pix and 0.28 µm/pix respectively. In total, the postmortem MRI acquisition constituted around 270 hours scanning time and the microscopy data took $>$ 1000 hours to acquire.

Postmortem MRI		Resolution (mm iso)	b-value (ms/μm^2)	# directions (linear)	# averages (spherical)	# b=0 volumes	δ/Δ (ms)	TE/TR (ms/s)
	Structural	0.3	-	-	-	-	-	7.8 / 0.097
	T_1 maps	0.6	-	-	-	-	-	8 / 10
a)	Diffusion MRI	0.6	4	128	-	8	7/13	25.4 / 10
	Diffusion MRI	1	4	250	-	10	14/24	42.5 / 3.5
b)	Diffusion MRI	1	7	1000	-	40	14/24	42.5 / 3.5
	Diffusion MRI	1	10	1000	-	40	14/24	42.5 / 3.5
	Diffusion MRI	1	4	-	30	1	-	42.5 / 6.4
	Diffusion MRI	1	4	50	-	2	14/24	42.5 / 6.4
c)	Diffusion MRI	1	7	-	30	1	-	42.5 / 6.4
	Diffusion MRI	1	7	50	-	2	14/24	42.5 / 6.4
	Diffusion MRI	1	10	-	30	1	-	42.5 / 6.4
	Diffusion MRI	1	10	50	-	2	14/24	42.5 / 6.4

a) High spatial resolution, b) Ultra-high angular resolution (ultra-HARDI), c) Combining linear & spherical tensor encoding

Microscopy	Thickness (μm)	Staining	Visualisation	Imaging resolution
Polarised Light Imaging	50	None	Myelinated fibres	4 μm/pix
Histology	50	Cresyl violet	Nissl bodies	-
Histology	50	-	-	-
Histology	50	Gallyas silver	Myelin	0.28 μm/pix
Histology	100	-	-	-
Unstained	50	-	-	-

Table 1: A summary of the postmortem MRI and multi-contrast microscopy data in the BigMac dataset. The postmortem diffusion data includes three protocols which can broadly be described as achieving high spatial resolution, ultra-high angular resolution, and combining linear and spherical tensor encoded data. The first data release includes PLI and myelin-stained histology (Gallyas); additional histological stains will be added to the BigMac resource as these data are acquired. iso=isotropic; #=number; DW=diffusion-weighted; TE=echo time; TR=repetition time; 'linear' and 'spherical' indicate diffusion MRI acquired with linear and spherical tensor encoding.

histology-stained slides were then imaged at high resolution (4 or 0.28 µm/pixel) and co-registered to the MRI [19, 20].

Through this extensive, multimodal, multi-contrast acquisition, the BigMac dataset aims to provide a detailed characterisation of the macaque connectome, where the data and analysis tools are openly disseminated to the scientific community.

2.2. Ultra-high angular resolution diffusion imaging

The postmortem diffusion MRI protocol had three main objectives (Table 5.1): i) high angular resolution imaging (orange), ii) high spatial resolution imaging (blue), and iii) combining linear and spherical tensor encoding (green). The diffusion MRI data includes two spatial resolutions: high resolution 0.6 mm data at $b = 4\,\mathrm{ms/\mu m^2}$ as well as 1 mm data at b-values of 4, 7 and $10\,\mathrm{ms/\mu m^2}$ (Figure 1). Here the 0.6 mm protocol was designed to strike a good balance between spatial and angular resolution. In comparison, the 1mm protocol facilitates the characterisation of complex fibre geometries through a more extreme sampling of q-space, with multiple b-values - where the reduced resolution provides increased SNR at high b - and either 250 or 1000 gradient directions per shell. In addition, at 1 mm we were able to acquire 1000 gradient directions ('ultra-HARDI data') in the outer two shells within an achievable scan time (< 1 week). Figure 2a shows the diffusion signal from a single voxel, demonstrating how with 1000 gradient directions we are able to examine the 3D signal profile in great detail.

We first examined whether having high angular resolution provided unique microstructural information with regards to fibre orientations and structural connectivity. Here, the Ball and Stick model [21, 22] was used to estimate the fibre orientations from data acquired at different b-values (b=4, 7 or 10 $\mathrm{ms/\mu m^2}$) and varied angular resolution ($64 - 1000$ gradient directions). Examining white matter voxels, Figure 2b shows how as the angular resolution of the data was increased, the Ball and Stick model estimated an increasing number of voxels with multiple fibre populations (Figure 2b top) and that the fibre orientations were estimated with increased precision (Figure 2b bottom). Both effects were most striking in the tertiary fibres, suggesting that high angular resolution data can resolve less dominant fibre populations which are overlooked in lower angular resolution data, and which could have important functional or structural profiles.

Differences in the estimated fibre populations from the Ball and Stick model from high angular resolution data may have downstream effects on tractography-based structural connectivity estimates. Here we hypothesised that increased angular resolution should result in superior fibre tracking through crossing fibre regions, improving the reconstruction of long range connections. Using probabilistic tractography (probtrackX [22]) we reconstructed a dense connectivity matrix for the 82 cortical regions of interest (ROIs) included in the Kötter and Wanke Regional Map (RM) parcellation [23, 24, 25]. Figure 2c shows the fractional increase in the number of streamlines from the matrix with 1000 gradient directions and that with 64 gradient directions ($(N_{1000} - N_{64})/N_{64}$). The ultra-high angular resolution dataset estimates more streamlines reaching almost every pair of ROIs. There is a particularly high number of additional streamlines between many interhemispheric connections (green box). A notable exception to this overall pattern is that parietal and premotor regions (coded yellow-orange on the depicted brain surface) do not show a large increase in interhemispheric connectivity with higher angular resolution (turquoise box) but do show increased connections to other regions in the same hemisphere (blue box). Many homotopic regions (where we expect there to be true connections) have a $>$ 1.8-fold increase, where we tend to see a larger effect in more lateral homotopic regions. Those with a $>$ 50-fold increase include the ventral part of the anterior visual area (VACv), the inferior temporal cortex (TCi) and the central temporal cortex (TCc). The high angular resolution data produced notably longer streamlines, with a $>$ 5-fold increase in the number of streamlines whose length was > 50 mm.

Finally, Figure 2d compares tractography reconstruction of the superior longitudinal fasciculus (SLF) II from data with high angular resolution (top), low angular resolution but high b-value (middle) or high spatial

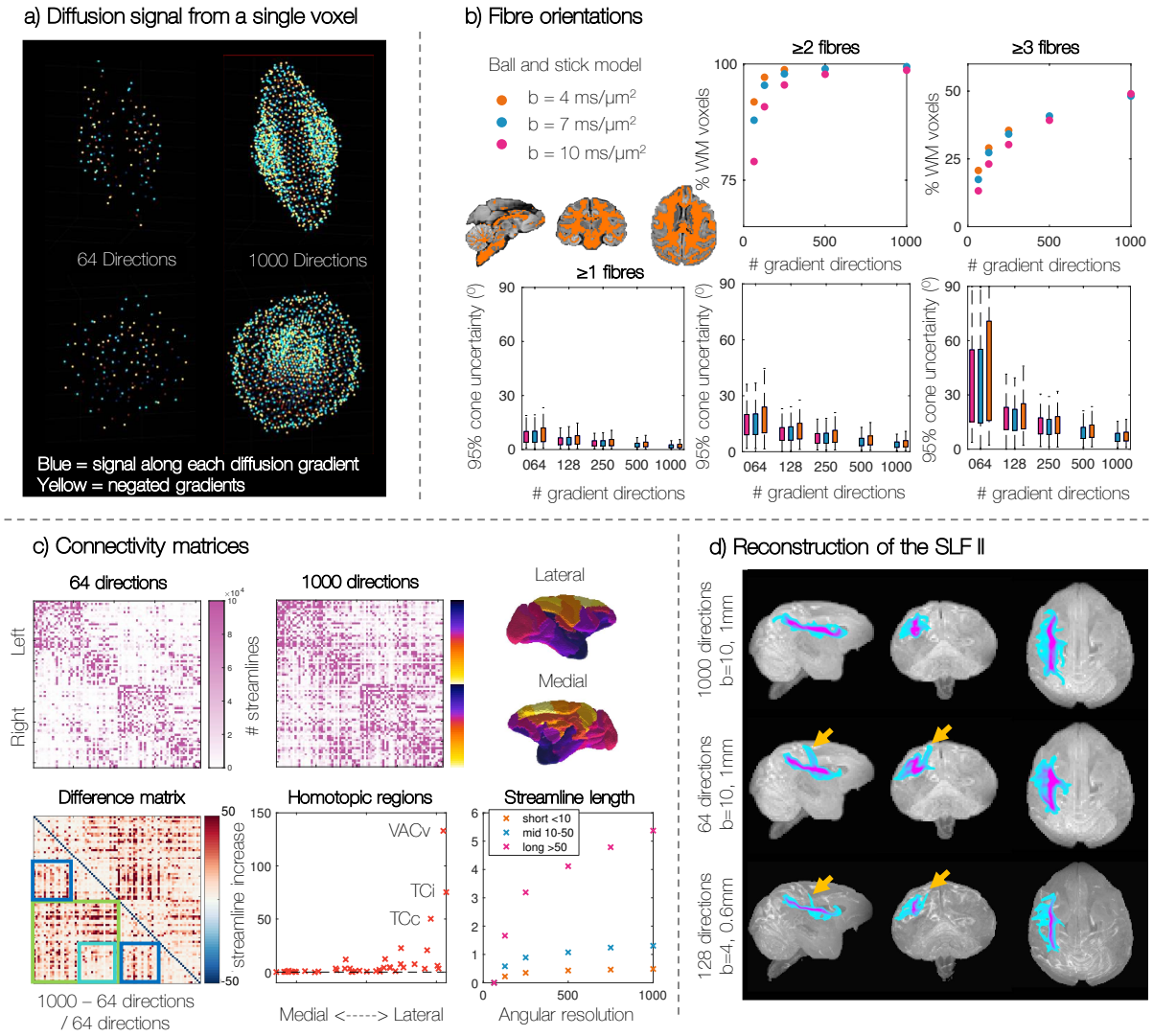


Figure 2: Ultra-high angular resolution data allows us to a) characterise the diffusion signal in great detail, b) find more tertiary populations and estimate fibre orientations with increased precision, c) reconstruct longer streamlines and increase inter-hemispheric connectivity, particularly to the occipital and temporal lobe, and d) reconstruct the SLF II with increased confidence. a) Each point represents the diffusion signal along a single gradient direction. The voxel is shown from two orthogonal views. b) The Ball and Stick model was fitted to data with varying angular resolution. Both the number of fibre populations, and the precision of the orientation estimates is plotted. c) Comparing connectivity matrices from 64 and 1000 gradient datasets. Since the difference matrix is symmetric, the bottom half is used to highlight interesting features: the green box indicates interhemispheric connectivity, with the turquoise box showing connections between parietal and premotor areas; the dark blue boxes show intrahemispheric connectivity of areas that are separated along the inferior-superior axis. Streamlines whose length < 10 mm are considered 'short', 10-50 mm are 'mid' length and > 50 mm are 'long'. d) Reconstruction of the superior longitudinal fasciculus (SLF) II using probabilistic tractography.

resolution (bottom). Although this tract is not difficult to reconstruct in human data, it's reconstruction can be challenging in the macaque. Here anatomically constrained probabilistic tractography was performed using XTRACT [26] with predefined seed, target and exclusion masks. In Figure 2d we use a fairly high threshold (0.01=1%) on the normalised tract density mask to isolate the tract centre (i.e. the voxels with the highest density of streamlines). In the ultra-HARDI data, the high angular resolution and contrast allows us to reconstruct the tract with a single, uniform core. We observe a higher density of streamlines extending the main tract body into the frontal and posterior regions, following the expected spatial extent and connectivity of the SLF II. Neither the tract from lower angular or high spatial resolution data have the same anterior-posterior reach. Instead these tracts appears to have a systematic false positive (yellow arrows) offshoot extending to the superior cortex, which may be indicative of streamlines crossing to the SLF I.

Together, these results suggest that higher angular resolution data may be advantageous when trying to reconstruct long range or inter-hemispheric connections which likely track through crossing fibre regions.

2.3. Microscopy data: characterising the detailed myeloarchitecture of the brain

After scanning, the whole brain was sectioned along the anterior-posterior axis, with consecutive sections allocated to either polarised light imaging [12, 13, 14] or one of five histology contrasts. Each microscopy contrast includes ~ 160 slides spanning the entire brain, with 350 µm separation between consecutive slides with the same contrast. As microscopy imaging takes many hundreds of hours, the first data release includes PLI and myelin-stained histology (Gallyas [15]) as detailed below; additional histological stains will be added to the BigMac resource as these data are acquired.

2.3.1. Polarised light imaging

Polarised light imaging (PLI) [12, 13, 14] utilises the birefringence of myelin to estimate the primary fibre orientation per microscopy pixel. Figure 3a shows example PLI from different locations along the anterior-posterior axis where image 1 is the most anterior and image 8 the most posterior. Here we see the myelinated fibres in high detail at a resolution of 4 µm per pixel. The orientations are colour-coded in HSV (hue-saturation-value) space, where the hue is dependent on the fibre orientation and the value is related to the tissue birefringence. The PLI data can, for example, track white matter fibres fanning across the cortex (A), bundles projecting between deep grey matter structures such as the globus pallidus and the putamen (B), as well as pontocerebellar fibres extending from the basilar sulcus and around the pons (C). Further the PLI can differentiate cortical layers which differ in the orientation and density of myelinated axons (A, white arrow).

In the cerebellum (D) we see i) clear delineation of the dentate nucleus (white arrow) and ii) how the PLI data separates the molecular and granular layers of the cerebellar cortex [27]; the molecular layer is the top, outermost layer of the cerebellar cortex (Figure 3b green arrows) and the granular layer (yellow arrows) lies between the molecular layer and the white matter. Notably, in many regions the molecular layer shows fairly strong birefringence, with in-plane angles parallel to the cortical surface. In comparison, the molecular layer at the gyri crown (blue arrows) tend to have low birefringence. Although the birefringence in the granular layer is low, we can observe individually identifiable fibres or fibre bundles which generally extend from the white matter and into the molecular layer. This pattern is observed in many but not all areas of the cerebellar cortex.

To hypothesise on the origin of the birefringence in the molecular layer, we should consider the known architecture of the cerebellar cortex (Figure 3b right) [27]. The granular layer (between the white matter and the molecular layer) contains cerebellar granule cells, some of the smallest but most numerous cells in the brain. The axon of the granule cell extends vertically into the molecular layer, where it then splits into

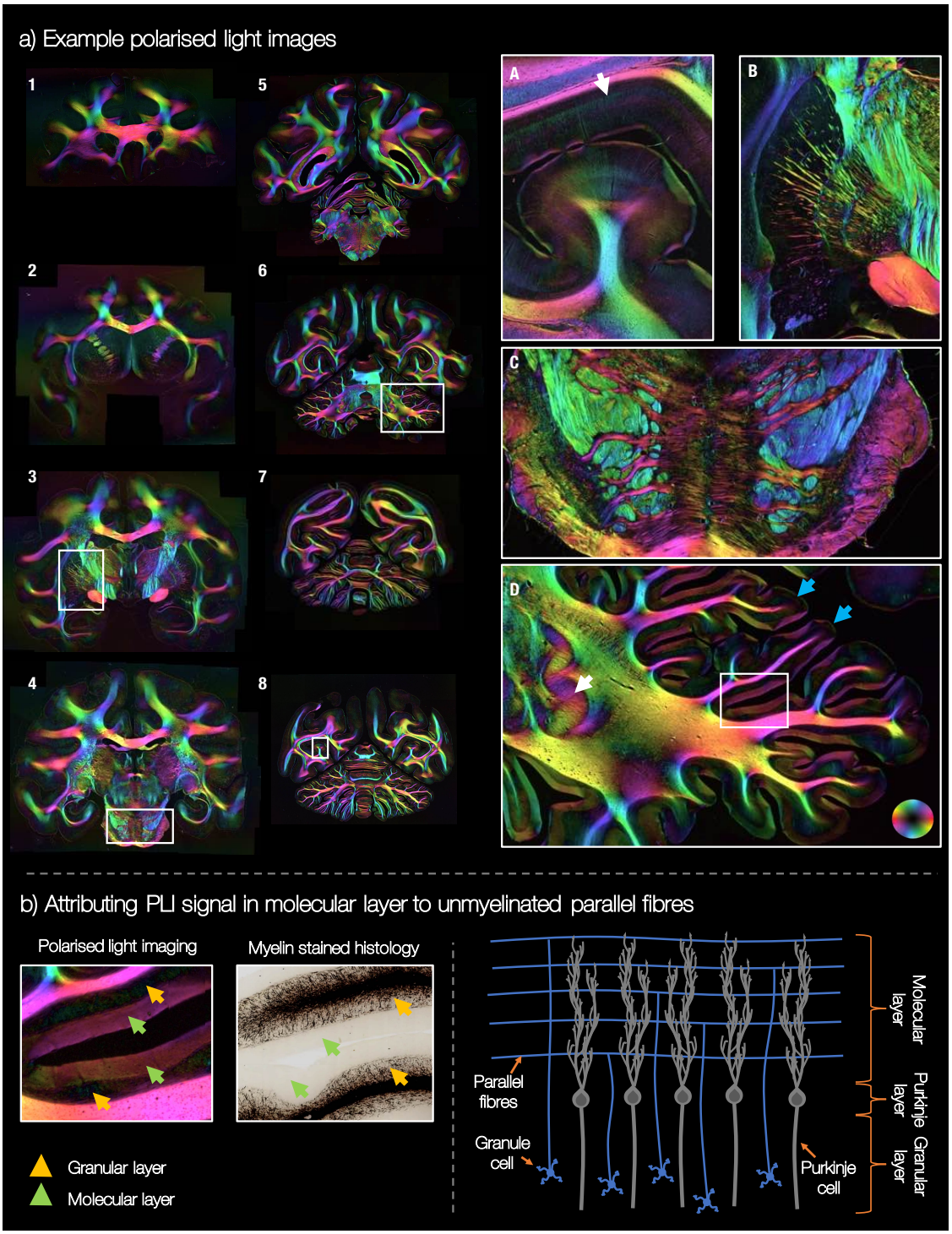


Figure 3: Polarised light imaging in BigMac. a) Example PLI throughout the brain. Image 1 = most anterior, 8 = most posterior of the images shown. The myeloarchitecture is viewed in great detail due to the 4 µm PLI resolution. We observe fibres projecting into the cortex (A), into and through subcortical structures (B), around the pons (C) and across the cerebellum (D). In D, the blue arrows point to the gyri crown. b) We hypothesise that the PLI signal in the cerebellar molecular layer can be attributed to parallel fibres. b, left) The yellow and green arrows point to the granular layer and the molecular layer respectively. b, right) The structure of the cerebellar cortex. This highly simplified schematic focuses solely on the granule and Purkinje cells, to illustrate the parallel fibres in the molecular layer. Note, the dendritic tree of the Purkinje cells has highly anisotropic dispersion. Here we see the axis of least dispersion, where the Purkinje dendritic tree fans out most in the through-page orientation.

two horizontal branches in a 'T-like' fashion. These branches are known as 'parallel fibres'. The Purkinje cells sit with their soma in the 'Purkinje layer', at the interface of the granular and molecular layer, and the Purkinje dendritic tree extends into the molecular layer. The parallel fibres of the granule cells run through the Purkinje dendrites, forming synaptic connections with the Purkinje dendritic spine. Consequently, it seems reasonable that these parallel fibres, running in line with the cortical surface, could cause the coherently oriented birefringence of the molecular layer. Similar conclusions were reached by Koike-Tani et al. [28] who attributed high birefringence in the molecular layer of a late stage chick embryonic cerebellum to the presence of densely packed, non-myelinated parallel fibres.

Although PLI signal from brain tissue is typically associated with the myelin, histological data did not support the presence of myelinated fibres in the molecular layer (Figure 3b left). Furthermore, non-myelinated fibres often exhibit positive birefringence due to the presence of oriented proteins within the thin myelin sheath, rather than negative birefringence associated with the lipid bilayer [29]. Such positive birefringence would result in PLI orientations tangential to those observed. Here there could be two effects: i) that the PLI signal is sensitive to not only myelin, but also other aspects of microstructure orientational coherence [30], or ii) that the parallel fibres have some small amount of myelin which is detected by PLI but not picked up by the histology stain. Indeed, though parallel fibres are often assumed to be exclusively unmyelinated [31], using electron microscopy Wyatt *et al.* [32] found myelinated fibres of 0.4-1.1 µm diameter in the Macaque molecular layer. Interestingly, Wyatt *et al.* found a larger number of myelinated fibres in the molecular layer near the Purkinje layer, and fewer towards the cortical surface. Consequently, we might expect to see a gradient in the PLI signal across the molecular layer, though this is not evident in the data. Future work is required to fully understand the origin of this birefringence in the molecular layer, though Figure 3 provides some evidence for PLI sensitivity to coherently orientated, anisotropic structures irrespective of their degree of myelination [28, 33], that may be omitted from classic myelin histology.

2.3.2. Myelin-stained histology

Figure 4a shows example images from myelin-stained histology (Gallyas silver, [15]) included in the anterior BigMac brain. We see interesting detail in both the white and grey matter where we can visualise single, thin axons in detail due to both the sensitivity of the stain and the sub-micrometre imaging resolution (0.28 µm / pixel). In the digitised images we can track complex patterns of fibre projection for example, from the white matter into the cortex (A), between sub-cortical structures (B), within the highly complex geometry of the hippocampus (C) or through the deep white matter (D). Furthermore, in the deep white matter we see different tissue 'textures' as well as large scale, 'wave-like' undulations of fibres in the corpus callosum (D). Though this a region which is often considered coherently ordered with little fibre dispersion, these data corroborate previous observations of fibre dispersion or incoherence in the corpus callosum [4].

The myelin-stained slides were analysed using structure tensor analysis [34, 35, 36, 37] to estimate the primary fibre orientation per microscopy pixel (Figure 4b). Figure 4c shows the primary fibre orientation derived from the structure tensor analysis per $\sim$ 40 µm superpixel. The orientations are colour-coded in HSV (hue-saturation-value) space similar to PLI. Here, pixels in heavily stained white matter are bright compared to those in the lightly stained grey matter as the "value" was set to $(1 - r)$, where r is the grey-scale pixel intensity. This image is then compared to an adjacent section imaged with PLI (Figure 4d). Despite the very different manner by which the orientation estimates were derived, the two methods provide corroborating information in both the white matter and the cortex, where myelinated fibre are less dense. Furthermore, we see how the two modalities provide subtly different information. For example, when a PLI pixel (at 4 µm/pix resolution) contains crossing fibre populations of approximately equal weighting and perpendicular orientations, the PLI-derived in-plane angle becomes uninformative, the PLI signal is low and the HSV image dark. In Figure 4e (right) we see how this results in a darker bands appearing between fibre tracts of different orientations. For example, where the pink and teal fibres intermingle, as indicated by the white arrows. In these crossing fibre regions, the histology data has an order of magnitude finer spatial resolution (at 0.28 µm/pixel) and so can likely resolve both fibre populations, the mean of which is shown

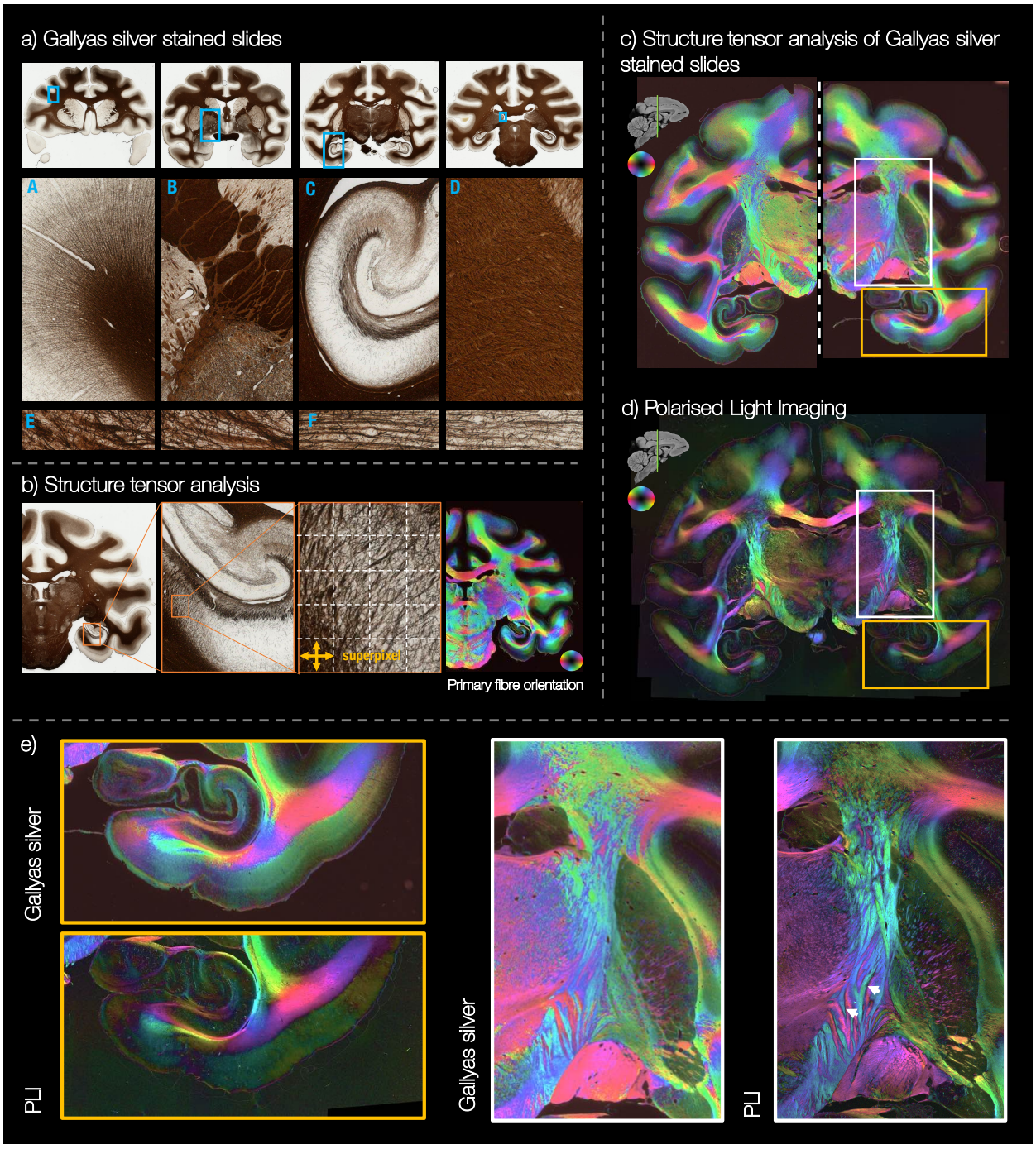


Figure 4: Gallyas silver stained histology in BigMac. a) Example digitised slides. With a spatial resolution of 0.28 µm / pixel, we see the myelinated fibres is great detail, visualising single axons at the grey/white matter boundary (A-F) and fibre undulations in the deep white matter (D). b) Structure tensor analysis was applied to the Gallyas silver stained slides to estimate a fibre orientation per 0.28 µm pixel. c,d) The fibre orientations derived from the Gallyas silver stained slide are compared to an adjacent PLI slide. We see remarkable consistency between the images with both modalities capturing the myeloarchitecture in detail. e) Both modalities show the detailed organisation of the hippocampus and surrounding white/grey matter (yellow box), as well as the corticospinal tract (white box).

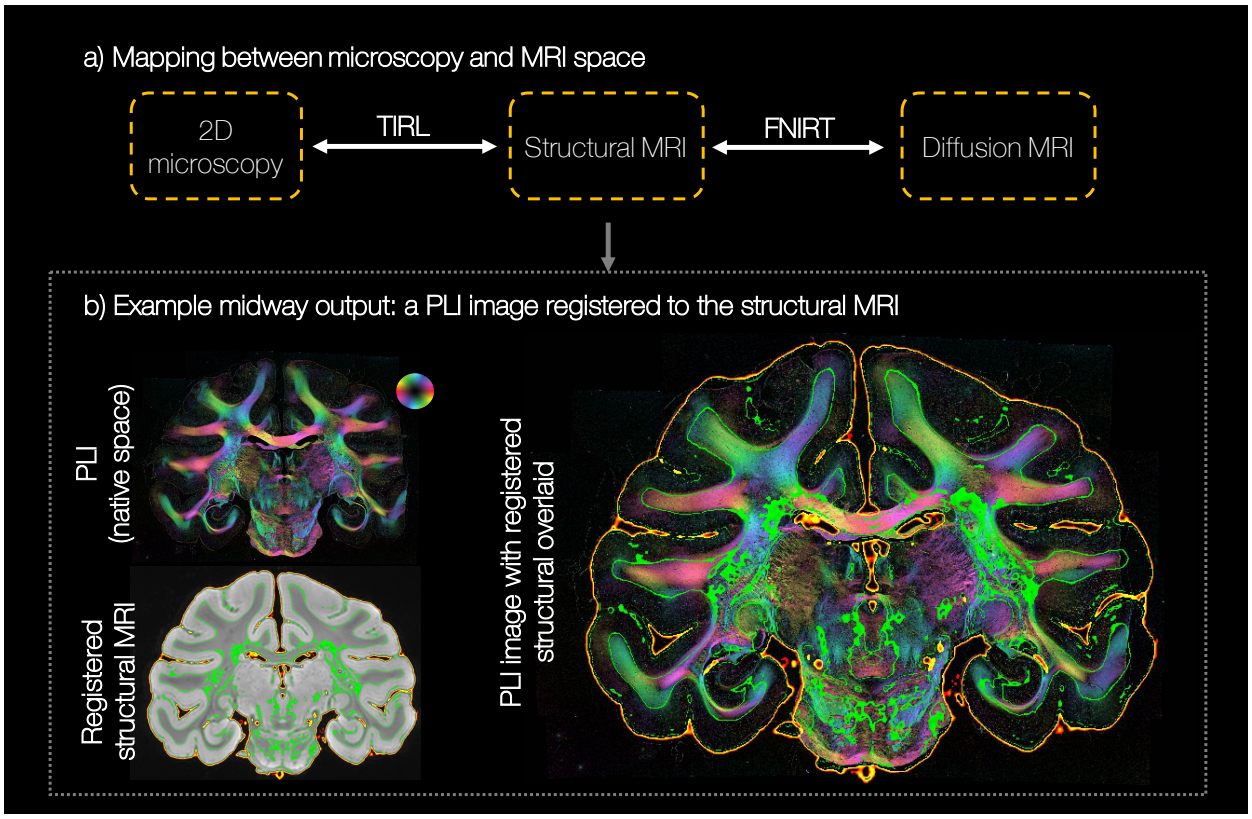


Figure 5: Co-registration of the BigMac PLI and MRI data. a) The PLI and diffusion MRI data were independently registered to the structural MRI, using multi-scale MRI-histo registration (TIRL) and more typical MRI-MRI non-linear registration (FNIRT) respectively. By combining the two warpfields, the PLI data can be directly transformed into diffusion MRI space or vica versa. b) An example output of TIRL. The structural image is clipped to indicate the tissue outline (orange) and white/grey matter boundary (green). These outlines are then overlaid on the PLI image (right).

in Figure 4b,c.

2.4. Co-registration of MRI and microscopy data

Our ability to meaningfully compare MRI and microscopy data is greatly enhanced by having high quality multimodal data registration, like that provided in BigMac. However, co-registration of the BigMac MRI and microscopy data was highly challenging for a number of reasons. Firstly, the spatial resolution of the data spans up to 4 orders of scale, from microscopy images with sub-micrometre resolution to the ∼millimetre MRI data. Secondly, the contrast between the MRI-microscopy images is substantially different and may highlight different tissue features. For example, in both the Gallyas and PLI images, neither the grey/white matter boundary nor outer tissue edges are consistently well defined. Thirdly, the thin tissue sections may be deformed during sectioning or microscopy preprocessing (e.g. staining and/or mounting). For example, the tissue can shrink or tear, or there may be dirt or bubbles in the microscopy slides. Finally, the inherently 2D microscopy must be registered into the 3D volume - a particularly difficult optimisation with many degrees of freedom.

Figure 5 shows an example BigMac registration using TIRL, a new MRI-microscopy registration tool by Huszar *et al.* [19, 20] that is specifically designed to overcome the above challenges and facilitate accurate MRI-microscopy co-registration. In BigMac, TIRL generated the mapping between microscopy and the

structural MRI, and FSL tools (FLIRT/FNIRT [38, 39]) were used for cross-modality registration within MRI (e.g. co-registering structural and diffusion MRI). TIRL utilises a sequence of linear and non-linear transforms to register 2D microscopy images into 3D MRI volumes. Further, the TIRL cost function is based on a modality independent neighbourhood descriptor [40], which is explicitly designed to capture correspondence in spatial information in a way that is agnostic to image contrast. By concatenating the TIRL and FSL transforms, the TIRL platform could then be used to map the high-resolution microscopy pixels into the 3D MR volume in any MR domain, or MR voxels onto a 2D microscopy plane. Figure 5b shows the quality of the TIRL transform where the MRI data maps very closely to the microscopy image, both at the white/grey matter boundaries (outlined in green) and the tissue edge (in orange).

2.5. Comparing fibre orientation distributions from microscopy and MRI

With co-registered MRI and microscopy data, BigMac facilitates voxelwise comparisons of quantitative microstructural metrics extracted from both modalities. Figure 6a shows example co-registered maps from MRI and microscopy. The primary axis of the diffusion tensor (V_1) [41] has been projected onto the microscopy plane and uses the same 2D colour map as the PLI. We observe excellent agreement between V_1 and the orientations derived from polarised light imaging. This facilitates direct validation of orientation information from MRI against microscopy, as illustrated in Figure 6b. Here, the diffusion signal has been modelled using a biophysical model, the Ball and Rackets model (BAR) [42], to extract a single, disperse fibre orientation distribution per voxel, which has then been projected onto the microscopy plane. This is compared to co-registered fibre orientation distributions from both PLI and Gallyas-stained histology. We observe good correspondence of the dominant fibre orientations in both coherent regions such as the corpus callosum and more complex relationships in crossing fibre regions such as the centrum semiovale. Here both orientational information, and the amount of fibre dispersion can be compared. Indeed, the slightly higher levels of dispersion away from the midline of the corpus callosum may be related to the 'wave like' fibre patterns observed in the myelin-stained histology (Figure 4 part d).

The fibre dispersion can then be quantified using the orientation dispersion index, which ranges from 0 for perfectly aligned fibres, to 1 for isotropic dispersion [43]. Figure 6c compares dispersion estimates across many white matter voxels (top, corresponding to the yellow mask), and a subset of voxels from the centrum semiovale (bottom, pink mask). We see good correspondence between the dispersion from myelin-stained histology and the diffusion model. Estimates of dispersion from PLI appear less reliable, in line with previous reports [4] and observations in 6b.

As well as comparing MRI-microscopy equivalents, microscopy can be used to understand indirect relationships with MR parameters. For example, histology dispersion is shown to have a clear negative correlation with fractional anisotropy (FA) from DTI (Figure 6d, [41]). In comparison, dispersion has a weaker correlation with microscopic FA (µFA) [44, 45] in the white matter (top) and almost no correlation in the centrum semioval (bottom), a known deep white matter region of complex dispersion. This is reassuring as the µFA parameter is explicitly meant to be independent of the fibre orientation distribution: in the centrum semiovale, µFA is approximately independent of dispersion, where the negative correlation across all of white matter is likely driven by partial volume effects. In future work when additional microscopy contrasts are added to BigMac, multivariate regressions can be performed to better understand how complex tissue microstructure relates to sensitive but not specific diffusion metrics such as those from the diffusion tensor or other signal models [46], which typically lack microstructural interpretation.

2.6. Towards the microscopy connectome: hybrid MRI-microscopy tractography

One of the primary limitations of the microscopy data in BigMac is that it only informs on the fibre orientations in the 2D plane of sampled slides, precluding 3D reconstruction of the microscopy connectome. In comparison, the diffusion data can provide orientational information in 3D, but with limited spatial

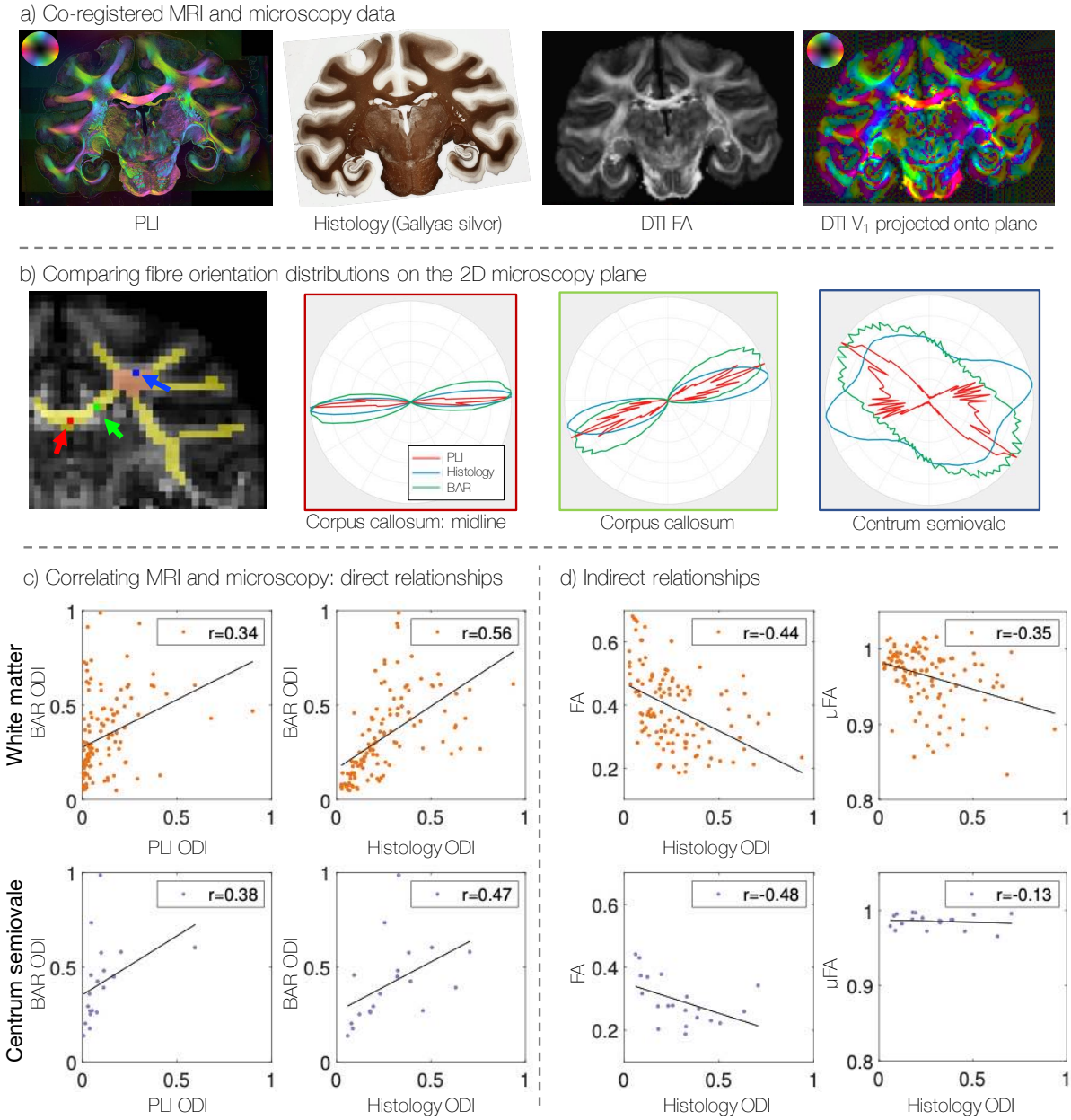


Figure 6: Comparing information from co-registered MRI and microcopy. a) Co-registration facilitates qualitative comparisons of microscopy and MRI (here DTI data) when both are warped to a common space. We see clear correspondence between the PLI orientations and the primary eigenvector of the diffusion tensor that has ben projected onto the microscopy plane. b) A 2D fibre orientation distribution is extracted on a voxelwise basis from both PLI, histology, and the Ball and Rackets model (BAR) for diffusion MRI. c) The orientation dispersion index (ODI) from microscopy is correlated with various diffusion metrics: the ODI from the Ball and Rackets model, fractional anisotropy (FA) from the diffusion tensor model, and µFA from data with multiple tensor encodings. The top row shows all data points from the white matter mask shown in yellow in part b. The bottom row shows only a subset of voxels from the centrum semiovale, shown in pink. The black lines show the line of best fit, and r is the correlation coefficient.

resolution (0.6-1mm isotropic). Figure 7 demonstrates one approach to joint modelling [10] where we combine in-plane orientations from microscopy with through-plane information from diffusion MRI (dMRI) to reconstruct 3D hybrid dMRI-microscopy fibre orientations at the resolution of the microscopy data. For each microscopy pixel, we compare the microscopy-derived in-plane orientation to orientations estimated from co-registered diffusion data using the Ball and Stick model [21] which have been projected onto the microscopy plane (Figure 7a). The microscopy through-plane angle is then approximated by that from the most similar BAS orientation. This produces a hybrid diffusion MRI-microscopy orientation that is both 3D and at the resolution of the microscopy data. These orientations can then be combined into a hybrid fibre orientation distribution (FOD) which can be directly compared with those from diffusion MRI and input into existing tractography algorithms for tract reconstruction. Here we reconstruct hybrid orientations using the PLI data in BigMac, though our current method is also applicable to histology slides analysed using structure tensor analysis.

Figure 7b shows example MRI-microscopy FODs at varying spatial resolutions. Reassuringly, the hybrid FODs show smoothly varying patterns in all three dimensions, even when the hybrid FODs are of higher spatial resolution than the diffusion MRI (0.6 hybrid reconstruction versus 1 mm diffusion data). Interestingly we observe notably fewer voxels with crossing fibre populations than we might expect from diffusion MRI in regions such as the centrum semiovale. The hybrid FODs can then be reconstructed at very high in-plane resolutions ($\gtrsim 4 \times 4$ µm in-plane, $\gtrsim 350$ µm through-plane) to observe fine structural details, or investigate the effect of spatial resolution on FOD reconstruction or downstream tractography (the subject of future work). As proof of concept, the corticospinal tract was reconstructed from the hybrid FODs using anatomically constrained tractography (Figure 7c). Future work will consider whole brain reconstruction of the 3D microscopy connectome at high spatial resolution. We expect this microscopy-inspired connectome to both provide new anatomical insight, and be a valuable resource for validating and advancing in vivo tractography [9].

3. Conclusion

The BigMac dataset aims to characterise a single connectome in exquisite detail, combining MR signals with high resolution microscopy data throughout the macaque brain. The postmortem diffusion MRI includes high b-value ultra-HARDI data to characterise the diffusion signal in great detail, estimate white matter fibre orientations with high precision, and improve the reconstruction of white matter tracts through crossing fibre regions. Furthermore, with 1000 gradient directions we retain dense sampling on an arbitrary 2D plane for direct comparison with 2D microscopy data. The microscopy includes coronal slides of polarised light imaging [12, 13, 14] and myelin-stained histology that has been densely sampled throughout the brain. This high-resolution, high specificity data allows us to visualise tissue myeloarchitecture in detail to provide neuroanatomical insight - such as the orientational coherence of the PLI signal in the cerebellar molecular layer - and act as a pseudo ground truth estimate of tissue microstructure against which MR metrics can be compared. Crucially, the MRI and microscopy data have been carefully co-registered, facilitating novel data fusion analyses including the hybrid diffusion MRI-microscopy tractography presented here.

The dataset has several limitations including: the 2D nature of the microscopy, which limits 3D reconstruction of the tissue microstructure at the micrometre level; tissue processing artefacts (*c.f. Section 5.4.1*) and inconsistent staining (*c.f. Section 5.4.2*) in some PLI and histology slides respectively; and the choice of MR data acquired where here we chose to comprehensively sample q-space in the long diffusion time regime, rather than, for example, acquire data at multiple diffusion times for sensitivity to restricted compartments or acquire MRI at very high spatial resolution. The latter could be overcome by acquiring additional MRI on a different monkey brain that was subsequently co-registered to the BigMac data, though such analyses would have to assume some level of anatomical consistency between brains.

Nonetheless, as a unique, multimodal resource that complements existing open data, the BigMac dataset

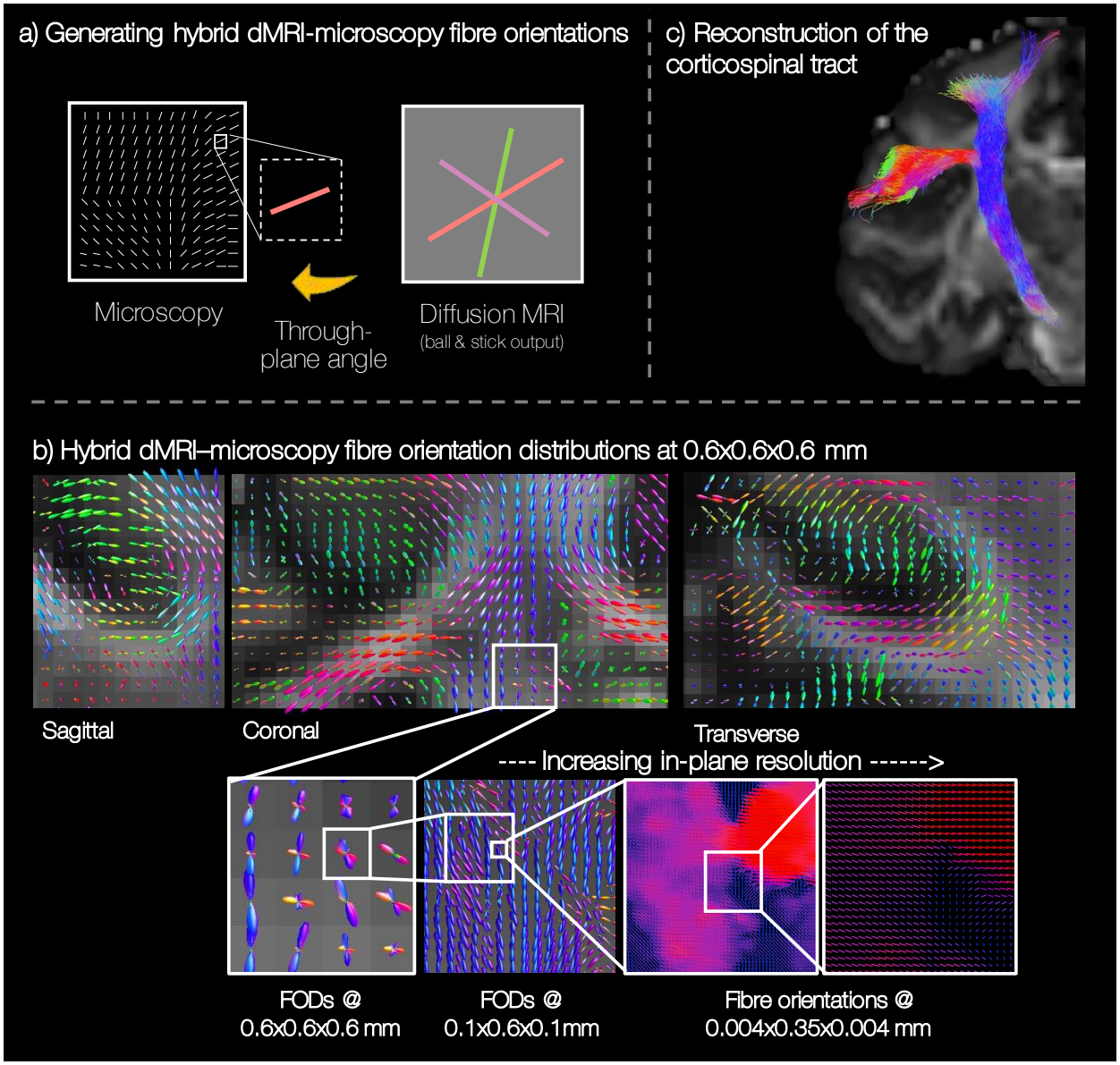


Figure 7: Estimating hybrid diffusion MRI-microscopy fibre orientations for future reconstruction of the microscopy-inspired connectome (see text for details). a) Where microscopy informs on fibre orientations within the microscopy plane, the through plane is approximated by that from diffusion MRI (Ball and Stick model [21]). b) The hybrid orientations can be combined into fibre orientation distributions at increasing in-plane resolutions. c) The hybrid FODs can be input into pre-existing tractography methods to to successfully reconstruct in white matter bundles such as the corticospinal tract.

will enable neuroscientists to ask new and fundamental questions. For example, it will enable researchers to i) link brain connectivity across spatial scales spanning four orders of magnitude; ii) understand the relationship between relatively crude but non-invasive MRI signals and cellular properties measured by invasive microscopy; and iii) develop novel approaches to extract rich information from in vivo MRI data by leveraging the specificity of microscopy-derived gold standards. The BigMac dataset provides an open access platform from which we can interconnect microstructural features with MRI signals throughout the brain.

Future work will include characterising the unique information provided by the hybrid MRI-microscopy tractography, as well as completing ongoing microscopy work to add Nissl-stained histology and other complementary stains (e.g neurofilament protein SMI32 or glia Iba-1 or GFAP) to the BigMac resource.

4. Data availability

The BigMac data (including minimally preprocessed data) are openly available via the Digital Brain Bank [6]: https://open.win.ox.ac.uk/DigitalBrainBank/#/datasets/anatomist. Example data can be accessed via the online viewer and the full dataset is available via a material transfer agreement to ensure the data is used for purposes which satisfy research ethics and funding requirements. Further documentation (including example images of all modalities) is available via https://open.win.ox.ac.uk/pages/amyh/bigmacdocumentation/index.html where we also provide basic MRI-microscopy tutorials and scripts relating to the analyses in this paper. Since the full dataset requires a considerable amount of memory ($\sim 1.8TB$), users may wish to request only a subset of relevant data. To facilitate this, the documentation includes an extensive file tree, listing available files as well as approximate memory requirements for different parts of the data. Additional data linking behaviour to fMRI (z-statistics maps) can be accessed via [17, 18] and similar in vivo MRI (structural and resting-state fMRI) are available for another 19 subjects for corss-subject comparisons via the primate data exchange (PRIME-DE): https://fcon_1000.projects.nitrc.org/indi/PRIME/oxford.html.

5. Materials and methods

5.1. Data acquisition

At the centre of the BigMac dataset is the brain of a single adult rhesus macaque (*Macaca mulatta*, male). During it's adult life, the macaque was scanned in vivo over multiple scan sessions. At 11.7 years of age, the brain was perfusion fixed in formalin after which extensive postmortem MRI data was acquired. After scanning, the entire brain was sectioned along the anterior-posterior axis, with consecutive slices processed for different microscopy contrasts.

The animal was cared for, and data were acquired by researchers at the University of Oxford, UK. All procedures were performed under licenses from the United Kingdom (UK) Home Office in accordance with the UK Animals (Scientific Procedures) Act 1986 and with European Union guidelines (EU Directive 2010/63/EU).

5.1.1. Tissue pathology

As part of a behavioural study (in preparation), the BigMac monkey underwent intraoperative bilateral lesioning of the orbitofrontal cortex. The postmortem MRI data in Supplementary Figure 1 shows the extent of the bilateral lesions approximately one year after surgery.

In addition to the planned lesion, inspection of the postmortem data (Supplementary Figure 2) shows a fairly substantial abnormality in the left hemisphere which extends from the inferior portion of the supramarginal

gyrus up through the post-central sulcus. This abnormality could relate to a cerebral bleed, which perhaps occurred post operatively, though no behavioural or other observations were made that would relate to this abnormality.

5.2. In vivo MRI

Prior to sacrifice, the animal partook in a number of studies [17, 18] in which behavioural and imaging data were acquired. One study [17] combines functional MRI with a decision-making task to investigate the of role surprising events (i.e. prediction errors) on reward-based learning. A second [18] links flexible behaviour to changes in both the MRI-derived structure and function of a fronto-cortical network.

The BigMac in vivo MRI data includes structural images, diffusion MRI, resting-state fMRI and task fMRI over a variety of tasks. The data were acquired at various time points throughout the animal's adult life. The in vivo data were acquired on a 3 T whole-body scanner (Gmax = 40 G/cm) with a four-channel phased-array receive coil and a local transmit coil (Windmiller Kolster Scientific). Here we include data acquired at two separate time points. "Session 1" includes diffusion, structural and resting-state fMRI, where complementary resting-state and structural data from another 19 animals has previously been made openly available through the PRIME-DE database for cross-subject comparisons. "Session 2" includes similar MRI from a shorter acquisition that occurred only 1 year before sacrifice (the last in vivo scan). As such, the age-induced atrophy between the in vivo and postmortem data should be roughly similar. During scanning the animal was kept under minimum anaesthetic using similar procedures to those previously described [47, 48, 49].

Structural MRI images were acquired using a T_1-weighted Magnetization Prepared - RApid Gradient Echo (MP-RAGE) sequence with 0.5 mm isotropic resolution, TE/TR = 4.01 ms/2.5 s and 128 slices. Whole brain fMRI data (BOLD) were acquired with echo planar imaging (EPI) and 2 mm isotropic resolution: TE/TR = 19 ms/2 s, 1600 volumes for Session 1 and 800 volumes for Session 2. This corresponds to 52 min 26 s and 26 min 13 s of data respectively. Diffusion MRI data were acquired using EPI with 1 mm isotropic resolution, TE/TR = 100 ms/8.2 s and a b-value of 1 ms/µm^2. 1100 diffusion weighted (81 unique gradient directions) and 144 non-diffusion weighted volumes were acquired with both ± phase encoding directions for Session 1, and 361 diffusion weighted (61 unique gradient directions) and 38 non-diffusion weighted volumes were acquired with both ± phase encoding directions for Session 2. The data were distortion corrected using bespoke pipelines and FSL tools [50, 51]. Additional task fMRI maps (z-stats) are available via [17, 18] .

To indicate the quality of the in vivo and postmortem data, Supplementary Figure 3 (top) shows example structural images from the most recent in vivo imaging session. This is then compared to postmortem data, as described below.

5.3. Postmortem MRI

At 11.8 years of age the animal was anaesthetised and the brain perfusion fixed with 90 % saline and 10 % formalin, extracted and then stored in 30% sucrose formalin. Postmortem data were then acquired on an 7 T small animal scanner (Agilent) fitted with a 40 G/cm gradient coil (Agilent, 205/120 mm) and a Birdcage receive/transmit RF coil (Rapid Biomedical, 72 mm). Prior to scanning the brain was rehydrated in phosphate-buffered saline to remove lasting fixative and somewhat restore both the diffusivity and T_2 of the tissue [52, 53]. The brain was then packed into a plastic holder filled with Fluorinert (FC-3283, 3 M, St. Paul, USA), a proton-free, susceptibility-matched fluid which is MR invisible and improves field homogeneity.

As the diffusion properties of brain tissue are highly dependent on the tissue temperature [54], the temperature was controlled by passing air at the constant temperature of 201°C.

The BigMac postmortem MRI data was acquired over three different scanning sessions:

1. 29th March - 5th April 2018: Acquisition of $b = 7$ and $10\,\mathrm{ms/\mu m^2}$ ultra-HARDI data with 1000 diffusion-weighted gradient directions per shell and 1 mm isotropic resolution.

2. 6th - 9th April 2018: Acquisition of the 0.3 mm structural MRI, T_1 mapping and 0.6 mm $b = 4\,\mathrm{ms/\mu m^2}$ diffusion-weighted data.

3. 20th - 23rd April 2018: Acquisition of 1 mm $b = 4\,\mathrm{ms/\mu m^2}$ diffusion-weighted data plus the protocol combining linear and spherical tensor encoding at $b = 4$, 7 and $10\,\mathrm{ms/\mu m^2}$.

The brain was not repacked between sessions, though slight deformations did occur as the tissue relaxed over time. In total, the postmortem MRI data acquisition took ~ 270 hours scanning time. All postmortem data were corrected in 3D for Gibbs ringing artefacts (mrdegibbs3D, MRtrix [55, 56, 57]) prior to other preprocessing.

5.3.1. Structural MRI

Two structural images were acquired with subtly different contrast: one a with multi gradient echo (MGE 3D) sequence, and one using balanced steady-state free procession (bSSFP).

The MGE parameters were: TE/TR = 7.8/97.7 ms, flip angle = 30°, 0.3 mm isotropic resolution, FOV = 76.8 x 76.8 x 76.8 mm. The structural image was subsequently corrected for bias field and segmented using FAST [58, 50, 51]. The white and grey matter masks were then hand edited to provide precise segmentation of the white and grey matter.

The bSSFP data were acquired using a TRUFI sequence with 16 frequency increments: TE/TR = 3.05/6.1 ms, flip angle = 30°, 0.3 mm isotropic resolution, FOV = 76.8 x 76.8 x 76.8 mm. The structural image was formed by averaging the data using root-mean sum of squares.

The Supplementary Figure 3 (bottom) shows example postmortem structural images from the BigMac dataset. Note how the contrast is inverted when related to the in vivo T1-weighted images. Here we purposefully acquire T2/T2*-weighted postmortem data as conventional T1w typically don't give good contrast postmortem due changes in relaxation times. Due to their high image quality and anatomical detail, the postmortem structural MRI act as a crucial intermediary in the co-registration of both the diffusion MRI and microscopy data (*c.f. Co-registration*) and the in vivo and postmortem MRI.

5.3.2. T_1 mapping

A T_1 map was acquired using similar imaging parameters as the 0.6 mm postmortem diffusion MRI data but now with an inversion recovery preparation: TE/TR = 8.6 ms/10 s, FOV = 76.8 x 76.8 x 76.8mm, resolution 0.6 mm isotropic and 12 inversion times (TI) from 10 to 6000 ms. The Barral model $S(\mathrm{TI}) = a + b\ \exp(-\mathrm{TI}/T_1)$ [59], where S is the MR signal and $[a, b, T_1]$ are unknowns, was fitted voxelwise to the data to obtain quantitative estimates of T_1 (inversion_recovery, qMRLab [60]).

5.3.3. Diffusion-weighted MRI

The diffusion-weighted data were acquired using a spin echo multi-slice (DW-SEMS) sequence and single-line readout. To ensure that data from different shells retain the same diffusion propagator, both the time between the gradients (i.e. the diffusion time, Δ) and gradient duration (δ) were kept constant for all data

with 1 mm isotropic resolution. The desired b-value was achieved by modifying the amplitude of magnetic gradient, G.

High spatial or angular resolution

In the 1 mm diffusion data with high angular resolution, data were acquired in batches of 26 volumes where one volume with negligible diffusion weighting ($b \sim 0\,\mathrm{ms/\mu m^2}$) was followed by 25 diffusion-weighted volumes. Two sets of gradient directions were used: one with 250 gradient directions ($b = 4\,\mathrm{ms/\mu m^2}$), the other with 1000 gradient directions ($b = 7, 10\,\mathrm{ms/\mu m^2}$). For both sets, the gradient directions were generated using GPS (an FSL tool, [61]) and were evenly distributed across the sphere. The directions were then ordered so that any consecutive subset of gradient directions (e.g the first 100 gradient directions) also gave good coverage across the sphere (orderpoints, Camino [62]). In this case, were the scan interrupted or prematurely stopped, we would retain reasonable angular coverage. Finally, evenly spread the heating of the magnetic gradients, the gradient directions within each batch of 25 were reordered to ensure that highly co-linear directions were not played out in close succession.

The 1mm data acquisition parameters were as follows: TE/TR = 42.4 ms/3.5 s; FOV=76 x 76 x 76 mm; $\delta/\Delta = 14/24$ ms; 1 mm isotropic resolution; time per gradient direction = 4.4 mins; $b = 4\,\mathrm{ms/\mu m^2}$ data had G = 12.0 G/cm, 250 gradient directions and 10 non-diffusion weighted volumes; $b = 7\,\mathrm{ms/\mu m^2}$ had G = 15.9 G/cm, 1000 gradient directions and 40 non-diffusion weighted volumes; $b = 10\,\mathrm{ms/\mu m^2}$ had G = 19.1 G/cm, 1000 gradient directions and 40 non-diffusion weighted volumes.

The 0.6 mm $b = 4\,\mathrm{ms/\mu m^2}$ data followed a different protocol. Here 128 diffusion-weighted gradient directions were acquired, followed by 8 volumes with negligible diffusion weighting. The acquisition parameters were as follows: TE/TR = 25.4 ms/10 s; FOV=76.8 x 76.8 x 76.8 mm; $\delta/\Delta = 7/13$ ms; time per gradient direction = 21.3 mins; $b = 4\,\mathrm{ms/\mu m^2}$; 0.6 mm isotropic resolution; G = 32 G/cm.

Preprocessing

The postmortem MRI data was found to have few distortions, so minimal preprocessing was applied. For example, the data did not need correcting for susceptibility or eddy current distortions. This is largely due to the brain being placed in a susceptibility-matched fluid and the data acquired with a single-line readout instead of the typical echo planar imaging (EPI). The main corrections were a) registration (both within and between session), b) correction of signal drift, and c) signal normalisation.

Registration

The ultra-HARDI data for both $b = 7$, and $10\,\mathrm{ms/\mu m^2}$ were acquired within the first scanning session. At specific time points throughout the week-long acquisition, the central scanner frequency was recalibrated. This occurred three times during the $b = 7\,\mathrm{ms/\mu m^2}$ acquisition and 4 times during $b = 10\,\mathrm{ms/\mu m^2}$. Because of the recalibration, images acquired with different scanner central frequencies are shifted (translated) with respect to one another. To correct for these translations, the data were rigidly registered to a reference S_0 image (i.e. a volume with negligible diffusion weighting) from the ultra-HARDI dataset. Here the reference image was taken to be the mean S_0 image from the first 'set' of images which were all acquired with the same central frequency. The registration was performed using FLIRT with spline interpolation of the data [63, 38].

Data from the second scan session includes high spatial resolution (0.6 mm isotropic) $b = 4\,\mathrm{ms/\mu m^2}$ diffusion data as well as the detailed (0.3 mm isotropic) structural scan. Upon inspection, the S_0 images associated with the 0.6 mm diffusion data ($b = 4\,\mathrm{ms/\mu m^2}$) appeared to slowly drift in position along the readout direction. To correct for signal drift, the S_0 images were linearly registered and intensity normalised to the first S_0 i.e. that which most likely represents the 'true' S_0 of the diffusion-weighted data. The data were aligned using FLIRT [63, 38] where the transformation was restricted to only consider translation along a single axis. The $b = 4\,\mathrm{ms/\mu m^2}$ 0.6 mm diffusion-weighted data were then co-registered to the postmortem structural image using linear registration (FLIRT, [63, 38]).

Supplementary Figure 4 describes how data acquired in different sessions was registered together using either linear or non-linear transforms (FLIRT/FNIRT) [63, 38, 39, 61]. Data acquired within the same scan session were registered using linear transforms. Upon inspection of the data, the brain shape appeared to change or 'relax' slightly between scanning sessions. To account for these deformations, non-linear transformations were generated both between the $b = 4\,\mathrm{ms/\mu m^2}$ 1 mm data and the ultra-HARDI data, and between the ultra-HARDI data and the postmortem structural image [39, 61]. Consequently, data users should take care to account for voxelwise rotations in the gradient directions according to the non-linear warpfield when combining non-linearly registered diffusion data from different shells.

Finally, to integrate the BigMac dataset with other datasets, non-linear transformations[39, 61] were computed between the postmortem structural image and the F99 standard template [64]. Here we utilise a T_1-like image, created from the structural MRI using hand-edited white and grey matter masks, because non-linear registration requires images with similar contrast and the BigMac ex vivo structural image has inverted contrast when compared to the in vivo F99 T_1.

Signal drift
In all experiments, the signal magnitude, measured as the mean signal across S_0 images, was seen to decrease over time. To correct for signal drift, a linear trend with respect to time was fitted to the S_0 images and subsequently regressed from the data (both the S_0 and diffusion-weighted volumes).

Data normalisation
Most diffusion models approximate the S_0 image by taking the mean S_0 image across all volumes with minimal diffusion weighting, assuming that the signal magnitude is constant across time. In contrast, here we found the S_0 signal to vary between scanning sessions and b-shells. Were diffusion models naively applied to concatenated data from the BigMac dataset, the results may be biased (e.g. the kurtosis would be misestimated). Consequently, the diffusion-weighted data was normalised to the mean S_0 of the $b = 10\,\mathrm{ms/\mu m^2}$ ultra-HARDI data.

Combining linear and spherical tensor encoding
Combining data with linear and spherical tensor encoding allows for the separation of effects due to the fibre orientation distribution and the diffusion properties, to estimate additional microstructural parameters such as µFA [45, 44]. In BigMac, data with spherical tensor encoding were also acquired at b-values of $4, 7$ and $10\,\mathrm{ms/\mu m^2}$. The gradient waveform was optimised using the NOW toolbox in Matlab [65]. Due to additional stress on the magnetic gradients when performing the spherical tensor encoding, the repetition time (TR) was increased with respect to the ultra-HARDI data: TE/TR = 42.5 ms/6.4 s; FOV = 76x76x76 mm; 1 mm isotropic resolution. For each b-value, 30 images were acquired with spherical tensor encoding and 1 with negligible diffusion weighting. Complementary data with linear tensor encoding and the same TR were also acquired: 50 gradient directions per shell with $\delta/\Delta = 14/24$ ms, plus 2 volumes with negligible diffusion weighting. The gradient amplitude G was adjusted to produced the required b-values of $b = 4,\ 7$ and $10\,\mathrm{ms/\mu m^2}$.

Data were corrected for Gibbs ringing and signal drift as above. All data were normalised to the linear tensor encoded $b = 10\,\mathrm{ms/\mu m^2}$ mean S_0. Maps of µFA as well as isotropic and anisotropic kurtosis were generated following the DIVIDE framework (by fitting the Laplace transform of the gamma distribution, *dtd_gamma* model) [44, 66, 67].

5.4. Microscopy

Using the BigMac dataset, we can link the MRI signal to microscopy data which has both micrometre resolution and high specificity. In BigMac, the brain was sectioned, stained, and imaged ('processed') in two batches. The brain was first cut around the level of the posterior tip of the central sulcus to create two tissue blocks, representing the anterior half and posterior half. First, the anterior block was sectioned, stained, and imaged, after which the posterior block was processed using a highly similar protocol.

Each tissue block was sectioned on a frozen microtome along the anterior-posterior axis to produce thin coronal tissue sections. Consecutive sections were allocated, in order, to one of six contrasts:

1. Polarised light imaging to visualise myelinated fibres (50 µm thick)
2. Cresyl violet staining of Nissl bodies (50 µm thick)
3. Unassigned section (50 µm thick)
4. Gallyas silver staining of myelin (50 µm thick)
5. Unassigned section (100 µm thick)
6. Unassigned section (50 µm thick)

Each stain was repeated every 350 µm throughout the brain. The unassigned sections were returned to formalin and stored for longevity.

The imaging of the tissue sections is very time consuming. Hence, slide digitisation is an ongoing process where the Nissl and other complementary stains will be released at a future date.

5.4.1. Polarised Light Imaging

Polarised light imaging (Figure 3) utilises the birefringence of myelinated axons to estimate the primary fibre orientation per pixel [12, 13, 14]. Here unstained tissue sections were imaged using a Leica DM4000B microscope (Leica, Germany) adapted for PLI with an LED light source, a polariser, a quarter wave plate with its fast axis at 45 degrees to the transmission axis of the polariser, and a rotatable polariser (the analyser).

Due to the large size of the BigMac tissue sections, multiple fields of view were acquired across each sample and later stitched together using bespoke software (adapted from [4]) to form a whole slide 'mosaic' image. For each field of view, images were taken as the analyser was rotated from 0 to 180 degrees in 9 equidistant steps. A 2.5x magnifying objective produced an imaging resolution of $\sim$ 4 µm per pixel. Background correction was performed [68, 4] to account for light source variations across the image, after which a sinusoid was fitted to the pixelwise image intensity as a function of the analyser rotation. Maps of transmittance, retardance, and in-plane angle were derived from the sinusoid phase and amplitude [12, 13, 14].

Supplementary Figure 5 shows example PLI mosaics from the BigMac dataset. The transmittance map is related to the amount of light extinguished by the sample. The retardance map is dependent on both the inclination and amount of birefringent material (i.e. myelinated fibres) within the PLI-pixel [12, 13, 14]. In the HSV image, the hue is dependent on the in-plane angle of the myelinated fibres, and the value is given by the tissue retardance.

By assuming that the amount of myelin is approximately constant across the white matter, an inclination angle can be estimated from the retardance map. This inclination estimate relies on knowing the myelin thickness and birefringence, which here was set to a somewhat arbitrary, constant value. Consequently, the estimated inclination angles are likely inaccurate and should not be used as a quantitative microscopy metric. In this work, the "inclination" map is used solely for MRI-microscopy co-registration.

The anterior PLI sections were mounted using a hard-set mounting medium (FluorSave, Merck) where over time we saw artefacts (bubbles) develop on the slides. This artefact is observed in the PLI transmittance image (Figure 5), though the retardance and in-plane maps do not appear to be substantially affected in the white matter apart from faintly visible edge effects (white arrows). In some anterior PLI sections we see background birefringence outside of the tissue which varies slowly across the slide (Figure 3 sections 1-5 where 1 and 3 are worst affected). This is due to the slides being coated in a small amount of gelatine which aids the mounting of tissue sections onto glass slides but which is also birefringent [69]. Nonetheless, the PLI orientations within the white matter do not appear greatly affected, where the birefringence of the myelin

appears to dominate [70]. The posterior sections (which were processed second) were instead mounted with an aqueous mounting medium (Polyvinylpyrrolidone, PVP) on plain glass slides without gelatine coating.

5.4.2. Gallyas silver staining

Gallyas silver staining [15] was used for histological visualisation of the myeloarchitecture [16] (Figure 4). In this method, colloidal silver particles bind to myelin and turn deep brown. After staining, the sections were cover-slipped, sealed and digitised using a Aperio Scanscope Turbo AT slidescanner (Leica) with a 20x/0.75 NA Plan Apo objective lens coupled with an x2 optical magnification lens to achieve a total magnification of 40x. This produced an imaging resolution of 0.28 µm/pix, where the histology image resolution is more than 10 times that of PLI. Due to the large slide size, many of the central sections were digitised in two images (labelled image 'a' and 'b').

Structure tensor analysis [34, 35, 36, 37] was applied to the digitised Gallyas images to extract the primary fibre orientation per microscopy pixel (Figure 4). Across a local neighbourhood of 150 x 150 pixels, the fibre orientations were then combined into a frequency histogram to produce a fibre orientation distribution for a $\sim$ 40 x 40 µm 'superpixel'. Summary statistics were also extracted at the level of the superpixel, where the superpixel parameters include:

1. The fibre orientation distribution: orientations within the 40 µm superpixel were combined into a frequency histogram (bin size = 2°).
2. The circular mean of the fibre orientation distribution.
3. The fibre orientation dispersion index at $\sim$ 40 µm: a Bingham distribution was fitted to the fibre orientations within the superpixel and the dispersion parameter κ was converted to the orientation dispersion index, $ODI = 2/\pi \, \mathrm{atan}(1/\kappa)$.
4. The mean RGB value over the superpixel.

Unfortunately many of posterior Gallyas silver sections exhibit a tissue processing artefact resulting in inconsistent or patchy staining (Supplementary Figure 6). This artefact is only observed in the posterior not anterior sections, and may be related to the formation of ice crystals during tissue processing. Remarkably, structure tensor analysis of slides with the staining artefact show smoothly varying orientations across the white matter that follow our neuroanatomical expectations (6b-e). Comparing structure tensor analysis of two adjacent slides, one without the artefact (6f) and one with the staining artefact (6g), we observe similar orientations, though with the artefactual slide showing reduced contrast in the grey matter (inset). As neither the PLI data nor the Nissl slides were affected by the same artefact, orientational information from either PLI or structure tensor analysis of the Nissl stained slides may be more reliable in these regions. One of the "unassigned" sets tissue sections (curently in formalin), will likely be used to repeat the Gallyas staining to obtain myelin-density estimates across the posterior brain.

5.5. Co-registration of MRI and microscopy data

The polarised light images and structure tensor output were registered to the postmortem structural MR image using TIRL (Figure 5). The resolution of the images were 4, 40 and 300 µm respectively. The structural MR image was chosen as the target image as i) it was the MR data with the highest spatial resolution and ii) it provides good grey/white matter contrast. To drive the registration, we selected the microscopy images with white/grey matter contrast most similar to the structural MR image and with the most well-defined tissue boundaries. Consequently, for the Gallyas slides we used the structure tensor RGB 'thumb' image in CIELAB or L*a*b space. The L*a*b space is based on the opponent colour model of human vision, where any given colour is represented as the combination of lightness ('L'), a position along a red-green axis ('a') and that along a blue-yellow axis ('b'). The 'b' image was used to drive the MRI-microscopy registration because is shows farly well defined tissue boundaries, that were difficult to determine in the

RGB space. For PLI we used the 'inclination' map which, when compared to the transmittance images, are relatively unaffected by the 'bubble' artefact (Supplementary Figure 5). The microscopy images were first resampled to the resolution of the structural MR data and then registered into the 3D imaging volume. TIRL outputs a series of transformations which allow the user to transform either pixel or voxel coordinates, or orientational vectors between domains. Scripts are provided to demonstrate how users can precisely map the high-resolution microscopy information to into the MRI volume, or vice versa.

The most anterior and posterior microscopy sections do not sample the corpus callosum meaning that there is no tissue directly connecting the two hemispheres. Once sectioned, the tissue from each brain hemisphere is fully disconnected and the distance between the hemispheres when mounted onto the slides is not meaningful. Therefore, each hemisphere was registered separately to the MR data with the aid of hand-drawn tissue masks. Similarly, the cerebrum was masked and registered separately to the cerebellum (ongoing work).

5.6. Comparing fibre orientation distributions from microscopy and MRI

For a qualitative comparison of fibre orientations from coregistered dMRI and PLI (Figure 6a), diffusion MRI data ($b = 4\,\mathrm{ms/\mu m^2}$, 250 gradient directions, 1mm isotropic) were processed using the diffusion tensor model (FSL's dtifit, [41]) to produce maps of fractional anisotropy, FA, and the primary eigenvector, V1. These maps were warped to PLI space using TIRL [19], and V1 was projected onto the microscopy plane for comparison with PLI.

Fibre orientation distributions from MRI and microscopy were then compared on a voxelwise basis (Figure 6b). Diffusion MRI data ($b = 10\,\mathrm{ms/\mu m^2}$, 1000 gradient directions, 1mm isotropic) were processed using the Ball and Rackets model (BAR) [42], to estimate a single, disperse fibre orientation distribution (FOD) per voxel. Following a previously published method [4], the FOD was then projected onto the microscopy plane for direct comparison with those from PLI and histology. The microscopy FODs were created by first warping the microscopy orientations to MR space. The fibre orientations (from PLI or structure tensor analysis of the Gallyas-stained slides) within each MR voxel were then combined into a frequency histogram with respect to orientation angle (resolution=2°), the output of which is shown in Figure 6b.

Finally, the orientation dispersion index (ODI, [43]) of each 2D FOD (from BAR, PLI and histology) was calculated according to [4]. An ODI of 0 indicates no dispersion, whilst an ODI of 1 describes isotropic dispersion. The ODI values were compared on to DTI FA ($b = 4\,\mathrm{ms/\mu m^2}$, 250 gradient directions, 1mm isotropic) [41] and microscopic FA from the simultaneous analysis of linear and spherical tensor encoded data [44, 45, 66, 67].

5.7. Hybrid diffusion MRI-microscopy 3D fibre reconstruction and tractography

For the hybrid orientations in Figure 7, PLI data informed on fibre orientations within the microscopy plane, whilst the diffusion data provided through plane information. This facilitated reconstruction of 3D hybrid orientations at spatial resolution of the microscopy data.

Diffusion MRI data ($b = 10\,\mathrm{ms/\mu m^2}$, 1000 gradient directions, 1mm isotropic) were analysed using the Ball and Stick (BAS) model to estimate 3 fibre populations per voxel, with 50 orientation estimates or samples per population [21, 22, 71]. The PLI images were co-registered to the diffusion MRI data using an optimised TIRL protocol [20]. The in-plane angle was warped into the diffusion space [39, 20] and compared to the BAS samples within the corresponding diffusion MRI voxel. To facilitate fair comparison, the BAS samples were projected onto the microscopy plane. Samples from BAS fibre populations with signal fractions < 0.05 were excluded. Finally, the microscopy through-plane angle was approximated by that from the most similar BAS sample. This produced a hybrid diffusion MRI-microscopy 3D fibre orientation per microscopy pixel.

The hybrid fibre orientations were then combined into 3D fibre orientation distributions (FODs). Here, a set of voxels were defined in diffusion space. In each voxel, the hybrid MRI-microscopy fibre orientations populated a 3D orientation histogram defined by 256 points evenly spaced across the sphere. Spherical harmonics of order 8 were then fitted to the normalised histogram. In spherical harmonic format, the hybrid diffusion MRI-microscopy FODs could then be visualised in standard MRI viewers [72, 57] and input into existing tractography methods [57]. Anatomically constrained streamline tractography was then performed using MRtrix (iFOD2) [57, 73] with anatomical masks adapted from XTRACT [26].

6. Acknowledgements

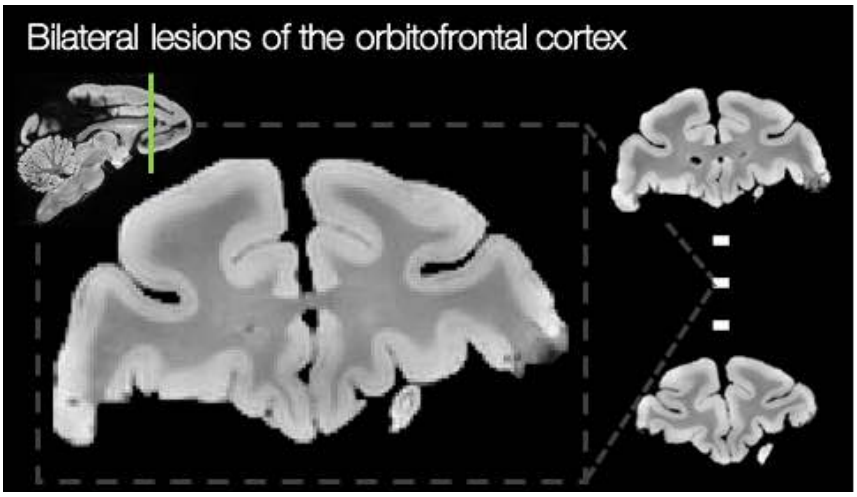
The authors would like to thank Katherine Bryant, Jesper Anderson, Frederik Lange, Moises Hernandez Fernandez, Paul McCarthy, Michiel Kleinnijenhuis and Menuka Pallebage Gamarallage for their advice at stages throughout this project. AFDH and IN were supported by the EPSRC, MRC and Wellcome Trust (grants EP/L016052/1, MR/L009013/1 and WT202788/Z/16/A). IN was also supported by the Clarendon Fund in partnership with the Chadwyck-Healey Charitable Trust at Kellogg College (Oxford). AK and NS were supported by Cancer Research UK (grant number C5255/A15935). RN was supported by the BBSRC (grant BB/N019814/1). JS was funded by ANR grant (ANR-21-CE37-0016). KLM and SJ were supported by the Wellcome Trust (grants WT202788/Z/16/A, WT215573/Z/19/Z and WT221933/Z/20/Z). The Wellcome Centre for Integrative Neuroimaging is supported by core funding from the Wellcome Trust (203139/Z/16/Z).

7. Contributions

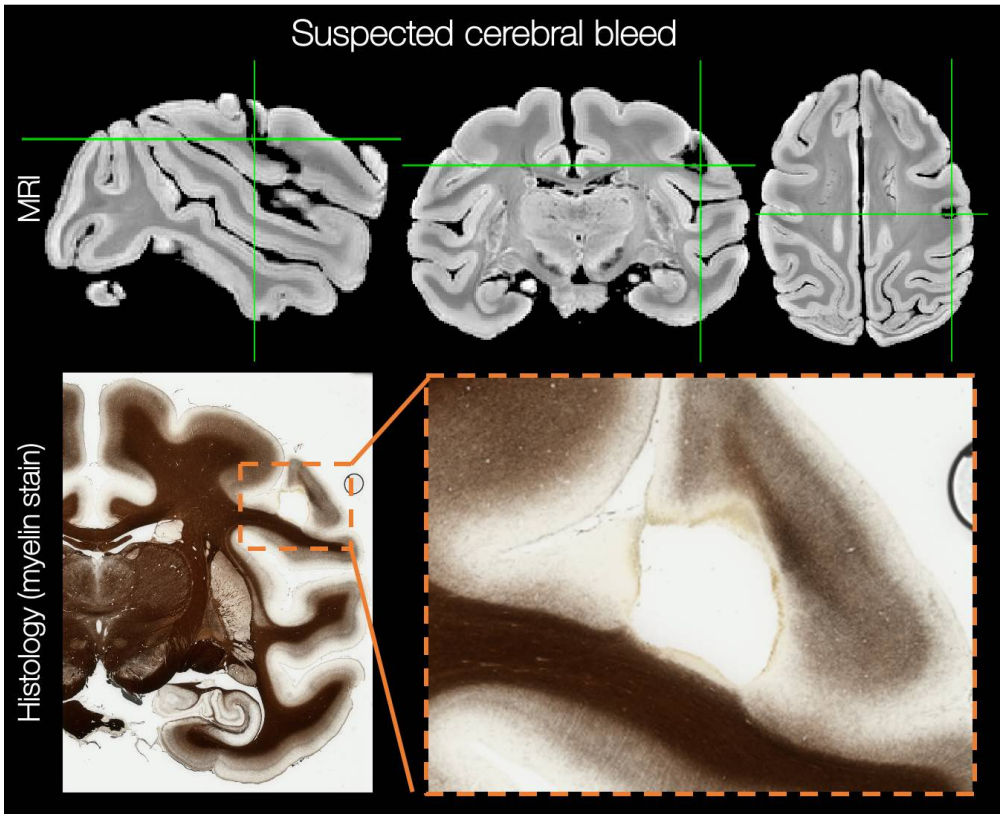
All authors reviewed the manuscript. Amy FD Howard: project lead; study conceptualisation; involved with acquisition of all postmortem MRI and microscopy data; developed acquisition protocols and processing pipelines; curated all data; conceived of and performed analyses; wrote and edited manuscript. Istvan N Huszar: developed the TIRL software and protocol for co-registration of the MRI-microscopy data. Adele Smart: acquired and processed the microscopy data, including co-registration to MRI with hand drawn masks. Michiel Cottaar: developed protocol for multiple diffusion tensor encoding. Greg Daubney: processed tissue for microscopy, including sectioning, staining and mounting. Taylor Hanayik: developed software for the Digital Brain Bank platform for downloading and viewing the data. Alexandre A Khrapitchev: optimised scanning protocols and acquired postmortem MRI. Rogier B Mars: study conceptualisation; provided high resolution postmortem MRI protocol; advised on data processing and analysis. Jeroen Mollink: developed software and protocols for polarised light imaging and structure tensor analysis. Connor Scott: facilitated slide scanning. Nicola R Sibson: facilitated postmortem MR scanning. Jerome Sallet: study conceptualisation; acquired in vivo data; provided postmortem brain & high resolution MRI protocol; facilitated MR acquisition and tissue processing for microscopy. Saad Jbabdi: study conceptualisation; advised on all acquisition protocols, processing pipelines and analyses; edited manuscript; provided resources and supervision. Karla L Miller: study conceptualisation; advised on all acquisition protocols, processing pipelines and analyses; edited manuscript; provided resources, supervision and funding.

8. Supplementary material

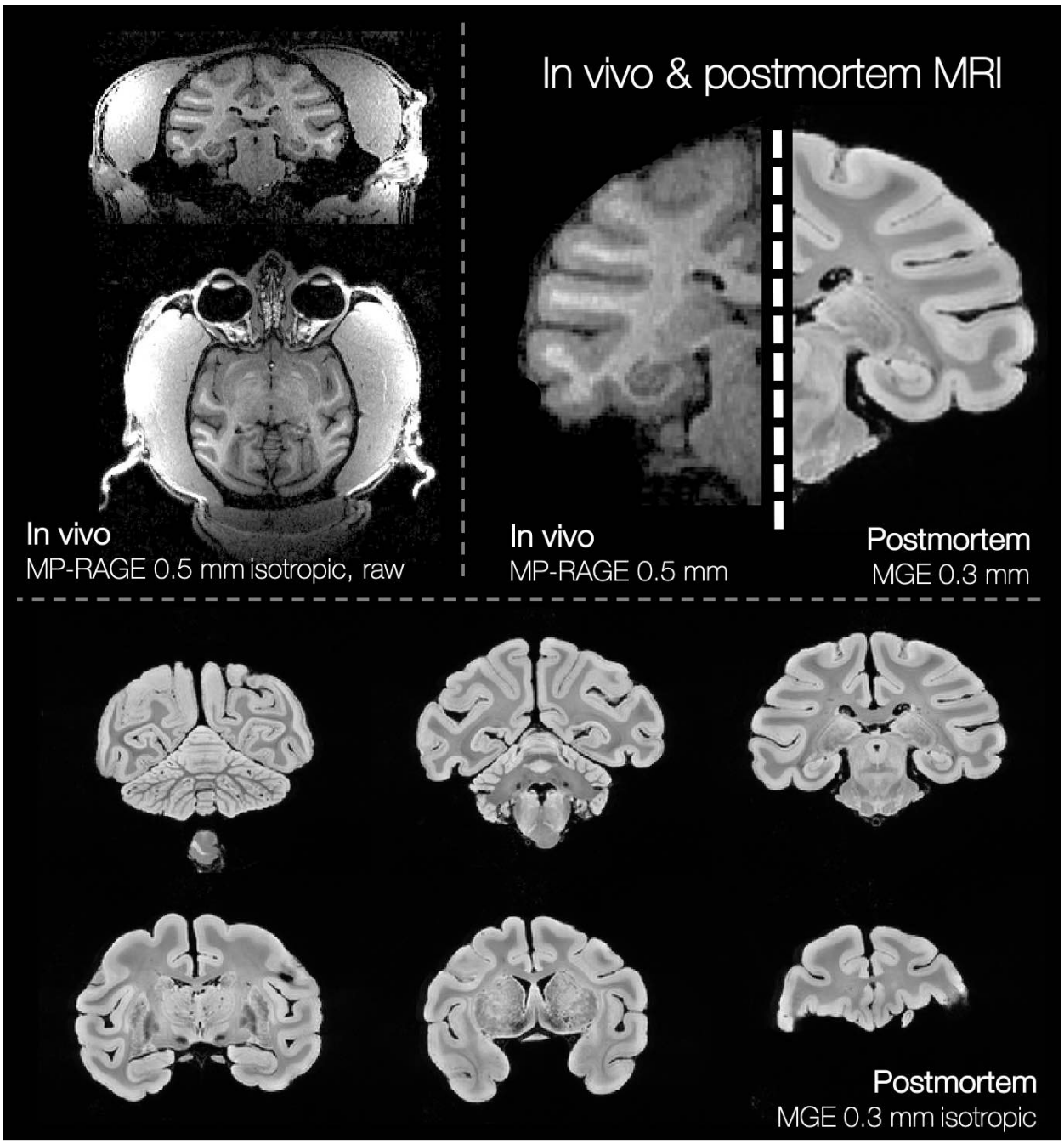
1. Miller, K.L., Alfaro-Almagro, F., Bangerter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L., Griffanti, L., Douaud, G., Okell, T.W., Weale, P., Dragonu, I., Garratt, S., Hudson, S., Collins, R., Jenkinson, M., Matthews, P.M., Smith, S.M.. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nature Neuroscience* 2016;19(11):1523–1536. doi:10.1038/nn.4393.



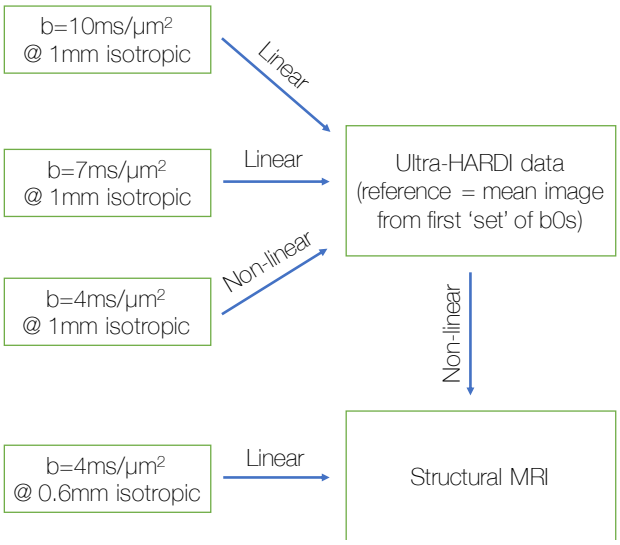
Supplementary Figure 1: Bilateral lesioning of the orbitofrontal cortex where postmortem structural MR images show the extent of the lesion. The top image is the most posterior, and the bottom the most anterior, of those shown. The lesion occurred approximately 1 year prior to the scans shown.



Supplementary Figure 2: Both the postmortem MRI (top) and histology data (bottom) show a region of tissue damage in the left hemisphere of the BigMac brain. This region is also visible in the most recent in vivo data. The tissue damage may be indicative of a postoperative cerebral bleed, though there was no indication of this at the time.

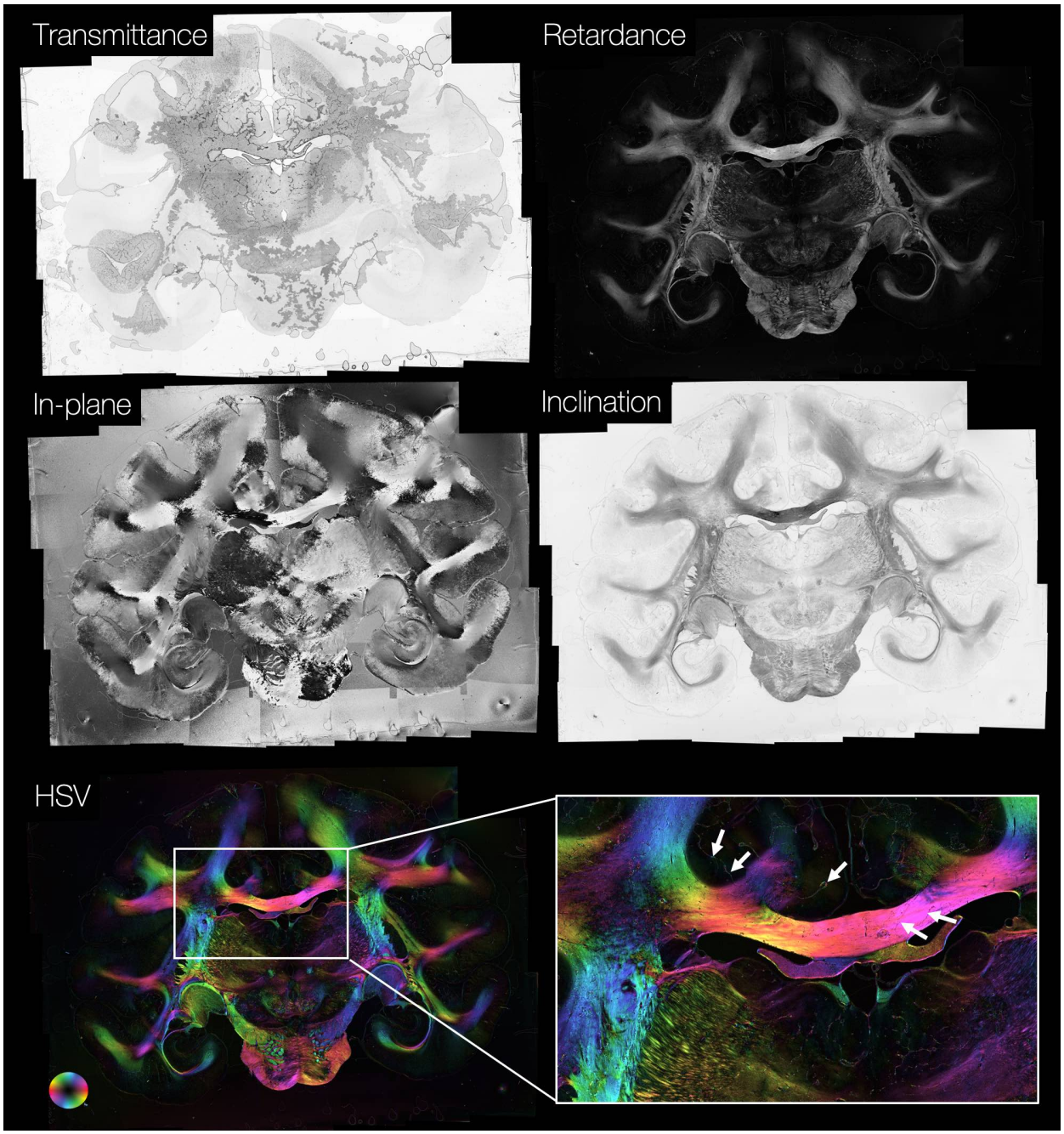


Supplementary Figure 3: The BigMac dataset includes both in vivo (structural, functional and diffusion MRI) and postmortem (structural MRI, diffusion MRI and microscopy) data in the same animal. Here we compare example structural images from both the in vivo and postmortem protocol. Though there is inverted contrast between the in vivo and postmortem structural images, the data is highly complementary. Furthermore, in the high-quality postmortem data (bottom) we can see the anatomy in detail due to the high 0.3 mm isotropic resolution. The top left image is the most posterior and the bottom right the most anterior, where we again see the bilateral lesion of the orbitofrontal cortex.

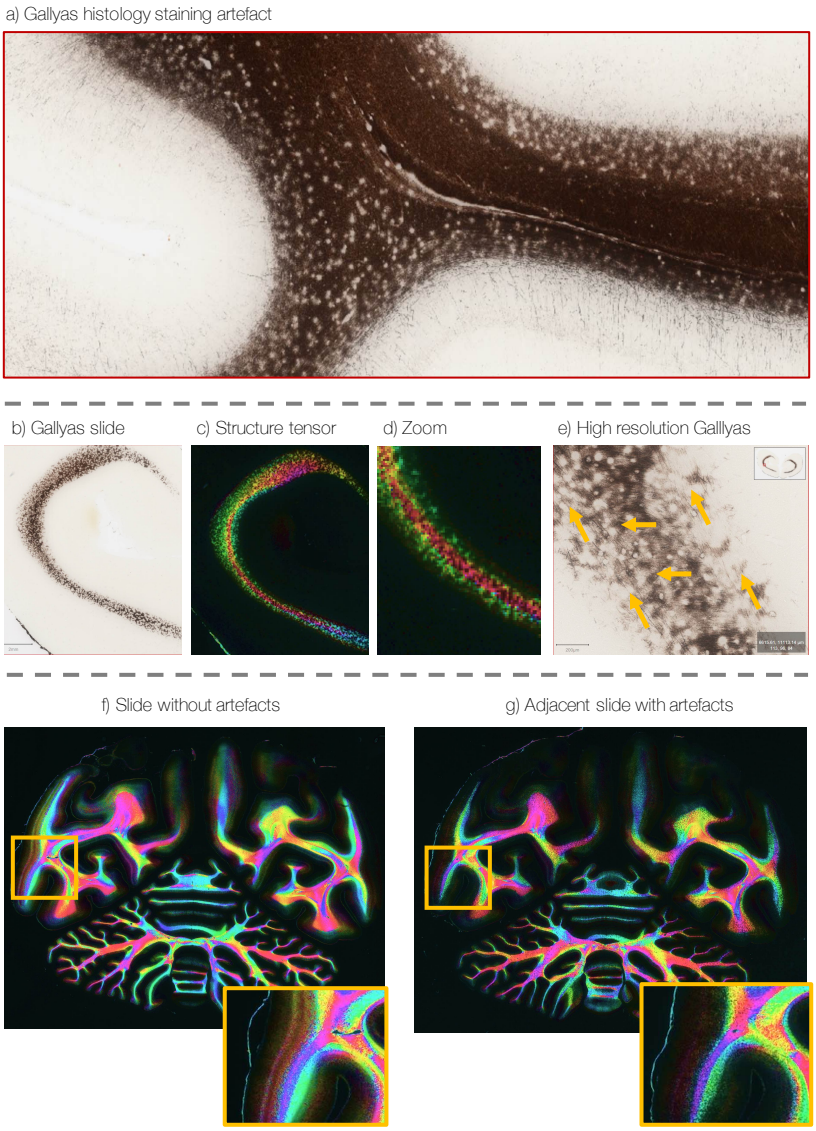


Supplementary Figure 4: Co-registration of the BigMac postmortem MRI data.

2. Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K.. The WU-Minn Human Connectome Project: An overview. *NeuroImage* 2013;80:62–79. doi:10.1016/j.neuroimage.2013.05.041.
3. Poldrack, R.A., Laumann, T.O., Koyejo, O., Gregory, B., Hover, A., Chen, M.Y., Gorgolewski, K.J., Luci, J., Joo, S.J., Boyd, R.L., Hunicke-Smith, S., Simpson, Z.B., Caven, T., Sochat, V., Shine, J.M., Gordon, E., Snyder, A.Z., Adeyemo, B., Petersen, S.E., Glahn, D.C., Mckay, D.R., Curran, J.E., Göring, H.H., Carless, M.A., Blangero, J., Dougherty, R., Leemans, A., Handwerker, D.A., Frick, L., Marcotte, E.M., Mumford, J.A.. Long-term neural and physiological phenotyping of a single human. *Nature Communications* 2015;6(1):1–15. doi:10.1038/ncomms9885.
4. Mollink, J., Kleinnijenhuis, M., van Cappellen van Walsum, A.M., Sotiropoulos, S.N., Cottaar, M., Mirfin, C., Heinrich, M.P., Jenkinson, M., Pallebage-Gamarallage, M., Ansorge, O., Jbabdi, S., Miller, K.L.. Evaluating fibre orientation dispersion in white matter: Comparison of diffusion MRI, histology and polarized light imaging. *NeuroImage* 2017;157:561–574. doi:10.1016/j.neuroimage.2017.06.001.
5. Pallebage-Gamarallage, M., Foxley, S., Menke, R.A., Huszar, I.N., Jenkinson, M., Tendler, B.C., Wang, C., Jbabdi, S., Turner, M.R., Miller, K.L., Ansorge, O.. Dissecting the pathobiology of altered MRI signal in amyotrophic lateral sclerosis: A post mortem whole brain sampling strategy for the integration of ultra-high-field MRI and quantitative neuropathology. *BMC Neuroscience* 2018;19(1). doi:10.1186/S12868-018-0416-1.
6. Tendler, B.C., Hanayik, T., Ansorge, O., Bangerter-Christensen, S., Berns, G.S., Bertelsen, M.F., Bryant, K.L., Foxley, S., Martijn van den Heuvel,, P., Howard, A.F., Huszar, I.N., Khrapitchev, A.A., Leonte, A., Manger, P.R., Menke, R.A., Mollink, J., Mortimer, D., Pallebage-Gamarallage, M., Roumazeilles, L., Sallet, J., Scholtens, L.H., Scott, C., Smart, A., Turner, M.R., Wang, C., Jbabdi, S., Mars, R.B., Miller, K.L.. The Digital Brain Bank, an open access platform for post-mortem datasets. *eLife* 2022;11. doi:10.7554/ELIFE.73153.
7. Amunts, K., Lepage, C., Borgeat, L., Mohlberg, H., Dickscheid, T., Rousseau, M.É., Bludau, S., Bazin, P.L., Lewis, L.B., Oros-Peusquens, A.M., Shah, N.J., Lippert, T., Zilles, K., Evans, A.C.. BigBrain: An ultrahigh-resolution 3D human brain model. *Science* 2013;340(6139):1472–1475. doi:10.1126/science.1235381.
8. Jelescu, I.O., Budde, M.D.. Design and Validation of Diffusion MRI Models of White Matter. *Frontiers in Physics* 2017;5:61. doi:10.3389/fphy.2017.00061.
9. Yendiki, A., Aggarwal, M., Axer, M., Howard, A.F., van Walsum, A.M.v.C., Haber, S.N.. Post mortem mapping of connectional anatomy for the validation of diffusion MRI. *NeuroImage* 2022;:119146doi:10.1016/J.NEUROIMAGE.2022.119146.
10. Howard, A., Jbabdi, S., AA, K., Sallet, J., Daubney, G., Mollink, J., Scott, C., Sibson, N., Miller, K.. The BigMac dataset: ultra-high angular resolution diffusion imaging and multi-contrast microscopy of a whole macaque brain. *ISMRM 27th Annual Meeting* 2019;.
11. Yendiki, A., Jones, R., Dalca, A., Wang, H., Fischl, B.. Towards taking the guesswork (and the errors) out of diffusion tractography. *ISMRM 28th Annual Meeting* 2020;.
12. Axer, H., Axerl, M., Krings, T., Keyserlingk, D.G.v.. Quantitative estimation of 3-D fiber course in gross histological sections of the human brain using polarized light. *Journal of Neuroscience Methods* 2001;105(2):121–31. doi:10.1016/S0165-0270(00)00349-6.
13. Larsen, L., Griffin, L.D., Gräßel, D., Witte, O.W., Axer, H.. Polarized light imaging of white matter architecture. *Microscopy Research and Technique* 2007;70(10):851–863. doi:10.1002/jemt.20488.
14. Axer, M., Grässel, D., Kleiner, M., Dammers, J., Dickscheid, T., Reckfort, J., Hütz, T., Eiben, B., Pietrzyk, U.,



Supplementary Figure 5: Example transmittance, retardance, in-plane and inclination maps from PLI. The in-plane and retardance images are combined in a composite HSV image. Though the slide is badly affected by bubbles and artefacts (see transmittance image), the artefacts appear to have little effect on the other PLI maps (retardance, in-plane and inclination). Consequently, only very minor artefacts are seen in the HSV image (white arrows). Whole slide PLI were acquired for > 160 sections throughout the BigMac brain.



Supplementary Figure 6: Artefacts seen in the posterior Gallyas-stained slides. a) The myelin stain appears patchy in the white matter with relatively little stain density in the grey matter. b-e) Structure tensor analysis of the slides with the staining artefact produces remarkably reasonable fibre orientations, particularly in the white matter. The orientations are smoothly varying and we are able to delineate the presence of multiple fibre bundles. f-g) When comparing structure tensor analysis of adjacent slides, one with and the other without the artefact, we see very similar orientations within the white matter, but reduced grey matter contrast in the slide with staining artefact.

Zilles, K., Amunts, K.. High-Resolution Fiber Tract Reconstruction in the Human Brain by Means of Three-Dimensional Polarized Light Imaging. *Frontiers in Neuroinformatics* 2011;5:34. doi:10.3389/fninf.2011.00034.
15. Gallyas, F.. Silver staining of Alzheimer's neurofibrillary changes by means of physical development. *Acta Morphologica Academiae Scientiarum Hungaricae* 1971;19(1):1–8.
16. Uchihara, T.. Silver diagnosis in neuropathology: Principles, practice and revised interpretation. *Acta Neuropathologica* 2007;113(5):483–499. doi:10.1007/s00401-007-0200-2.
17. Grohn, J., Schüffelgen, U., Neubert, F.X., Bongioanni, A., Verhagen, L., Sallet, J., Kolling, N., Rushworth, M.F.. Multiple systems in macaques for tracking prediction errors and other types of surprise. *PLOS Biology* 2020;18(10):e3000899. doi:10.1371/JOURNAL.PBIO.3000899.
18. Sallet, J., Noonan, M.A.P., Thomas, A., O'Reilly, J.X., Anderson, J., Papageorgiou, G.K., Neubert, F.X., Ahmed, B., Smith, J., Bell, A.H., Buckley, M.J., Roumazeilles, L., Cuell, S., Walton, M.E., Krug, K., Mars, R.B., Rushworth, M.F.. Behavioral flexibility is associated with changes in structure and function distributed across a frontal cortical network in macaques. *PLoS Biology* 2020;18(5):e3000605. doi:10.1371/journal.pbio.3000605.
19. Huszar, I., Pallebage-Gamarallage, M., Tendler, B.C., Foxley, S., Heinrich, M.P., Turner, M.R., Ansorge, O., Miller, K.L., Jenkinson, M.. Automated slice-to-volume registration between histology and whole-brain post-mortem MRI. In: *Intl. Soc. Mag. Reson. Med. 27*. 2019:.
20. Huszar, I.N., Pallebage-Gamarallage, M., Foxley, S., Tendler, B.C., Leonte, A., Hiemstra, M., Mollink, J., Smart, A., Bangerter-Christensen, S., Brooks, H., Turner, M.R., Ansorge, O., Miller, K.L., Jenkinson, M.. Tensor Image Registration Library: Automated Non-Linear Registration of Sparsely Sampled Histological Specimens to Post-Mortem MRI of the Whole Human Brain. *bioRxiv* 2019;:849570doi:10.1101/849570.
21. Behrens, T.E.J., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., Clare, S., Matthews, P.M., Brady, J.M., Smith, S.M.. Characterization and Propagation of Uncertainty in Diffusion-Weighted MR Imaging. *Magnetic Resonance in Medicine* 2003;50(5):1077–1088. doi:10.1002/mrm.10609.
22. Behrens, T.E.J., Berg, H.J., Jbabdi, S., Rushworth, M.F.S., Woolrich, M.W.. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage* 2007;34(1):144–155. doi:10.1016/j.neuroimage.2006.09.018.
23. Kötter, R., Wanke, E.. Mapping brains without coordinates. *Philosophical Transactions of the Royal Society B: Biological Sciences* 2005;360(1456):751–766. doi:10.1098/RSTB.2005.1625.
24. Shen, K., Bezgin, G., Matthew Hutchison, R., Gati, J.S., Menon, R.S., Everling, S., McIntosh, A.R.. Information Processing Architecture of Functionally Defined Clusters in the Macaque Cortex. *Journal of Neuroscience* 2012;32(48):17465–17476. doi:10.1523/JNEUROSCI.2709-12.2012.
25. Shen, K., Bezgin, G., Schirner, M., Ritter, P., Everling, S., McIntosh, A.R.. A macaque connectome for large-scale network simulations in TheVirtualBrain. *Scientific Data 2019 6:1* 2019;6(1):1–12. doi:10.1038/s41597-019-0129-z.
26. Warrington, S., Bryant, K., Khrapitchev, A., Sallet, J., Charquero-Ballester, M., Douaud, G., Jbabdi, S., Mars, R., Sotiropoulos, S.. XTRACT - Standardised protocols for automated tractography in the human and macaque brain. *NeuroImage* 2020;217:116923. doi:10.1016/j.neuroimage.2020.116923.
27. Palay, S.L., Chan-Palay, V.. Cerebellar Cortex: Cytology and Organization. Springer Berlin Heidelberg; 1974. doi:10.1007/978-3-642-65581-4.
28. Koike-Tani, M., Tani, T., Mehta, S.B., Verma, A., Oldenbourg, R.. Polarized light microscopy in reproductive and developmental biology. *Molecular Reproduction and Development* 2015;82(7-8):548–562. doi:10.1002/MRD.22221.
29. Schmitt, F.O., Bear, R.S.. The Ultrastructure of the Nerve Myelin Sheath. *Biological Reviews* 1939;14(1):27–50. doi:10.1111/J.1469-185X.1939.TB00922.X.
30. Huang, X.R., Knighton, R.W.. Microtubules Contribute to the Birefringence of the Retinal Nerve Fiber Layer. *Investigative Ophthalmology & Visual Science* 2005;46(12):4588–4593. doi:10.1167/IOVS.05-0532.
31. Lange, W.. The myelinated parallel fibers of the cerebellar cortex and their regional distribution. *Cell and Tissue Research* 1976;166:489–496. doi:10.1007/BF00225913.
32. Wyatt, K.D., Tanapat, P., Wang, S.S.. Speed limits in the cerebellum: Constraints from myelinated and unmyelinated parallel fibers. *European Journal of Neuroscience* 2005;28(1):2285–2290. doi:10.1111/j.1460-9568.2005.04053.x.
33. Koike-Tani, M., Tominaga, T., Oldenbourg, R., Tani, T.. Birefringence Changes of Dendrites in Mouse Hippocampal Slices Revealed with Polarizing Microscopy. *Biophysical Journal* 2020;118(10):2366–2384. doi:10.1016/J.BPJ.2020.03.016.
34. Bigun, J., Bigun, T., Nilsson, K.. Recognition by symmetry derivatives and the generalized structure tensor. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2004;26(12):1590–1605. doi:10.1109/tpami.2004.126.
35. Budde, M.D., Frank, J.A.. Examining brain microstructure using structure tensor analysis of histological sections. *NeuroImage* 2012;63(1):1–10. doi:10.1016/j.neuroimage.2012.06.042.
36. Budde, M.D., Annese, J.. Quantification of anisotropy and fiber orientation in human brain histological sections. *Frontiers in Integrative Neuroscience* 2013;7:3. doi:10.3389/fnint.2013.00003.
37. Seehaus, A., Roebroeck, A., Bastiani, M., Fonseca, L., Bratzke, H., Lori, N., Vilanova, A., Goebel, R., Galuske, R.. Histological validation of high-resolution DTI in human post mortem tissue. *Frontiers in Neuroanatomy* 2015;9:1–12. doi:10.3389/fnana.2015.00098.
38. Jenkinson, M., Bannister, P., Brady, M., Smith, S.. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage* 2002;17(2):825–841. doi:10.1016/S1053-8119(02)91132-8.
39. Andersson, J.L.R., Jenkinson, M., Smith, S.. Non-linear registration aka spatial normalisation. Tech. Rep. June; 2007.
40. Heinrich, M.P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F.V., Brady, S.M., Schnabel, J.A.. MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration. *Medical Image Analysis* 2012;16(7):1423–1435. doi:10.1016/j.media.2012.05.008.
41. Basser, P.J., Mattiello, J., LeBihan, D.. MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*

1994;66(1):259–267. doi:10.1016/S0006-3495(94)80775-1.
42. Sotiropoulos, S.N., Behrens, T.E., Jbabdi, S.. Ball and rackets: Inferring fiber fanning from diffusion-weighted MRI. *NeuroImage* 2012;60(2):1412–1425. doi:10.1016/j.neuroimage.2012.01.056.
43. Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., Alexander, D.C.. NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage* 2012;61(4):1000–1016. doi:10.1016/j.neuroimage.2012.03.072.
44. Lasič, S., Szczepankiewicz, F., Eriksson, S., Nilsson, M., Topgaard, D.. Microanisotropy imaging: quantification of microscopic diffusion anisotropy and orientational order parameter by diffusion MRI with magic-angle spinning of the q-vector. *Frontiers in Physics* 2014;2:11. doi:10.3389/fphy.2014.00011.
45. Szczepankiewicz, F., Lasič, S., van Westen, D., Sundgren, P.C., Englund, E., Westin, C.F., Ståhlberg, F., Lätt, J., Topgaard, D., Nilsson, M.. Quantification of microscopic diffusion anisotropy disentangles effects of orientation dispersion from microstructure: Applications in healthy volunteers and in brain tumors. *NeuroImage* 2015;104:241–252. doi:10.1016/j.neuroimage.2014.09.057.
46. Kor, D.Z., Jbabdi, S., Huszar, I.N., Mollink, J., Tendler, B.C., Foxley, S., Wang, C., Scott, C., Smart, A., Ansorge, O., Pallebage-Gamarallage, M., Miller, K.L., Howard, A.F.. An automated pipeline for extracting quantitative histological metrics for voxelwise MRI-histology comparisons. *bioRxiv* 2022;doi:10.1101/2022.02.10.479718.
47. Sallet, J., Mars, R.B., Noonan, M.P., Andersson, J.L., O'Reilly, J.X., Jbabdi, S., Croxson, P.L., Jenkinson, M., Miller, K.L., Rushworth, M.F.. Social network size affects neural circuits in Macaques. *Science* 2011;334(6056):697–700. doi:10.1126/SCIENCE.1210027.
48. Mars, R.B., Jbabdi, S., Sallet, J., O'Reilly, J.X., Croxson, P.L., Olivier, E., Noonan, M.A.P., Bergmann, C., Mitchell, A.S., Baxter, M.G., Behrens, T.E., Johansen-Berg, H., Tomassini, V., Miller, K.L., Rushworth, M.F.. Diffusion-Weighted Imaging Tractography-Based Parcellation of the Human Parietal Cortex and Comparison with Human and Macaque Resting-State Functional Connectivity. *Journal of Neuroscience* 2011;31(11):4087–4100. doi:10.1523/JNEUROSCI.5102-10.2011.
49. Noonan, M.A.P., Sallet, J., Mars, R.B., Neubert, F.X., O'Reilly, J.X., Andersson, J.L., Mitchell, A.S., Bell, A.H., Miller, K.L., Rushworth, M.F.. A Neural Circuit Covarying with Social Hierarchy in Macaques. *PLOS Biology* 2014;12(9):e1001940. doi:10.1371/JOURNAL.PBIO.1001940.
50. Mars, R.B., Verhagen, L., Gladwin, T.E., Neubert, F.X., Sallet, J., Rushworth, M.F.S.. Comparing brains by matching connectivity profiles. *Neuroscience and Biobehavioral Reviews* 2016;60:90–97. doi:10.1016/j.neubiorev.2015.10.008.
51. Mars, R.B., Jbabdi, S., Rushworth, M.F.. A Common Space Approach to Comparative Neuroscience. *https://doiorg/101146/annurev-neuro-100220-025942* 2021;44:69–86. doi:10.1146/ANNUREV-NEURO-100220-025942.
52. Dyrby, T.B., Baaré, W.F., Alexander, D.C., Jelsing, J., Garde, E., Søgaard, L.V.. An ex vivo imaging pipeline for producing high-quality and high-resolution diffusion-weighted imaging datasets. *Human Brain Mapping* 2011;32:544–563. doi:10.1002/hbm.21043.
53. Shepherd, T.M., Thelwall, P.E., Stanisz, G.J., Blackband, S.J.. Aldehyde fixative solutions alter the water relaxation and diffusion properties of nervous tissue. *Magnetic Resonance in Medicine* 2009;62(1):26–34. doi:10.1002/mrm.21977.
54. Le Bihan, D., Delannoy, J., Levin, R.L.. Temperature mapping with MR imaging of molecular diffusion: application to hyperthermia. *Radiology* 1989;doi:10.1148/radiology.171.3.2717764.
55. Kellner, E., Dhital, B., Kiselev, V.G., Reisert, M.. Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magnetic Resonance in Medicine* 2016;76(5):1574–1581. doi:10.1002/MRM.26054. `1501.07758`.
56. Bautista, T., O'Muircheartaigh, J., Hajnal, J.V., Tournier, J.D.. Removal of Gibbs ringing artefacts for 3D acquisitions using subvoxel shifts. *Proceedings of the International Society for Magnetic Resonance in Medicine* 2021;.
57. Tournier, J.D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.H., Connelly, A.. MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage* 2019;202:116137. doi:10.1016/j.neuroimage.2019.116137.
58. Zhang, Y., Brady, M., Smith, S.. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging* 2001;20(1):45–57. doi:10.1109/42.906424.
59. Barral, J.K., Gudmundson, E., Stikov, N., Etezadi-Amoli, M., Stoica, P., Nishimura, D.G.. A robust methodology for in vivo T1 mapping. *Magnetic Resonance in Medicine* 2010;64(4):1057–1067. doi:10.1002/MRM.22497.
60. Karakuzu, A., Boudreau, M., Duval, T., Boshkovski, T., Leppert, I.R., Cabana, J.F., Gagnon, I., Beliveau, P., Pike, G.B., Cohen-Adad, J., Stikov, N.. qMRLab: Quantitative MRI analysis, under one umbrella. *Journal of Open Source Software* 2020;5(53):2343. doi:10.21105/JOSS.02343.
61. Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M.. Fsl. *NeuroImage* 2012;62(2):782–790. doi:10.1016/j.neuroimage.2011.09.015.
62. Cook, P.A., Symms, M., Boulby, P.A., Alexander, D.C.. Optimal acquisition orders of diffusion-weighted MRI measurements. *Journal of Magnetic Resonance Imaging* 2007;25(5):1051–1058. doi:10.1002/jmri.20905.
63. Jenkinson, M., Smith, S.. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis* 2001;5(2):143–156. doi:10.1016/S1361-8415(01)00036-6.
64. Van Essen, D.C.. Surface-based atlases of cerebellar cortex in the human, macaque, and mouse. In: *Annals of the New York Academy of Sciences*; vol. 978. New York Academy of Sciences; 2002:468–479. doi:10.1111/j.1749-6632.2002.tb07588.x.
65. Sjölund, J., Szczepankiewicz, F., Nilsson, M., Topgaard, D., Westin, C.F., Knutsson, H.. Constrained optimization of gradient waveforms for generalized diffusion encoding. *Journal of Magnetic Resonance* 2015;261:157–168. doi:10.1016/J.JMR.2015.10.012.
66. Szczepankiewicz, F., Sjölund, J., Ståhlberg, F., Lätt, J., Nilsson, M.. Tensor-valued diffusion encoding for diffusional variance decomposition (DIVIDE): Technical feasibility in clinical MRI systems. *PLOS ONE* 2019;14(3):e0214238. doi:

10.1371/JOURNAL.PONE.0214238.
67. Nilsson, M., Szczepankiewicz, F., Lampinen, B., Ahlgren, A., de Almeida Martins, J.P., Lasič, S., Westin, C., Topgaard, D.. An open-source framework for analysis of multidimensional diffusion MRI data implemented in MATLAB. 2018.
68. Dammers, J., Axer, M., Grassel, D., Palm, C., Zilles, K., Amunts, K., Pietrzyk, U.. Signal enhancement in polarized light imaging by means of independent component analysis. *NeuroImage* 2010;49(2):1241–1248. doi:10.1016/j.neuroimage.2009.08.059.
69. Tobolsky, A.V.. Stress relaxation, birefringence and the structure of gelatin and other polymeric gels. *Journal of Physical Chemistry* 1955;59(6):575. doi:10.1021/j150528a026.
70. Menzel, M., Michielsen, K., De Raedt, H., Reckfort, J., Amunts, K., Axer, M.. A Jones matrix formalism for simulating three-dimensional polarized light imaging of brain tissue. *Journal of the Royal Society Interface* 2015;12(111):20150734. doi:10.1098/rsif.2015.0734.
71. Hernandez-Fernandez, M., Reguly, I., Jbabdi, S., Giles, M., Smith, S., Sotiropoulos, S.N.. Using GPUs to accelerate computational diffusion MRI: From microstructure estimation to tractography and connectomes. *NeuroImage* 2019;188:598–615. doi:10.1016/j.neuroimage.2018.12.015.
72. McCarthy, P.. FSLeyes. *Zenodo* 2020;doi:10.5281/ZENODO.3858136.
73. Tournier, J.D., Calamante, F., Connelly, A.. Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions. 2010.